Evaluating the Transition to College Mathematics Course in Texas High Schools: Examining Heterogeneity across Schools and Student Characteristics

James E. Pustejovsky[1] & Megha Joshi[2]

[1] University of Wisconsin - Madison

[2] University of Texas at Austin

April 20, 2021

Author Note

This research was conducted while Pustejovsky was a faculty member at the University of Texas at Austin. The project was supported through a Greater Texas Faculty Fellows grant from Greater Texas Foundation, with additional support from the Charles A. Dana Center and the Educational Psychology Department of the University of Texas at Austin. The opinions expressed are those of the authors and do not represent the views of the Greater Texas Foundation, the Charles A. Dana Center, the University of Texas at Austin, or the University of Wisconsin - Madison.

Correspondence concerning this article should be addressed to James E. Pustejovsky, 1082C Educational Sciences, 1025 West Johnson St., Madison, WI 53706. E-mail: pustejovsky@wisc.edu

Abstract

Texas House Bill 5 introduced requirements that school districts partner with institutions of higher education to provide college preparatory courses in mathematics and English language arts for high school seniors who are not yet college ready. In response to House Bill 5 requirements, the Charles A. Dana Center developed a model college preparatory mathematics course, Transition to College Mathematics Course (TCMC), which has been adopted by dozens of school districts across Texas over the past several school years. In prior work, we examined the effects of TCMC on students' progress into post-secondary education by comparing two student cohorts who participated in TCMC to observationally similar students from the same cohort but who did not enroll in the course. In this report, we investigate the extent of heterogeneity in the effects of participating in TCMC. We find little evidence that the program was differentially effective for students from different socio-economic backgrounds, nor do we find evidence that program effects varied by the number of years that it had been offered. However, for key outcomes such as rates of passing a college-level math course, the effects of participating in TCMC may have varied across the schools that offered the course. Just as in prior work, these findings must be interpreted cautiously because we were unable to fully assess and account for students' college-readiness status at the start of their senior year.

Evaluating the Transition to College Mathematics Course in Texas High Schools: Examining Heterogeneity across Schools and Student Characteristics

Introduction

Post-secondary education has become a gateway for economic and social mobility in the United States. At both the federal and state levels, policy makers are increasingly focused on improving student degree completion. In pursuing this goal, one challenge is that many students exit high school without adequate preparation for college-level coursework. Under-preparation is a particular problem in mathematics, where college-level courses are often sequenced through rigid pre-requisites. Post-secondary institutions offer developmental coursework to address deficits in student preparation, but a growing body of evidence indicates that traditional developmental coursework can, contrary to its apparent purpose, create barriers to pursuing college-level coursework and making progress towards a degree.

In Texas, the state legislature responded to these problems by introducing changes to curriculum and graduation requirements in the form of Texas House Bill 5 (HB5) from the 2013 legislative session. Among its provisions, HB5 required school districts to offer a college preparatory mathematics course for students not meeting college readiness standards in mathematics by the end of their third year of high school. This course has to be offered through a partnership with a community college or other institution of higher education. HB5 further stipulated that successful completion of the college preparatory course had to satisfy the partner institution's requirements for entry into college-level mathematics coursework, so that students would not have to later pass proficiency exams.

In response to the goals and requirements of HB5, the Charles A. Dana Center designed a model college preparatory math course, called the Transition to College Mathematics Course (TCMC). The Dana Center developed TCMC by melding previously developed secondary-level course materials with strategies they had used to build college-level developmental courses. TCMC differs from conventional remedial math courses in several important respects. First, the course content aligns with the multiple mathematics pathways framework (Charles A. Dana Center, 2016, 2019) adopted by many Texas higher education institutions, providing a coherent sequence of work for students as they transition from high school to higher education.[1] Second, the course involves novel material and instructional strategies, rather than repetition of content that students have already encountered. Third, the course incorporates evidence-based pedagogical approaches including putting greater emphasis on richly contextualized applications, developing students' self-regulated learning strategies and productive persistence, and varying instructional activities. Taken together, these differences provide reason to expect that student participation in TCMC could have immediate and longer-term impacts on student outcomes.

TCMC was offered at high schools in eight districts across central Texas during the 2016-17 school year and in thirty districts during the 2017-18 school year. In previous research (Pustejovsky & Joshi, 2019, 2020), we aimed to assess the average effects of participating in TCMC by comparing students who participated in TCMC to observationally similar students who did not enroll in the course. We applied this evaluation strategy with two cohorts of students who participated in TCMC during the 2016-17 or 2017-18 school years (the first and second years of implementation). From analysis of the first cohort of students, we found that those who took TCMC graduated at slightly higher rates than comparison students, but had lower rates of enrollment in post-secondary education, driven by lower rates of enrollment in 4-year colleges or universities (Pustejovsky & Joshi, 2019). We also found that, relative to students in the comparison groups, students who took TCMC were less likely to pass college-level and

[1] The TCMC materials implemented reflect the student learning outcomes contained in the HB5 College Preparatory Mathematics Content Framework, the result of a collaborative process organized by the Texas Success Center (TSC) at the Texas Association of Community Colleges (TACC). A committee consisting of content experts from both Higher Education and K-12 educators worked collaboratively to develop and revise this framework, seeking input from educators across the state. This feedback, along with the Learning Outcomes in the Lower-Division Academic Course Guide Manual (ACGM), the Texas College and Career Readiness Standards (CCRS), and other relevant materials informed the committee's work.

developmental math courses. In analysis of the second cohort of students, we again found that those who took TCMC graduated at higher rates than comparison students (Pustejovsky & Joshi, 2020). They had similar rates of overall enrollment in post-secondary education, but enrolled in community colleges at higher rates and in 4-year colleges or universities at lower rates than did comparison students. Relative to comparison students, second-cohort students who took TCMC were also less likely to take and less likely to pass college-level math coursework. However, results for both cohorts must be interpreted cautiously because we were unable to fully assess and account for students' college-readiness status at the start of their senior year.

Our previous evaluation work on TCMC was focused on understanding the *average* effects of participating in TCMC for the set of students who enrolled in the course during the first two years of implementation. In this report, we move beyond our initial focus in two ways. First, we investigate the proximal consequences of enrolling in the course in order to more fully characterize the difference in the educational experiences of students who took the course versus students in the comparison group. We pose two specific research questions:

1. What fraction of students who enrolled in TCMC passed the course, and how did passing rates vary across participating schools? Passing the course is an initial step that is necessary in order for students to bypass proficiency exams if they pursue post-secondary education. Thus, course passage rates are an important indicator of the extent to which participating in TCMC might affect college-level course enrollment and course passage.

2. What, if any, mathematics courses did students in the comparison group take during their senior year? The course-taking patterns of the comparison group are relevant to understanding the courses that students who enrolled in TCMC might *otherwise* have taken, had TCMC not been available.

Second, moving beyond our initial focus on average effects of participating in TCMC, we investigate the extent to which the effects of participating in TCMC may have *varied* across individuals and contexts. We pose three specific research questions:

3. To what extent did students from different socio-economic status (SES) levels benefit differentially from participation in TCMC? This question is important to consider due to its equity implications, as well as because prior research has indicated that related forms of intervention, such as early college coursework, might have stronger effects for students from lower SES backgrounds (An, 2013).

4. Did the effects of participating in TCMC vary depending on whether the school was in the first or second year of implementing the program? To the extent that differences exist, it would suggest that the effects of TCMC might change as schools and teachers gain experience implementing the program.

5. Finally, to what extent did the effects of participating in TCMC vary across schools that offered the program? Cross-site variation in effects indicates that some conditions or circumstances influence whether the program has its intended impacts (Bloom, Raudenbush, Weiss, & Porter, 2017; Raudenbush & Bloom, 2015). For instance, the characteristics of teachers, the demographic composition of a school, or the proximity of the implementing school to the partner community college might influence whether students who participate in TCMC are more likely than students who do not participate to subsequently enroll in and pass college-level math coursework. The presence of meaningful cross-site variation would point towards a need for further investigation of explanatory factors.

Just as in our previous work, we seek to identify the effects of TCMC by comparing students who took the course to a group of students who did not enroll in TCMC, but who closely resemble students who took the course during the 2017-18 school year.

Methods

During the 2017-18 school year, students' decisions to enroll in TCMC did not follow any systematic rule across schools. Schools were provided guidance that the course was intended for students who were not yet college ready at the beginning of twelfth grade, but schools did not use a single approach for determining college readiness. Rather, they relied on college readiness tests, standardized test scores, grades, or some combination of those data sources. Furthermore, advising practices differed from school to school, and so it is reasonable to infer that enrollment in TCMC was subject to individual discretion and idiosyncratic factors. Most students who enrolled in the course were in 12th grade.

As in our previous work (Pustejovsky & Joshi, 2019, 2020), we sought to evaluate the effects of participating in TCMC by identifying a comparison group of students who did not enroll in TCMC, but who were observationally similar to the group of students who did enroll in the course. By observationally similar, we mean that students in the comparison group had a distribution of background characteristics similar to that of students who enrolled in the course. Ideally, the two groups would be similar on the set of any background characteristics that could have influenced students' enrollment and that may have been associated with subsequent outcomes. If we succeeded in creating groups that have very similar distributions for all such confounding variables, then any differences in outcomes between the groups can be interpreted as the impact of actually participating in TCMC for students who chose to enroll in it.

For this approach to provide valid evidence about the causal impacts of TCMC, the set of background characteristics used in developing the comparison group have to include a set of variables that is sufficient to account for all other sources of confounding bias. As we describe further below, we included an extensive set of control variables, including a standardized measure of math achievement, information on prior math course-taking patterns, and a rich set of demographic information. However, we were unable to obtain access to information about student performance on the Texas Success Initiative (TSI) or other standardized measures of college-readiness, assessed prior to the start of students' senior year. It is quite plausible that these variables influenced student enrollment decisions, and they may also have associations with subsequent student outcomes, such as student decisions to pursue post-secondary education or enroll in college-level math courses—even after accounting for the observed measures of mathematics achievement and course-taking patterns—the impact estimates may include bias from omitted confounders. We must therefore be cautious in interpreting differences between TCMC students and comparison students as indicative of causal effects.

TCMC Implementation

As part of TCMC, each participating district partnered with a community college in order to satisfy the college readiness course requirements of HB5. Participating schools received free curriculum materials, a two-day, in-person professional development training, a further one-day training mid-way through the school year, and ongoing technical support. Participating schools agreed to assist with the Dana Center's evaluation activities by administering surveys to and providing further administrative data on students enrolled in the course.

Data Sources

Data sources for the present analysis were the same as in our previous analysis of the second year of implementation (Pustejovsky & Joshi, 2020). We used statewide longitudinal data collected by the Texas Education Agency (TEA), and Texas Higher Education Coordinating Board (THECB). The data include student demographics, course enrollment and completion in elementary and secondary schools, performance on State of Texas Assessments of Academic Readiness (STAAR) assessments, high school graduation, student matriculation into in-state colleges and universities, and student course-taking and course-performance information from in-state colleges and universities.

Analytic Samples. Our analytic sample is drawn from students in thirty districts in central Texas, who offered TCMC during the 2017-18 school year. In seven of the participating districts, schools were offering TCMC for the second year, having participated in the first year of implementation, while schools in the remaining districts were offering the course for the first time. Within each participating school, the analytic sample comprised two groups: students who enrolled in TCMC and students from the same cohort, within schools that offered TCMC, but who never enrolled in the course.[2] To identify sections of TCMC, we cross-checked enrollment data provided by the Dana Center against TEA course completion data for courses labeled as College Preparatory Mathematics or Independent Study, based on class name, identification number, class period, (anonymized) teacher identification number, and the number of sections offered. The treatment group consisted of 12^{th} grade students enrolled in sections of TCMC in the thirty districts that offered the course during 2017-18. We excluded a small number of students who appeared to be enrolled in sections of TCMC but were not actually enrolled in the focal campuses according to the TEA attendance data. The comparison group was drawn from among all 12^{th} grade students who were enrolled at the focal campuses during the 2017-18 school year but not enrolled in TCMC.

Outcomes. We examined the effects of participating in TCMC on several outcomes related to students' post-secondary success. Just as in our previous work (Pustejovsky & Joshi, 2020), our primary interest was in rates of enrollment and passage of college-level math coursework. In order to provide a more complete picture, we also examined several more proximal outcomes that could indicate how TCMC influences student trajectories, including math course-taking and course passage patterns during students' senior year, high school graduation rates, and college enrollment rates. We disaggregated college enrollment rates by sector because passing TCMC or another college preparatory math

[2] For students who were enrolled at multiple campuses during a single year (i.e., because they switched schools mid-year), we retained the record from the school with the maximum number of eligible days.

Table 1
Outcome definitions and sources

Outcome	Definition	Source
Senior year course-taking	Enrollment and passage of selected mathematics courses during senior year.	p__course__complete
High School Graduation	Graduating high school.	p__graduate
Post-Secondary Enrollment	Enrollment in community college, public and private four year-institutions, or health programs.	cbm__001
Post-Secondary Enrollment: Public Four Year	Enrollment in public four-year colleges.	cbm__001
Post-Secondary Enrollment: Private Four Year	Enrollment in private four-year colleges.	cbm__001
Post-Secondary Enrollment: Community	Enrollment in community colleges.	cbm__001
Post-Secondary Enrollment: Partner	Enrollment in community colleges that partnered with the focal districts to offer TCMC.	cbm__001
Post-Secondary Math Enrollment: College-Level	Enrollment in non-developmental college math courses. For all the outcomes below based on cbm__00s, we excluded students who took courses for dual credit. We excluded lab, co-op, internship, clinical and practicum courses.	cbm__00s
Post-Secondary Math Enrollment: Developmental	Enrollment in developmental college math courses.	cbm__00s
Post-Secondary Math Passing: College-Level	Passing non-developmental college math courses. For duplicated records, we kept the passing grade.	cbm__00s
Post-Secondary Math Passing: Developmental	Passing developmental college math courses. For duplicated records, we kept the passing grade.	cbm__00s

course provides exemption from developmental math coursework only at partner community colleges. Thus, the design of the college math course requirements in HB5 tended to incentivize enrollment at partner community colleges. Table 1 provides definitions of the outcomes and lists the data sources from which they were obtained.[3]

Covariates. As in our previous evaluation of the average effects of TCMC (Pustejovsky & Joshi, 2020), we used an extensive set of background characteristics for purposes of developing a comparison group and estimating effects of participating in

[3] Post-secondary outcome data were only available for students who enrolled in institutions of higher education within the state of Texas, as recorded in THECB data.

TCMC. Our main covariates included current demographic status and history, program and service enrollment history, prior math course enrollments, prior math performance, and current enrollment in Advanced Placement math or dual-credit math courses. Because TCMC is designed for students who are under-prepared for college-level math courses, it is highly likely that prior math course-taking and achievement influenced whether students were advised or required to take TCMC, and prior math achievement is also clearly related to later student outcomes. We also account for current enrollment in Advanced Placement math courses or dual-credit math courses because enrollment in these courses is generally indicative of college readiness, and so students who enroll in them would be very unlikely to enroll in TCMC.

Our data included information on student socio-economic background drawn from school administrative records. For the present analysis, we measured SES using indicators from multiple points in time. Specifically, our key measure of SES was the proportion of the previous eight years in which a student was eligible for free meals, based on TEA demographic records. Evidence indicates that this approach of using longitudinal measures of disadvantage better captures underlying socio-economic background, and more strongly correlates with educational outcomes, compared to using point-in-time measures (Michelmore & Dynarski, 2017). This variable plays a key role in answering our first research aim, regarding whether TCMC has different effects for students from different SES levels.

Table 2 reports the full set of control variables and the data sources from which they were obtained.[4] For categorical variables with missing data, we created an additional category indicating missingness. For continuous variables, missing data were imputed with the mean of the variable in the TCMC group within the given high school. For the continuous variables, we also created additional variables indicating missing values

[4] Pustejovsky & Joshi (2020, Tables 3 through 5) reported descriptive statistics on the distribution of the covariates among students who did or did not enroll in TCMC (prior to constructing the comparison group).

(Rosenbaum & Rubin, 1984).

Table 2
Covariate definitions and sources

Variable	Definition	Source
Sex	Sex- male/female.	p_attend_demog
Race/Ethnicity	Race/ethnicity- Asian American, African American, Hispanic, American Indian, Pacific Islander, Multiracial, and White.	p_attend_demog
Economic disadvantage	Indicates economic disadvantage status: free lunch status, reduced lunch status, no disadvantage or other disadvantage. We dummy coded the variable and took the average of the data from the attend and enroll datasets for the 12th grade year.	p_attend_demog & p_enroll_demog
At risk for dropping out	Indicates whether a student was at risk for dropping out of school according to state-defined criteria as of the beginning of the 12th grade year .	p_enroll_demog
Giftedness	Indicates whether a student participated in state-approved gifted and talented program. We took the average of the data from the attend and enroll datasets for the 12th grade year.	p_attend_demog & p_enroll_demog
Immigrant status	Indicates whether a student was identified as an immigrant according to the definition in Title III of No Child Left Behind Act of 2001- individuals who are aged 3 through 21, were not born in any state, and have not been attending one or more schools in any one or more states for more than three full academic years. The data is from the beginning of the 12th grade year.	p_enroll_demog
Special education status	Indicates whether a student participated in special education instructional and related services program or general education program using special education services, supplementary aids, or other special arrangements. We took the average of the data from the attend and enroll datasets for the 12th grade year.	p_attend_demog & p_enroll_demog
Limited English proficiency	Indicates whether a student was limited English proficient as determined by Language Proficiency Assessment Committee (LPAC) as of the end of the 12th grade year.	p_attend_demog
Prior math course-taking	Indicates whether a student took Grade 8 Mathematics (four years prior), Algebra I (three, four, or five years prior), Geometry (one, two, or three years prior), Algebra II (one, two, or three years prior) Pre-calculus (one or two years prior), Algebraic Reasoning (one or two years prior), and Mathematical Models with Applications (one, two, or three years prior). Variables for these courses indicated if the student passed, failed, or did not take the class in the given year.	p_course_complete
Prior math performance	STAAR end-of-course exam score for Algebra I, completed between 2014 and 2017. For duplicate scores (i.e., if the students took the test in multiple years), we kept the earliest score.	staareoca1
AP and DC enrollment	Indicates whether a student took an Advanced Placement math course or a dual-credit math course during their senior year.	p_course_complete
History of economic disadvantage, at-risk for dropping out, giftedness, immigrant status, and special education status	We tracked these variables from 2010 through 2017. Variables include (1) the number of years of available tracked data; (2) the number of years that a student was indicated as being in any of the categories for economic disadvantage and the number of years that the student was indicated as being in special education program, in gifted program, an immigrant, and at risk; (3) the proportion of years (the number of years the student was in the category divided by the number of years of record available) for the economic disadvantages categories; and, (4) if the student was ever indicated as being in special education program, in gifted program, an immigrant, and at risk.	p_enroll_demog

Constructing the Comparison Group

The five research questions that we investigate in this study used different analytic models. However, all of the analyses entailed first constructing comparison groups of students from each school who did not participate in TCMC. We summarize these methods before describing our approach to analysis of each research question.

To the construct comparison group, we estimated propensity scores using a generalized boosted regression model [GBRM; McCaffrey, Ridgeway, and Morral (2004)]. GBRM differs from standard methods (such as logistic regression) in that it is a non-parametric model, which does not impose strong assumptions about the functional form of the relationship between the propensity score and the covariates. Furthermore, rather than estimating the model by optimizing predictive fit, GBRM directly optimizes a measure of comparability between treated and untreated units. Thus, GBRM is particularly well-suited for estimating propensity scores based on a large set of covariates, as we use here (Lee, Lessler, & Stuart, 2009; McCaffrey, Ridgeway, & Morral, 2004).

We estimated propensity scores via GBRM with the twang package [version 1.4-9.5; Ridgeway, McCaffrey, Morral, Griffin, and Burgette (2016)] in R [version 3.3.1; R Core Team (2016)] with default settings for the number of trees, interaction depth, and shrinkage. The propensity score model included all of the covariates listed in Table 2. We included school and district IDs as additional covariates in the propensity score model, which has the effect of allowing the GBRM algorithm to discover school-by-covariate interactions that improve balance. Based on the propensity score model estimated using GBRM, average treatment effect for the treated (ATT) weights were calculated as

$$w_{ij} = D_{ij} + (1 - D_{ij}) \left(\frac{\hat{p}_{ij}}{1 - \hat{p}_{ij}} \right), \tag{1}$$

for student $i = 1, ..., n_J$ in school $j = 1, ..., J$, where D_{ij} is an indicator term equal to one if the student was enrolled in TCMC, and $\hat{p}_{ij}$ is the estimated propensity score for the student. Weights were standardized to sum to one within the TCMC and comparison groups within each school.

One way to understand the consequences of using these weights is to consider the implications for effective sample size (ESS). The ESS corresponds to the number of observations from an unweighted sample that would yield the same level of precision as a weighted sample (Ridgeway, McCaffrey, Morral, Griffin, & Burgette, 2016). Table 3 reports the ESS for the pool of possible comparison students (prior to weighting), for the weighted comparison group, and for the group of students who participated in TCMC at each school. School names are anonymized.

Analytic Models: TCMC Course Passing Rates

Our first research question involved purely descriptive analysis of course outcomes for the set of students who participated in TCMC. For each school that offered the course, we calculated the proportion of students who passed each semester of the course, relative to the total number of students enrolled. We summarized the course passing rates across schools using student-level weighted averages of the passing rates from each school. In other words, the overall course passing rate represents the proportion of *all* students who passed the course, relative to the total number of students enrolled across schools.

Analytic Models: Other Math Courses

Our second research question examined rates of enrollment and rates of passing math courses other than TCMC during students' senior year. For this analysis, we selected all math courses where there was substantial enrollment among students in the analytic sample, including Geometry, Algebra I, Algebra II, Algebraic Reasoning, Advanced Quantitative Reasoning, Mathematical Models with Applications, Pre-Calculus, AP Calculus (AB and BC), and Multivariable Calculus. For students enrolled in TCMC, we calculated the percentage of the analytic sample from each school that a) enrolled in a course section or b) passed the course section. For students in the comparison group, we

Table 3

Effective sample sizes by school and group

School	Comparison Pool (unweighted)	Comparison Group (weighted)	TCMC
AA	597	384.6	17
AB	631	272.4	130
AC	296	209.5	<5
AD	75	24.9	34
AE	546	124.5	11
AF	507	242.9	6
AG	259	109.5	43
AH	643	162.0	29
AI	434	175.4	63
AJ	292	139.2	42
AK	447	228.1	131
AL	23	6.4	5
AM	507	178.0	62
AN	573	233.8	30
AO	536	122.0	16
AP	500	152.9	146
AQ	327	115.2	19
AR	248	74.0	150
AS	664	364.8	121
AT	598	299.0	19
AU	1113	678.3	12
AV	329	153.7	53
AW	134	65.8	12
AX	471	175.7	106
AY	278	106.8	7
AZ	501	228.1	8
BA	523	141.1	19
BB	397	137.6	132
BC	74	16.9	7
BD	494	195.5	13
BE	71	38.2	8
BF	630	147.0	98
BG	406	110.0	49
BH	32	17.1	10
BI	426	234.7	65
BJ	86	43.8	32
BK	152	60.5	13
BL	253	61.7	13
BM	314	133.8	38
BN	822	380.6	70
BO	406	143.5	21
BP	884	374.6	330
BQ	86	57.5	<5
BR	604	268.8	25
BS	829	261.7	28
BT	244	143.5	49
BU	704	332.4	27
BV	75	26.4	29
BW	682	285.1	29
BX	36	9.7	39
BY	95	40.4	37
Total	20854	7970.6	2456

calculated the *weighted* percentage of the analytic sample from each school that enrolled or passed each course section, with weights given by Equation (1). To obtain an overall summary, we weighted the school-specific percentages based on the number of students enrolled in TCMC at each school. As a result, the enrollment rates and passing rates for the comparison group are based on a sample that matches the TCMC group in terms of student demographics and composition across schools. We therefore interpret these rates as estimates of the course-taking patterns that would have been observed among students enrolled in TCMC, had they not taken the course.

Analytic Models: Student-level Moderation

Our third research question was whether the effects of TCMC varied depending on students SES levels—a student-level characteristic. We addressed this question using a moderated regression analysis. For student i in school j, let Y_{ij} denote a measured outcome, let D_{ij} be an indicator equal to one if the student was enrolled in TCMC and equal to zero if the student was in the comparison group, let SES_{ij} be the measure of student SES[5], and let $\mathbf{X}_{ij}$ denote a set of additional covariates encoding student background characteristics other than SES. We centered each covariate at its unweighted mean in the treated group within each school, so that

$$\frac{1}{n_j}\sum_{i=1}^{n_j} D_{ij}\mathbf{X}_{ij} = \mathbf{0} \tag{2}$$

for $j = 1, ..., J$.

After centering, we applied the following analytic model:

$$Y_{ij} = \alpha_j + \beta_j D_{ij} + \boldsymbol{\gamma}_X \mathbf{X}_{ij} + \gamma_{SES} SES_{ij} + \boldsymbol{\delta}\mathbf{X}_{ij} D_{ij} + \delta_{SES} SES_{ij} D_{ij} + e_{ij}. \tag{3}$$

We estimated the model using weighted least squares, with weights given by Equation (1).

[5] The SES measure is a continuous variable that captures to the proportion of the previous eight years during which the student was eligible for free lunch. Thus, the variable ranges from zero to one, where zero corresponds to never being eligible for free lunch, one corresponds to always being eligible, and intermediate values correspond to being eligible during some years but not others.

We obtained standard errors for coefficient estimates HC0-type sandwich estimators (Zeileis, 2004, version 2.3-4), which are robust to heteroskedasticty in the regression errors of Equation (3).

Model (3) included school-specific intercept terms (α_j), which allowed average outcomes to differ across schools, as well as school-specific treatment effects (β_j), which allowed the difference in outcomes between TCMC students and comparison students to differ across schools. Furthermore, the model allowed the effect of treatment to interact with SES and with each of the remaining covariates. The term γ_{SES} captures the association between the outcome and SES in the absence of treatment, while holding constant the remaining covariates. Specifically, its magnitude corresponds to the difference in expected outcomes among matched comparison students who were always eligible for free meals ($SES_{ij} = 1$) versus matched comparison students who were never eligible ($SES_{ij} = 0$) but who were identical in terms of their remaining background characteristics.

The key parameter in Model (3) is δ_{SES}, which captures the extent to which the average effect of participating in TCMC *varies* depending on student SES. Specifically, it represents the difference in expected effects (i.e., average outcomes among TCMC students versus average outcomes among matched comparison students) for students who were always eligible for free meals ($SES_{ij} = 1$), relative to the expected effects for students who were never eligible ($SES_{ij} = 0$). Thus, a positive value of δ_{SES} would indicate that TCMC has *equity-enhancing* effects because it benefits lower SES students more strongly than higher-SES students; a value of $\delta_{SES} = 0$ would indicate that the effects of TCMC are similar for students from different socio-economic backgrounds; and a negative value of δ_{SES} would indicate that TCMC benefits higher SES students more strongly than lower-SES students.

Analytic Models: School-level Moderation

Our fourth and fifth research questions pertain to school-level heterogeneity in the effects of participating in TCMC. Consequently, we moved to a different analytic approach to address them—that of random effects meta-analysis (Hedges & Vevea, 1998). Specifically, we used a random effects meta-regression model of the school-specific treatment effect estimates to investigate: a) whether there were differences in average effects for schools implementing TCMC for the second year versus those implementing the program for the first time (Research Question 4) and b) the extent of additional heterogeneity in effects across schools (Research Question 5).

The random effects meta-analysis model provides a way to summarize and analyze the distribution of school-level effects (i.e., the β_j estimates from Model 3). Because the covariates $\mathbf{X}_{ij}$ and SES_{ij} were centered at school-specific means among TCMC students (following Equation 2), the estimate of β_j can be interpreted as the average effect of participating in TCMC for students in school j who chose to take the course. Let $\hat{\beta}_j$ be the weighted least squares estimate for school j, with estimated standard error $\hat{\sigma}_j$. Let F_j be an indicator variable equal to 1 if school j was implementing TCMC for the first time during the 2017-18 school year and equal to 0 if school j was implementing for the second year. With this notation, we estimated the following meta-analytic model:

$$\hat{\beta}_j = \mu_{First} F_j + \mu_{Second}(1 - F_j) + \nu_j + e_j, \tag{4}$$

where ν_j captures school-level variation in the effects of TCMC (apart from random sampling error) and e_j captures the sampling error associated with using the estimated value $\hat{\beta}_j$ rather than the true effect β_j. Following meta-analytic conventions, we assume that ν_j has expectation zero and variance τ^2 and that e_j has expectation zero and variance equal to $\hat{\sigma}_j^2$. We estimated Model (4) using restricted maximum likelihood methods via the metafor package for R (Viechtbauer, 2010). Following best practice, we obtained standard errors using the Knapp-Hartung method (Knapp & Hartung, 2003).

Regarding our fourth research question, the key parameters of Model (4) are μ_{First} and μ_{Second}, which represent the average effects of participating in TCMC among schools offering the course for the first or second year, respectively. Thus, a difference between μ_{First} and μ_{Second} would indicate that the effects of the program differ across the two groups of schools. It is also important to note that these average effects across each group of schools differ—and have a distinct interpretation—from the overall average effects reported in our previous evaluations. In our previous work, our focus was on estimating average effects across students who took TCMC within the set of schools that participated—a *student-level* average across a *sample* of schools. In contrast, the present analysis aims to estimate average effects across a set of schools that could have decided to offer TCMC—a *school-level* average across a hypothetical *population* of schools. This shift towards focusing on a hypothetical population of schools is one way to consider the *generalizability* of our previous findings. However, it does mean that estimates from the present analysis will tend to have a higher degree of uncertainty associated with them because we are now seeking to draw inferences about a broader population of schools—not just those that offered the course.

Regarding our fifth research question, the key parameter of Model (4) is τ, which represents the standard deviation (i.e., spread) of the distribution of school-specific effects. A larger standard deviation would indicate greater school-to-school variability in the effects of participating in TCMC; a standard deviation of zero (or near zero) would indicate that there is little variation in the effects. We calculated 95% confidence intervals for τ using the Q-profile method (Viechtbauer, 2007), as implemented in the metafor package.

Results

Covariate Balance

Before estimating the effects of participating in TCMC and investigating variation in those effects, we first used propensity score weighting to obtain a comparison group of students who did not take TCMC, but who closely resembled students from the same school who participated in the course. Overall, the re-weighting process resulted in a comparison group that was very similar to the set of students who took TCMC in terms of demographic characteristics, academic performance, and past math coursework. Our previous evaluation (Pustejovsky & Joshi, 2020) provided further, detailed information about the degree to which the re-weighted comparison group resembled the group of TCMC students. To the extent that these background characteristics explain whether a student enrolled in TCMC, differences in outcomes between students who took TCMC and students in the comparison group can be attributed to the impact of the program.

TCMC Passing Rates (Research Question 1)

The rates at which participating students passed the TCMC is an important initial consideration because students must pass the course in order to benefit from its administrative provisions. Of the schools that offered TCMC during the 2017-18 school year, the vast majority offered it as a two-semester sequence. Out of a total 2,416 students from 49 schools who enrolled in the first semester of the course, 95% passed the first semester of the sequence. Among the 2,422 students who enrolled in the second semester of the course, the overall passing rate was 91%. Two additional schools offered the course in a single-semester format; at these schools, the overall passing rate exceeded 98%.

Figure 1 depicts the course passing rates at participating schools, by semester; schools with a larger number of students who enrolled in the course are represented with larger bubbles.[6] As can be seen, passing rates varied across schools. A large majority of

[6] A small number of schools are not depicted because the number of enrolled students was so small that

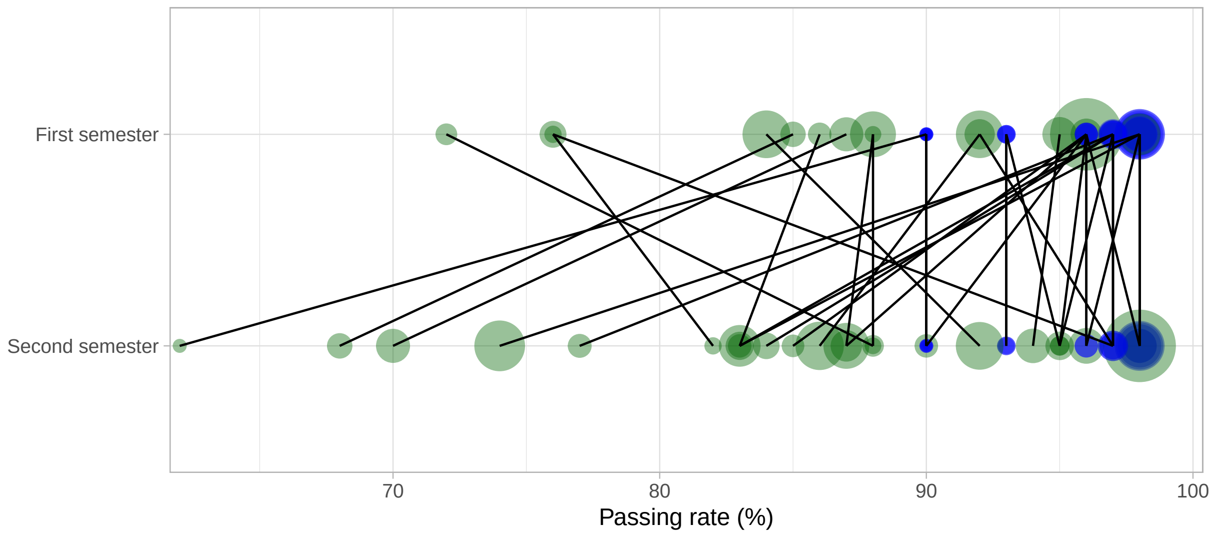

Figure 1. Passage rates of TCMC by semester and by school. The size of each point corresponds to the number of students enrolled in TCMC at the school. A black lines connect the passing rates from first semester and second semester at each unique school. Points in blue are masked (top-coded) in order to prevent disclosure of individually identifying information.

schools had passing rates of at least 80%, and approximately half of schools had passing rates of at least 90% for the second semester of the sequence. A small number of schools, most of which had small enrollments, saw passing rates below 80%.

Course-Taking Patterns During Senior Year (Research Question 2)

In addition to understanding course passage rates, we can gain a fuller picture of the experiences of students who took TCMC by considering the courses that they *did not* take—that is, the courses that they forwent due to enrolling in TCMC. We examined this question by comparing rates of enrollment and rates of passing math courses during their senior year among TCMC students to rates among students in the comparison group. To the extent that the comparison group accounts for all important confounding factors, then their rates of enrollment and passing provide an estimate of the course-taking patterns that TCMC students would have experienced, if TCMC were not available.

passing rates had to be masked. For some additional schools, depicted in blue, passing rates were top-coded due to disclosure limitations.

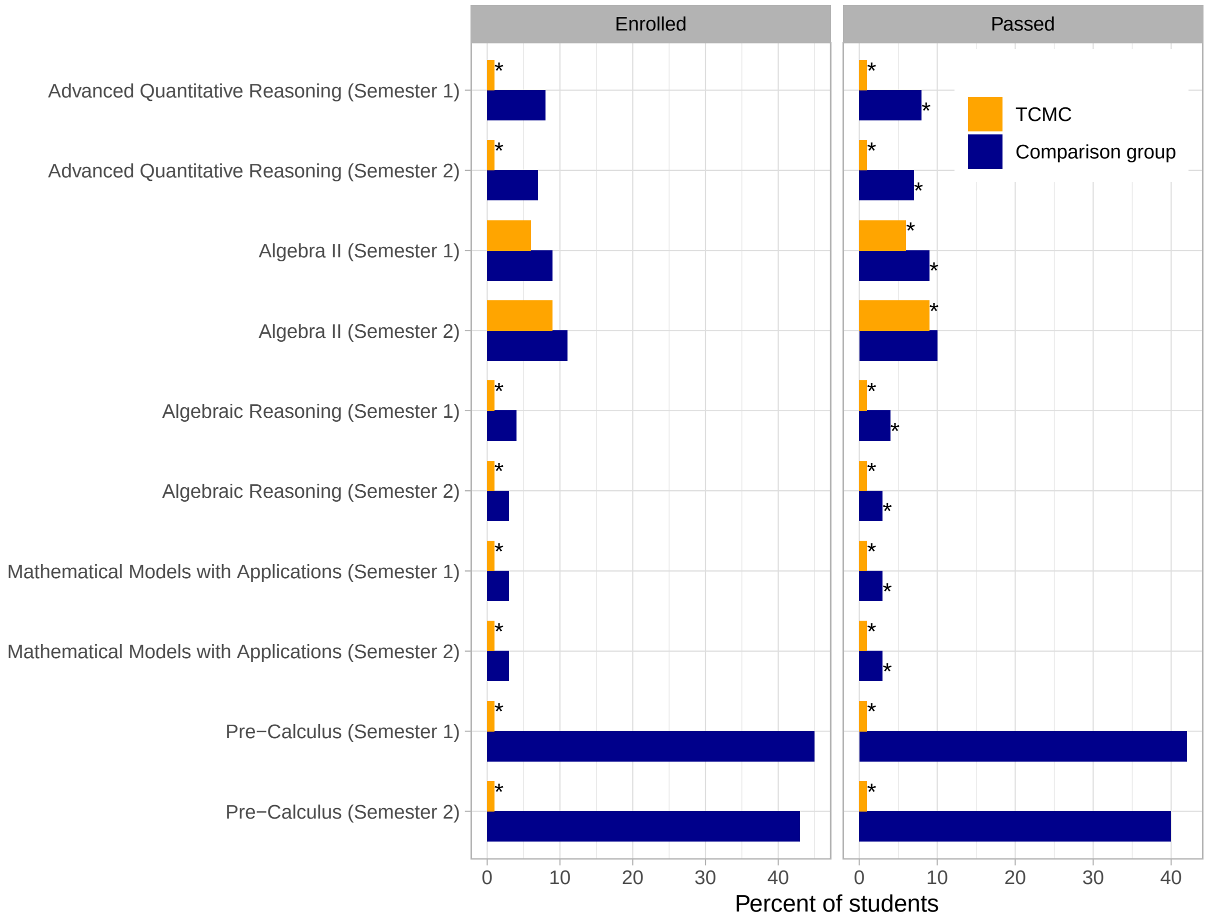

Figure 2. Percentage of students who enrolled in and passed selected math courses during senior year, among students taking TCMC and students in the comparison group. Percentages marked '*' indicate that the exact percentage is masked (bottom-coded) in order to prevent disclosure of individually identifying information.

Figure 2 depicts rates of enrollment (left-hand panel) and rates of course-passing (right-hand panel) during students' senior year for selected math courses.[7] Notably, over 40% of comparison students enrolled in Pre-Calculus during their senior year, making it the most common alternative math course to TCMC. Smaller numbers of comparison group students enrolled in Advanced Quantitative Reasoning, Algebraic Reasoning, or Mathematical Models with Applications. Very few TCMC students enrolled in these courses, which suggests that they might have served as potential alternatives to taking

[7] The figure does not depict results for Algebra I, Geometry, AP Calculus, or Multivariable Calculus because only a very small fraction of students enrolled in any of these courses during their senior year.

TCMC. Both TCMC students and comparison students enrolled in Algebra II, although comparison students enrolled at slightly higher rates than TCMC students—particularly for the first semester of the sequence. For each of these courses, passage rates follow a very similar pattern to that of enrollment rates, and Pre-Calculus was the primary alternative math course passed by students in the comparison group.

Student-Level Moderation by Socio-Economic Status (Research Question 3)

We now turn to investigating variation in the effects of participating in TCMC, across several dimensions. We first considered whether differences in outcomes between students who participated in TCMC versus comparison students varied depending on students' socio-economic status (SES)—an individual-level moderating factor. Table 4 reports estimates of the moderating relation with rates of high school graduation and post-secondary enrollment. For reference, the first set of columns provides the estimates of the average differences between TCMC students and comparison students, as reported in our previous evaluation (Pustejovsky & Joshi, 2020). The second set of columns (labeled γ_{SES}) provides estimates of the relation between student SES and each outcome among students in the comparison group. These estimates indicate that student SES is strongly predictive of rates of post-secondary enrollment overall, enrollment in community college, and enrollment in public four-year colleges or universities. For instance, by 4 semesters after high school graduation, students who were always eligible for free lunch ($SES_{ij} = 1$) were 12.70 percentage points less likely to have enrolled in community college, compared to students who were never eligible for free lunch ($SES_{ij} = 0$).

The final set of columns in Table 4 (labeled δ_{SES}) report estimates of the moderating effect of SES. Each estimate represents a *difference* in the estimated average effect of participating in TCMC, comparing students who were always eligible for free lunch to students who were never eligible for free lunch. Across outcomes, none of the moderating effects are statistically distinguishable from zero, meaning that we cannot rule out the

Table 4

Moderating effects of SES on high school graduation and post-secondary enrollment rates

	Average effect of TCMC			SES slope (γ_{SES})			SES moderation (δ_{SES})		
Outcome	Est.	SE	p	Est.	SE	p	Est.	SE	p
High School Graduation									
1 year	4.7	0.4	<.001	-0.7	1.3	.566	-0.9	1.4	.530
2 year	3.8	0.3	<.001	-0.5	1.2	.669	-0.2	1.2	.879
Post-Secondary Enrollment: Overall									
1 semester	0.6	1.2	.603	-16.4	2.4	<.001	-2.9	4.0	.465
2 semester	2.1	1.2	.074	-17.6	2.4	<.001	-1.7	4.0	.681
3 semester	1.9	1.2	.105	-17.3	2.4	<.001	-1.6	4.0	.694
4 semester	2.0	1.2	.093	-17.4	2.4	<.001	-1.4	4.0	.727
Post-Secondary Enrollment: Community									
1 semester	3.2	1.1	.003	-9.0	2.3	<.001	-5.2	3.8	.177
2 semester	4.8	1.2	<.001	-10.5	2.4	<.001	-4.3	4.0	.283
3 semester	4.0	1.2	<.001	-12.7	2.4	<.001	-3.5	4.0	.384
4 semester	4.0	1.2	<.001	-12.7	2.4	<.001	-3.5	4.0	.384
Post-Secondary Enrollment: Public Four Year									
1 semester	-2.5	0.6	<.001	-6.8	1.5	<.001	1.1	2.5	.652
2 semester	-2.4	0.7	<.001	-6.9	1.6	<.001	1.6	2.5	.514
3 semester	-2.5	0.7	<.001	-6.8	1.6	<.001	1.6	2.5	.517
4 semester	-2.5	0.7	<.001	-8.3	1.6	<.001	1.1	2.6	.664
Post-Secondary Enrollment: Private Four Year									
1 semester	-0.3	0.3	.417	-1.2	0.9	.173	0.7	1.3	.559
2 semester	-0.3	0.3	.391	-1.1	0.9	.209	0.8	1.3	.554
3 semester	-0.3	0.3	.391	-1.1	0.9	.209	0.8	1.3	.554
4 semester	-0.3	0.3	.367	-1.1	0.9	.206	0.8	1.3	.570
Post-Secondary Enrollment: Partner									
1 semester	1.9	1.0	.055	-7.5	2.1	<.001	-0.5	3.5	.890
2 semester	2.8	1.1	.007	-8.5	2.2	<.001	-0.5	3.7	.885
3 semester	2.1	1.1	.049	-10.1	2.3	<.001	-0.5	3.8	.905
4 semester	2.0	1.1	.059	-10.1	2.3	<.001	-0.4	3.8	.922

possibility that effects of participating in TCMC were consistent across students from different socio-economic backgrounds. Across outcomes, the largest moderator relation was $\hat{\delta}_{SES} = -5.2$ percentage points (SE = 3.8) for community-college enrollment within one semester after high school graduation, meaning that the estimated average effect of participating in TCMC was about 5 percentage points smaller among students who were always free-lunch eligible than the average effect among students who were never free-lunch eligible. However, this estimate was very uncertain and not statistically distinguishable from zero. Across outcomes, estimates of the moderating effects of SES had the opposite sign from the corresponding estimates of overall average effects. This suggests that, even if the effects of TCMC did differ by SES, participating in the course may have had stronger effects for higher-SES students.

Table 5 reports results of the moderator analysis for post-secondary math course enrollment and course passage rates. It is laid out following the same format as Table 4. As with the post-secondary enrollment outcomes, student free-lunch eligibility was negatively related to math course enrollment and math course passing rates. Across outcomes, SES did not appear to moderate the effects of participating in TCMC. Few of the moderating relations were statistically distinguishable from zero, indicating that it was possible that the effects of participating in TCMC were consistent across students from varying socio-economic backgrounds.

Table 5
Moderating effects of SES on post-secondary math course-taking and course-passing rates

	Average effect of TCMC			SES slope (γ_{SES})			SES moderation (δ_{SES})		
Outcome	Est.	SE	p	Est.	SE	p	Est.	SE	p
Post-Secondary Math Enrollment: College-Level									
1 semester	-1.7	0.5	.001	-2.2	1.1	.039	-1.2	1.8	.498
2 semester	-2.4	0.8	.004	-7.6	1.9	<.001	0.9	3.0	.763
3 semester	-2.4	0.8	.003	-9.0	1.9	<.001	1.1	3.0	.720
4 semester	-2.4	0.8	.003	-9.0	1.9	<.001	1.1	3.0	.720
Post-Secondary Math Enrollment: Developmental									
1 semester	-0.2	0.3	.523	-2.3	1.0	.023	2.5	1.3	.047
2 semester	0.1	0.5	.775	-2.5	1.2	.042	1.4	1.7	.422
3 semester	0.2	0.5	.695	-2.4	1.2	.052	1.5	1.7	.397
4 semester	0.2	0.5	.695	-2.4	1.2	.052	1.5	1.7	.397
Post-Secondary Math Passing: College-Level									
1 semester	-1.6	0.4	<.001	-1.6	0.9	.074	-0.7	1.5	.643
2 semester	-2.4	0.6	<.001	-3.6	1.5	.014	-1.2	2.4	.614
3 semester	-2.5	0.7	<.001	-4.6	1.5	.003	-1.7	2.5	.501
4 semester	-2.5	0.7	<.001	-4.6	1.5	.003	-1.7	2.5	.501
Post-Secondary Math Passing: Developmental									
1 semester	-0.1	0.2	.805	-2.0	0.9	.031	2.5	1.1	.020
2 semester	0.1	0.4	.784	-2.1	1.1	.049	1.4	1.5	.340
3 semester	0.1	0.4	.792	-2.2	1.1	.042	1.6	1.5	.289
4 semester	0.1	0.4	.792	-2.2	1.1	.042	1.6	1.5	.289

School-Level Moderation by Year of Implementation (Research Question 4)

Tables 6 and 7 organize our main results regarding school-level variation in the effects of TCMC, for enrollment outcomes and math course-related outcomes, respectively. In each table, the first two sets of columns report estimated average effects for the set of schools that were implementing TCMC for the first time (μ_{First}) or for the smaller set of schools that were offering the course for the second year (μ_{Second}).[8] The third set of columns reports the difference in average effects between the two groups of schools.

Across the outcomes of high school graduation and post-secondary enrollment (Table 6), the estimated effects of participating in TCMC were largely consistent across the two groups of schools. None of the differences between school types were statistically distinguishable from zero. Furthermore, the estimated effects within each group of schools were generally consistent in terms of sign and magnitude. For example, the estimated effects of participating in TCMC on post-secondary enrollment in a community college were all positive and on the order of two to four percentage points, depending on the follow-up time. Similarly, the estimated effects on post-secondary enrollment in a four-year college or university were all negative and on the order of three to five percentage points, depending on the follow-up time.

We found a very similar pattern for the math course-taking and course-passing outcomes (Table 7). None of the differences between school types were statistically distinguishable from zero, and the sign and magnitude of average effects were generally similar for schools in the first year or second year of implementing the program. The largest difference between the two groups of schools was only one percentage point, for the estimated effects of enrolling in a college-level math course. Overall, there was little indication that the effects of participating in TCMC varied across the two groups of schools.

[8] Each of these average effect estimates is based on a random effects meta-analysis of the school-level average effects. The estimates are therefore for a school-level average effect, rather than an individual-level average across the sample of students who participated in the course.

Table 6

School-level average effects of TCMC on high school graduation and post-secondary enrollment rates, by year of implementation

	First year of implementation (μ_{first})			Second year of implementation (μ_{second})			Difference ($\mu_{second} - \mu_{first}$)			Heterogeneity (τ)	
Outcome	Est. (SE)	95% CI	p	Est. (SE)	95% CI	p	Est. (SE)	95% CI	p	Est.	95% CI
High School Graduation											
1 year	3.7 (0.6)	[2.5, 4.9]	<.001	5.4 (1.2)	[3.1, 7.8]	<.001	1.7 (1.3)	[-0.9, 4.4]	.197	2.8	[2.2, 4.3]
2 year	2.9 (0.5)	[2.0, 3.8]	<.001	4.1 (0.9)	[2.4, 5.9]	<.001	1.2 (1.0)	[-0.7, 3.2]	.209	2.0	[1.3, 3.2]
Post-Secondary Enrollment: Overall											
1 semester	-0.7 (1.3)	[-3.3, 2.0]	.617	-0.1 (2.2)	[-4.4, 4.3]	.979	0.6 (2.5)	[-4.5, 5.7]	.811	2.1	[0.0, 6.8]
2 semester	1.7 (1.2)	[-0.8, 4.2]	.173	-0.5 (2.0)	[-4.5, 3.5]	.793	-2.2 (2.3)	[-6.9, 2.5]	.342	0.0	[0.0, 5.6]
3 semester	2.1 (1.2)	[-0.3, 4.5]	.086	-1.1 (1.9)	[-4.9, 2.7]	.553	-3.2 (2.2)	[-7.7, 1.3]	.155	0.0	[0.0, 4.5]
4 semester	2.2 (1.2)	[-0.2, 4.6]	.072	-1.2 (1.9)	[-5.0, 2.6]	.534	-3.4 (2.2)	[-7.9, 1.1]	.137	0.0	[0.0, 4.7]
Post-Secondary Enrollment: Community											
1 semester	2.1 (1.4)	[-0.8, 4.9]	.152	4.0 (2.3)	[-0.7, 8.7]	.092	1.9 (2.7)	[-3.5, 7.4]	.480	4.2	[0.0, 7.6]
2 semester	4.0 (1.5)	[1.0, 7.0]	.010	4.1 (2.5)	[-0.9, 9.1]	.105	0.1 (2.9)	[-5.7, 5.9]	.973	4.8	[0.0, 7.8]
3 semester	4.2 (1.3)	[1.6, 6.8]	.002	1.8 (1.9)	[-2.1, 5.7]	.356	-2.4 (2.3)	[-7.1, 2.3]	.308	1.5	[0.0, 5.5]
4 semester	4.2 (1.3)	[1.6, 6.8]	.002	1.8 (1.9)	[-2.1, 5.7]	.356	-2.4 (2.3)	[-7.1, 2.3]	.308	1.5	[0.0, 5.5]
Post-Secondary Enrollment: Public Four Year											
1 semester	-3.9 (0.9)	[-5.7, -2.1]	<.001	-4.8 (1.7)	[-8.1, -1.4]	.006	-0.9 (1.9)	[-4.7, 2.9]	.638	3.7	[2.5, 6.2]
2 semester	-3.6 (0.9)	[-5.5, -1.8]	<.001	-4.9 (1.7)	[-8.3, -1.5]	.005	-1.3 (1.9)	[-5.2, 2.6]	.501	3.7	[2.6, 6.4]
3 semester	-3.7 (0.9)	[-5.5, -1.8]	<.001	-5.0 (1.7)	[-8.4, -1.5]	.005	-1.3 (1.9)	[-5.2, 2.6]	.502	3.8	[2.6, 6.4]
4 semester	-3.9 (0.9)	[-5.8, -2.1]	<.001	-3.5 (1.8)	[-7.0, 0.0]	.052	0.4 (2.0)	[-3.6, 4.4]	.832	3.6	[2.3, 6.4]
Post-Secondary Enrollment: Private Four Year											
1 semester	-0.5 (0.2)	[-0.9, -0.1]	.007	-0.5 (0.3)	[-1.1, 0.2]	.159	0.0 (0.4)	[-0.7, 0.8]	.894	0.5	[0.0, 1.1]
2 semester	-0.5 (0.2)	[-0.9, -0.1]	.008	-0.5 (0.3)	[-1.2, 0.2]	.176	0.1 (0.4)	[-0.7, 0.9]	.876	0.5	[0.0, 1.2]
3 semester	-0.5 (0.2)	[-0.9, -0.1]	.008	-0.5 (0.3)	[-1.2, 0.2]	.176	0.1 (0.4)	[-0.7, 0.9]	.876	0.5	[0.0, 1.2]
4 semester	-0.6 (0.2)	[-1.0, -0.2]	.007	-0.5 (0.4)	[-1.3, 0.2]	.162	0.0 (0.4)	[-0.8, 0.9]	.925	0.6	[0.0, 1.2]
Post-Secondary Enrollment: Partner											
1 semester	1.3 (1.3)	[-1.3, 4.0]	.317	1.8 (2.1)	[-2.4, 6.0]	.383	0.5 (2.5)	[-4.5, 5.5]	.842	3.6	[0.0, 7.9]
2 semester	2.8 (1.4)	[0.0, 5.6]	.048	1.2 (2.2)	[-3.2, 5.6]	.587	-1.6 (2.6)	[-6.8, 3.5]	.529	3.5	[0.0, 7.9]
3 semester	2.8 (1.4)	[0.0, 5.5]	.049	0.1 (2.1)	[-4.1, 4.3]	.963	-2.7 (2.5)	[-7.6, 2.3]	.288	2.8	[0.0, 7.3]
4 semester	2.6 (1.3)	[-0.1, 5.3]	.056	0.1 (2.0)	[-4.0, 4.1]	.966	-2.5 (2.4)	[-7.4, 2.3]	.299	2.6	[0.0, 7.0]

Table 7

School-level average effects of TCMC on post-secondary math course-taking and course-passing rates, by year of implementation

	First year of implementation (μ_{first})			Second year of implementation (μ_{second})			Difference ($\mu_{second} - \mu_{first}$)			Heterogeneity (τ)	
Outcome	Est. (SE)	95% CI	p	Est. (SE)	95% CI	p	Est. (SE)	95% CI	p	Est.	95% CI
Post-Secondary Math Enrollment: College-Level											
1 semester	-2.8 (0.7)	[-4.2, -1.5]	<.001	-2.7 (1.3)	[-5.4, -0.1]	.044	0.1 (1.5)	[-2.8, 3.1]	.931	2.9	[2.1, 4.7]
2 semester	-3.4 (0.9)	[-5.2, -1.6]	<.001	-2.3 (1.6)	[-5.6, 0.9]	.151	1.1 (1.8)	[-2.6, 4.7]	.565	2.4	[0.0, 5.3]
3 semester	-3.6 (1.0)	[-5.6, -1.7]	<.001	-2.6 (1.7)	[-6.0, 0.9]	.138	1.1 (2.0)	[-2.9, 5.0]	.587	2.8	[0.0, 5.8]
4 semester	-3.6 (1.0)	[-5.6, -1.7]	<.001	-2.6 (1.7)	[-6.0, 0.9]	.138	1.1 (2.0)	[-2.9, 5.0]	.587	2.8	[0.0, 5.8]
Post-Secondary Math Enrollment: Developmental											
1 semester	-0.1 (0.2)	[-0.4, 0.2]	.571	-0.3 (0.3)	[-0.9, 0.3]	.253	-0.3 (0.3)	[-0.9, 0.4]	.454	0.3	[0.0, 1.5]
2 semester	-0.9 (0.4)	[-1.6, -0.1]	.028	-1.3 (1.0)	[-3.2, 0.7]	.209	-0.4 (1.1)	[-2.5, 1.7]	.705	1.1	[1.0, 3.3]
3 semester	-1.0 (0.4)	[-1.8, -0.1]	.027	-1.0 (1.0)	[-3.1, 1.1]	.346	-0.0 (1.1)	[-2.3, 2.2]	.975	1.4	[0.9, 3.4]
4 semester	-1.0 (0.4)	[-1.8, -0.1]	.027	-1.0 (1.0)	[-3.1, 1.1]	.346	-0.0 (1.1)	[-2.3, 2.2]	.975	1.4	[0.9, 3.4]
Post-Secondary Math Passing: College-Level											
1 semester	-2.4 (0.6)	[-3.5, -1.3]	<.001	-2.1 (1.0)	[-4.2, 0.0]	.052	0.3 (1.2)	[-2.0, 2.7]	.778	2.3	[1.7, 4.1]
2 semester	-3.4 (0.8)	[-5.0, -1.8]	<.001	-3.0 (1.4)	[-5.9, -0.2]	.039	0.4 (1.6)	[-2.9, 3.7]	.824	2.9	[1.7, 5.3]
3 semester	-3.4 (0.9)	[-5.1, -1.6]	<.001	-3.2 (1.5)	[-6.2, -0.1]	.042	0.2 (1.8)	[-3.4, 3.7]	.925	3.1	[1.9, 6.0]
4 semester	-3.4 (0.9)	[-5.1, -1.6]	<.001	-3.2 (1.5)	[-6.2, -0.1]	.042	0.2 (1.8)	[-3.4, 3.7]	.925	3.1	[1.9, 6.0]
Post-Secondary Math Passing: Developmental											
1 semester	-0.0 (0.1)	[-0.2, 0.1]	.765	-0.1 (0.2)	[-0.5, 0.2]	.371	-0.1 (0.2)	[-0.5, 0.2]	.521	0.1	[0.0, 0.9]
2 semester	-0.4 (0.2)	[-0.7, 0.0]	.052	-0.6 (0.5)	[-1.6, 0.3]	.185	-0.3 (0.5)	[-1.3, 0.7]	.597	0.0	[0.0, 1.8]
3 semester	-0.4 (0.2)	[-0.8, -0.0]	.035	-0.6 (0.5)	[-1.6, 0.4]	.205	-0.2 (0.5)	[-1.3, 0.9]	.704	0.2	[0.0, 1.8]
4 semester	-0.4 (0.2)	[-0.8, -0.0]	.035	-0.6 (0.5)	[-1.6, 0.4]	.205	-0.2 (0.5)	[-1.3, 0.9]	.704	0.2	[0.0, 1.8]

School-Level Heterogeneity of Effects (Research Question 5)

Our final research question was whether the effects of participating in TCMC varied across participating schools, above and beyond any differences by year of implementation. If school-to-school variation is evident, it suggests that local contextual factors (other those that we have examined) may have influenced the program's implementation. To address this question, we used a random effects meta-analysis model to estimate the *standard deviation* of school-specific effects, denoted $\hat{\tau}$, after accounting for year of implementation and for the sampling uncertainty of the school-specific effect estimates.[9] Larger standard deviations indicate greater heterogeneity in effects across schools.

The final set of columns in Tables 6 and 7 report the estimated standard deviations of school effects, along with 95% confidence intervals, for each outcome. We found evidence of heterogeneous effects on high school graduation rates and on rates of enrollment in public four-year colleges or universities (Table 6). For example, four semesters after high school graduation, we estimated that rates of enrollment in public four-years varied across schools, with a standard deviation of 3.6 percentage points, 95% CI [2.3, 6.4]. Based on this estimate and on the estimated average effect of -3.50 percentage points among schools in their second year of implementation, we would expect that approximately two thirds of such schools would have effects between -7.10 and 0.10 percentage points—a non-trivial degree of variation. For other post-secondary enrollment outcomes, such as enrollment in community colleges, we found lower degrees of heterogeneity that were not statistically distinguishable from zero. For example, we estimated a standard deviation of 1.5 percentage points, 95% CI [0.0, 5.5] for effects on community college enrollment, four semesters after high school graduation. Thus, for these outcomes, we cannot rule out the possibility that the effects of participating in TCMC were consistent across schools, nor can

[9] Some degree of variation in the school-specific effect estimates is to be expected simply because each estimate is based on a limited–often fairly small–sample of data from each school. The set of estimates will exhibit variation simply because of this sampling uncertainty, and this would be true even if the underlying true effect was constant across schools. The random effects meta-analysis approach that we apply takes this sampling uncertainty into account.

we rule out the possibility of moderate heterogeneity.

Regarding enrollment and passing rates in post-secondary math courses (Table 7), we found evidence of a small degree of heterogeneity in effects on rates of enrollment in a developmental math course ($\hat{\tau} = 1.4$ percentage points, 95% CI [0.9, 3.4] by four semesters after high school graduation) and a moderate degree of heterogeneity in effects on rates of passing a college-level course ($\hat{\tau} = 3.1$ percentage points, 95% CI [1.9, 6.0] by four semesters after high school graduation). Given the estimate of heterogeneity in effects on college-level math course passage rates, along with an estimated average effect of -3.20 percentage points among schools in their second year of implementation, we would expect that approximately two thirds of such schools would have effects between -6.30 and -0.10 percentage points. This indicates that the implementation of TCMC at different schools may have led to varied outcomes, with some contexts leading to little or no effects on success in college-level math courses, while other contexts led to negative effects. Taken as a whole, these findings suggest that there were heterogeneous effects of participating in TCMC on some key outcomes that the program intended to affect.

Discussion

In our previous evaluation work, we focused on understanding differences in *average* outcomes, comparing students who participated in TCMC during their senior year to students from the same schools who did not take the course but who closely resembled students that did. In the present evaluation, we shifted perspective by considering, first, proximal consequences of enrollment in TCMC and, second, the extent to which differences between TCMC students and comparison students may have *varied* depending on student-level or school-level factors.

Regarding proximal outcomes of enrollment in TCMC, we found that enrolled students passed the course at high rates, on average, although there was some school-to-school variation in passage rates. Furthermore, Pre-Calculus was the most common senior year math course taken by students in the comparison group. Smaller numbers of comparison students took courses such as Advanced Quantitative Reasoning, Algebraic Reasoning, and Mathematical Models with Applications. If the comparison group was constructed in a way to account for all important confounding factors, then we can interpret the course-taking patterns among the comparison group as representing courses that students who enrolled in TCMC would have taken, if TCMC had not been available. However, as we noted in previous evaluations, it is possible that the comparison group was not well-matched with the group of TCMC students on measures of students college readiness in mathematics (such as the Texas Success Initiative, SAT, or ACT exam) that could have influenced their course placement during senior year. In this case, the course-taking patterns of the comparison group would likely be skewed towards more advanced courses than what TCMC students would have otherwise taken, simply because the comparison group would tend to include students with stronger math skills. In light of the pattern of course-taking patterns that we observed, we consider this quite plausible.

The other focus of our investigation was heterogeneity in outcome comparisons between TCMC students and comparison students. Here, we found little indication that the difference between TCMC students and comparison students varied by socio-economic backgrounds, as measured by the proportion of years that a student was eligible for free lunch. Neither did we find evidence that effects varied by whether a school was offering TCMC for the first time or second time. However, for key outcomes such as passing a college-level math course, the degree of difference between TCMC students and comparison students varied across the schools that offered the course. One possible interpretation of this heterogeneity is that the effects of the program depended on aspects of the school context that varied across schools. Such contextual factors might include the course instructor's level of experience, aspects of the coordinating agreements between districts where the course was offered and their partner community colleges, or the degree of alignment between the TCMC curriculum and college-level math coursework at partner community colleges. In further studies on TCMC or other college readiness math courses, it would be valuable to track and investigate factors such as these.

Another possible explanation of school-to-school heterogeneity is that the approach we took to constructing a comparison group may have been more successful in reducing bias for some schools than for others. If students who took TCMC at some schools were better matched with comparison students, while those at other schools were not as closely matched on important background characteristics, then the school-specific effect estimates would include varying degrees of bias. Even if the true effects were entirely consistent across schools, this varying degree of bias could show up as heterogeneity in our random effects meta-analysis. This possibility points towards a limitation of our overall evaluation strategy, discussed further below, with respect to estimation of school-specific treatment effects.

Limitations

Just as in our previous evaluations, findings from this analysis must be interpreted cautiously due to our inability to fully account for initial differences between TCMC students and comparison students at the start of their senior year. Notably, we were unable to access data on whether students had already achieved college readiness (i.e., based on the Texas Success Initiative or some other standardized assessment) at the start of their senior year, and therefore unable to adjust for differences between comparison students and TCMC students on direct markers of college readiness, nor to exclude comparison students who were already college ready and therefore not eligible for TCMC.

Relative to students in TCMC, college-ready students in the comparison group would be more likely to apply to and gain admission into four-year colleges and more likely to enroll in and pass college-level math courses. In turn, this would tend to make the estimated average impacts of participating in TCMC appear more negative (detrimental) than they really are. It is much less clear how our inability to adjust for college readiness indicators might have influenced findings regarding effect moderation. In order for these results to be affected, it would have to be the case that the bias due to college readiness status was associated with the potential moderator of interest—for example, that the bias due to college readiness status was more pronounced among higher-SES students than among lower-SES students. This might have been the case for the analysis of student-level socio-economic status, although it is hard to say how likely. It seems less plausible for the analysis of moderation by year of implementation.

Another important limitation of the analysis is that our evaluation approach was not designed to generate school-level estimates of the effects of participating in TCMC. Estimates for individual schools were subject to much greater sampling uncertainty than were our reported estimates of average effects across schools. Furthermore, our approach to constructing the comparison group aimed to reduce covariate imbalances *in the aggregate*, across schools, rather than to create comparison groups that were balanced at the level of individual schools. As a result, school-specific estimates of the effects of participating in TCMC were more influenced by our approach to regression adjustment (i.e., Equation 3), and thus more susceptible to bias from model mis-specification. For all of these reasons, we have refrained from reporting school-specific estimates, instead focusing on meta-analytic summaries of their distribution.

Implications

Findings from this study and our previous evaluation work suggest that participating in TCMC might have shifted some students who might otherwise have enrolled in four-year institutions towards enrolling in community college instead. Furthermore, the strength of these effects may have varied across schools in different contexts. These impacts would be consistent with the incentive structure of the college preparatory course requirements created by HB 5, in that the successful completion of the course provided exemption from proficiency exams only at the partner institution. However, it is also possible that the pattern of results could be due to pre-existing differences between TCMC students and comparison students—particularly differences in achieving college readiness milestones or differences in aspirations or goals for attending four-year colleges—for which our analysis was not able to adjust. Despite the possibility of such bias, it seems nonetheless prudent for further evaluations of TCMC and other college preparatory math courses to attend to the possibility that the incentive structures created by HB5 math courses could lead to shifts in enrollment patterns across sectors.

In order to more effectively evaluate TCMC–and other similar programs–it will be critical to consider and account for students' college readiness status while students are still in high schools. Incorporating such data into the Texas PK-20 Workforce Database would enhance the capacity of the research community for understanding the impacts of policy changes, programs, and other strategies aimed at improving student success across the transition between secondary and post-secondary education.

A final policy recommendation is to continue investigating the implementation of college preparatory math courses implemented in response to HB5. Further work should investigate how schools assess college readiness and determine whether students should take college readiness courses. It would also be valuable to investigate students' perspectives and experiences in such courses as they navigate the transition from high school to college.

References

An, B. P. (2013). The Impact of Dual Enrollment on College Degree Attainment: Do Low-SES Students Benefit? *Educational Evaluation and Policy Analysis*, *35*(1), 57–75. https://doi.org/10.3102/0162373712461933

Bloom, H. S., Raudenbush, S. W., Weiss, M., & Porter, K. (2017). Using multi-site experiments to study cross-site variation in effects of program assignment. *Journal of Research on Educational Effectiveness*.

Charles A. Dana Center. (2016). *The case for Mathematics Pathways*. Austin, TX: University of Texas at Austin. Retrieved from https://dcmathpathways.org/resources/case-mathematics-pathways

Charles A. Dana Center. (2019). *The case for Mathematics Pathways - updated*. Austin, TX: University of Texas at Austin. Retrieved from https://dcmathpathways.org/resources/case-mathematics-pathways

Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods*, *3*(4), 486.

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, *22*(17), 2693–2710.

Lee, B. K., Lessler, J., & Stuart, E. A. (2009). Improving propensity score weighting using machine learning. *Statistics in Medicine*, n/a–n/a. https://doi.org/10.1002/sim.3782

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*, *9*(4), 403–425. https://doi.org/10.1037/1082-989X.9.4.403

Michelmore, K., & Dynarski, S. (2017). The gap within the gap: Using longitudinal data to understand income differences in educational outcomes. *AERA Open*, *3*(1), 1–18. https://doi.org/10.1177/2332858417692958

Pustejovsky, J. E., & Joshi, M. (2019). *Evaluating the transition to college mathematics course in texas high schools: Findings from the first year of implementation*. Bryant, TX: Greater Texas Foundation. Retrieved from https://www.greatertexasfoundation.org/wp-content/uploads/2019/06/Pustejovsky-Final-2019.pdf

Pustejovsky, J. E., & Joshi, M. (2020). *Evaluating the transition to college mathematics course in texas high schools: Findings from the second year of implementation.* Bryant, TX: Greater Texas Foundation. Retrieved from https://www.greatertexasfoundation.org/wp-content/uploads/2020/10/PustejovskyWhitepaper_v1.pdf

R Core Team. (2016). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Raudenbush, S. W., & Bloom, H. S. (2015). Learning about and from a distribution of program impacts using multisite trials. *American Journal of Evaluation*, *36*(4), 475–499.

Ridgeway, G., McCaffrey, D., Morral, A., Griffin, B. A., & Burgette, L. (2016). *Twang: Toolkit for weighting and analysis of nonequivalent groups.* Retrieved from https://CRAN.R-project.org/package=twang

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, *79*(387), 516–524.

Viechtbauer, W. (2007). Confidence intervals for the amount of heterogeneity in meta-analysis. *Statistics in Medicine*, *26*(1), 37–52.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. Retrieved from https://www.jstatsoft.org/v36/i03/

Zeileis, A. (2004). Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software*, *11*(10), 1–17.