International Journal of Educational Research 95 (2019) 118–130

Contents lists available at ScienceDirect

International Journal of Educational Research

journal homepage: www.elsevier.com/locate/ijedures

Private sector inputs into public school preparation for careers in health care and STEM: An examination of cognitive skills development in the Johnson & Johnson Bridge-to-Employment program in the United States☆

Michael J. Camasso[a], Radha Jagannathan[b,*], Michael J. Bzdak[c]

[a] Rutgers, the State University of New Jersey, School of Environment and Biological Science, Department of Agriculture, Food and Resource Economics, United States

[b] Rutgers, the State University of New Jersey, Edward J. Bloustein School of Planning & Public Policy, United States

[c] Corporate Contributions, Johnson & Johnson, United States

ARTICLE INFO

Keywords:
School-to-career
Corporate social responsibility
STEM enrichment program
Cognitive skills development
Multi-level model

ABSTRACT

We examine the impact of a STEM enhancement program (SEP) sponsored by Johnson & Johnson (J&J) entitled Bridge-to-Employment (BTE). Since 1992, J&J has funded and supervised 80 BTE programs in 19 countries aiming to increase the academic performance of disadvantaged students, enabling these individuals to successfully pursue higher education and careers in the health industry. We study the science, math and language arts grades of 236 BTE and 308 comparison students from ten BTE program sites in the U.S, and find that BTE participation slows the deterioration in math and language arts grades, but has no impact on science grades. We discuss the implications this research has for future evaluations of SEPs and for the formulation of future SEP initiatives.

1. Introduction

Employment projections made by the U.S. Bureau of Labor Statistics show that skilled occupations with the most job growth potential through 2026 are in the areas of science, technology, engineering and mathematics (STEM). Almost all these STEM jobs require either a bachelor's degree (statistical analysts, biomedical engineers, software developers, applied mathematicians, etc.) or an associate degree (web-designers, environment engineering technicians, computer support specialists, technical writers, etc.) and all require a solid grounding in basic mathematics and scientific inquiry (Fayer, Lacey, & Watson, 2017).

Against this backdrop of expanding opportunity in the U.S., is a public-school system that consistently has failed to produce sufficient levels of high quality STEM education at either the elementary or at the high school levels (DeSilver, 2017; Committee on Highly Successful Schools or Programs for K-12 STEM Education, 2011). Since at least the early 1980′s, the math and science scores of American students have consistently lagged behind their European and Asian counterparts (DeSilver, 2017; Organization for Economic Cooperation & Development (OECD), 2017; National Commission on Excellence in Education, 1983). Moreover, the

☆ An earlier version of this paper was presented at the 27th annual meeting of the Society for the Advancement of Socio-Economics (SASE) held at London School of Economics and Political Science (July 2–4), 2015. We would also like to thank Julie Solomon and Pam Carter for their helpful comments on an earlier draft of this paper.

* Corresponding author.
E-mail addresses: mcamasso@sebs.rutgers.edu (M.J. Camasso), radha@rutgers.edu (R. Jagannathan), mbzdak@its.jnj.com (M.J. Bzdak).

https://doi.org/10.1016/j.ijer.2019.01.006
Received 30 May 2018; Received in revised form 20 December 2018; Accepted 22 January 2019
Available online 22 February 2019

mathematics and science performance of black and Hispanic students has been especially disheartening with only 13 percent of black students and 26 percent of Hispanic students meeting mathematics benchmarks on the American College Testing (ACT) national college readiness examination and 11 percent of blacks and 22 percent of Hispanics reaching readiness in science (American College Testing (ACT), 2017). This poor performance in high school has carried over into poor preparation for the STEM labor market. A recent report from the PEW Research Center finds that while blacks make up 11 percent of workforce, they comprise 7 percent of STEM workers with bachelor's degree. For Hispanics, this disparity is even more pronounced, while comprising 16 percent of the workforce, Hispanics account for 6 percent of all STEM workers with a four-year college degree (Graf, Fry, and Funk, 2018).

One approach aimed at addressing the STEM preparation problem in the US has been the STEM enrichment program (SEP). SEPs are typically partnerships of high schools (and less often elementary schools) with higher education institutions, community organizations and business/industry to provide authentic learning activities in STEM subjects. The goal of SEPs is to engage students in a comprehensive system of mentoring, tutoring, enhancement activities and learning opportunities both in and outside classroom (Berger, Turk-Bicakci, & Garett, 2013; Kemple & Willner, 2008; Jobs For the Future, 2017). While some SEPs focus primarily on improving technical skills in math and science, others attempt to improve students' social skills as well as academic knowledge, skills and abilities (Jagannathan, Camasso, Delacalle, 2018; Alvarado & Muniz, 2018).

In this paper, we examine the impact of one SEP, the Johnson & Johnson Bridge-to-Employment (BTE) program on the math and science performance of a panel of 10^{th}, 11^{th}, and 12^{th} grade students from 10 BTE programs located in the states of New Jersey, Pennsylvania, and Delaware. Launched in 1992, BTE has created business-community partnerships in over 80 communities in 19 countries with the objective of preparing students from disadvantaged and minority schools with the academic readiness and career awareness necessary for the successful pursuit of employment in the health industry and in the science, technology, engineering, mathematics, manufacturing and design ($STEM^2D$) labor sectors (Bzdak, 2007; Detgen, 2017). To our knowledge, our study is the first assessment of BTE impact using longitudinal data and one of the very few to employ across-site trajectory analysis of SEP effectiveness (See Kemple & Willner, 2008; Wai, Lubinski, Benbow, & Steiger, 2010; Berger et al., 2013 for several other examples of multi-site, longitudinal evaluations of other SEPs).

2. STEM enrichment programs and minority students

Extra-school programs designed to close the gap in science and mathematics competencies of black and Hispanic students on the one hand, and white and Asian students on the other, have been a steadfast policy response in the US since at least the release of the controversial Coleman report in 1966 (Coleman et al., 1966; Hill, 2017). The genesis of the extra-school SEP, moreover, can be traced to the early 1990′s and American impetus to create a joint education reform and workforce agenda that integrated academic education with vocation instruction and work-based learning (Commission on the Skills of the American Workforce, 1990; Hughes, Bailey, & Mechur, 2001).

The School-To-Work Opportunities Act (STWOA) of 1994 (U.S. 103rd Congress, 1994) demonstrated some commitment by the federal government to strengthen school-to-work transitions; however, this commitment was limited to 29, one time, five year partnership grants with all funding expiring in 2001. The stated goal of STWOA was to support states in their efforts to establish statewide school-to-work transition systems. The Act mandated three core system components: viz., school-based learning linking secondary and post-secondary education, work-based learning providing a planned program of work experience, and connecting activities to ensure coordination between work and school learning by involving employers. Critical, data-driven assessments of STWOA generally support the notion that School-to-Work is a promising strategy to engender work-force readiness and labor force participation (Hamilton & Hamilton, 1999; Hughes et al., 2001). Conclusive evidence of success, however, was not an inference that could be drawn from the numerous evaluations of STWOA. While school-to-work programs helped students better define their career interests and goals for the future (Larson & Vandergrift, 2000), effects on student mathematics or science test scores were mixed. For example, while Maxwell and Rubin (2000) report a positive influence on Grade Point Average (GPA) and grades, Hughes et al. (2001) found no effect. And while employers typically demonstrated great enthusiasm for school-to-work programs, the impact of these programs on college completion and labor market success remained unclear.

One positive effect of STWOA was the creation of a number of career academy model high schools, most notably Sir Francis Drake High School in Marin County, California and the Health and Biosciences Academy in Oakland, California (Steinberg, 1998). Career academies are organized as small learning communities serving approximately 200 students from grades 9 thru 12. Most partner with local businesses to provide work-based learning opportunities. In a rigorous 16-year evaluation of nine (9) academies employing an experimental research design (Kemple & Willner, 2008) conclude that while program increased the earnings of young men by 11 percent, they had no impact on high school completion, post-secondary enrollment or post-secondary academic attainment. Effects on low income, minority students, moreoever, were estimated to be modest.

The Early College High School Initiative (ECHSI) launched by the Gates Foundation in 2002, attempts to enrich the high school education of disadvantaged students by providing the opportunity to pursue a high school diploma while simultaneously earning college credits. In this extra-school model, public or private schools partner with a college/university and community groups to earn one to two years of college credits. Partnerships with business/industry play a smaller role in ECHSI which is guided by the philosophy that a college degree is the best pathway to a high paying job (Jobs For the Future, 2017; Berger et al., 2013).

Anecdotal and descriptive accounts of ECHSI describe a successful initiative that has expanded to 210 early college schools serving 50,000 students in 24 U.S. states (Webb and Mayka, 2011). An impact evaluation of ten (10) early college schools where retrospective lotteries were employed to identify matched comparison groups (Berger et al., 2013) report that ECHSI students were much more likely (than comparison group students) to earn an associate degree (22% vs 1.4%) or a bachelor's degree (2.3% vs 0%).

With respect to grade point average (GPA) or mathematics grades in high school, however, ECHSI students performed no better than their non-ECHSI counterparts. A Jobs For the Future (JFF) study of English language learners also reports no differences in mathematics scores between college bridge programs attendees and non-participants (Jobs For the Future, 2017).

While programs like Career Academies and ECHSI encourage extra-school partnerships that promote STEM, the missions of these programs are broader than SEP and support a more generic definition of college success and/or labor force attachment. Therefore it is quite reasonable to conclude that the small impacts of Career Academies and ECHSI programs on mathematics and science performances are a result of adulterated STEM curriculum, attenuated teaching intensity and related factors. SEPs like the 21st Century Community Learning Centers Program (21st CCLC) (Levine and Zimmerman, 2010); Mathematics, Engineering, Science Achievement (MESA) (Alvarado & Muniz, 2018) and Bridge-to Employment (Family Health International) (Detgen, 2017; FHI-360, 2017) have a more concentrated focus and it is also reasonable to expect that these programs would produce better math and science results for disadvantaged and minority students.

The 21st CCLC program was authorized by the U.S. Congress in 1994 to improve the academic performance of students primarily in minority school districts through before-school, after-school school hours and in-summer tutoring, teaching and mentoring. In a randomized, controlled field experiment conducted by Mathematica Policy Research from 2000 through 2003, the researchers found few impacts when they examined 2308 elementary school students from twenty-six 21st Century Learning Centers. No effects were reported for math or reading test scores or on science, math or reading grades (James-Burdumy, Dynarski, Moor, Deke, and Mansfield, (2005) and Levine and Zimmerman, 2010). Since this report, several states including Texas, New Jersey and Washington have reported that the Century program has led to higher state assessment scores in mathematics and reading (American Institutes for Research, 2018).

MESA currently operates in 12 U.S. states and employs a combination of enrichment activities, academic support and industry involvement to aid disadvantaged high school students on their path toward STEM fields in college and the labor market (MESA (Mathematics, Engineering, Science Achievement), 2017). Using the ongoing High School Longitudinal Study transcript file from 2009 to 2013 (Alvarado & Muniz, 2018) compared the odds of minority students (1) taking advanced placement courses and (2) planning to major in a STEM field in college if they attended a MESA high school or a matched non-MESA school (N = 25, 210). These researchers did not find a MESA effect for either black or Hispanic students on plans to major in STEM and found that only blacks were more likely to take an AP course (7 percent increase).

Unlike 21st CCLC and MESA, BTE has not been the subject of a rigorous impact evaluation. A 2017 report authored by FHI-360, the organization that coordinates BTE activities for Johnson & Johnson found that while 47 percent of BTE graduates plan to pursue career in STEM, only 27 percent of a comparison group announced similar plans. This report also notes that from 2011 to 2016, BTE accounted for a 4 percent increase in grade point average and a 2 percent increase in science grades; however, the data and analysis supporting these outcomes are not provided (FHI-360, 2017). In a second in-house report describing the effect that BTE had on program graduates, Detgen (2017) discusses results from an online survey of 522 students noting that 24 percent said they gained some technical skills and 55 percent said they planned to work in STEM. A low response rate (10 percent) and the absence of any information or respondent – non-respondent differences limit the utility of the survey to determine program impact.

3. The johnson & johnson Bridge-to-Employment program model

The core mission of BTE is to introduce high school students with disadvantaged and minority backgrounds to the expanding array of high skills careers in health care that are science and/or math based (Bzdak, 2007). The vehicle used to carry out this mission is the community-corporate partnership comprising a Johnson & Johnson local operating company, a secondary public school, and an institution of higher education. The BTE model distinguishes between school-to-work and school-to-career, placing emphasis on the educational pathways provided by college degree programs that lead to higher paying health care occupations. BTE also stresses the importance of the programs as a structure within which J&J employees can mentor and tutor students and demonstrate their own sense of social responsibility (Aakhus & Bzdak, 2015).

The program theory underpinning BTE follows a set of propositions that are fundamental to most SEPs: viz., (1) participation will increase STEM exposure and skills, (2) this, in turn, will translate into stronger orientations and help develop technical and soft skill competencies, (3) the results are in alignment with and a commitment to pursue STEM education and careers. Examples of technical skill development are ability to read a technical manual, follow a set of basic assembly instructions, conduct a simple statistical analysis, master basic field observation techniques, conduct a simple lab experiment, write a narrative summary of descriptive data, apply basic knowledge of arithmetic and algebra. Soft skill development includes communication effectiveness, capacity to work within a team, critical thinking, self-control, empathy, and future goal orientation. Mastery of both sets of abilities is deemed necessary in the creation of STEM identities, and essential to the successful pursuit of STEM and healthcare careers.

Although each BTE program takes on the unique character of the local J&J operating company, public school and higher education partners that comprise the collaboration, all programs are required to operationalize this theory by following a common template or logic model. As Frechtling (2007) notes, logic models can be thought of as theories of change which guide the program operations that need to be made, the hypotheses that need to be tested and the empirical predictions that are suggested. An example of this program model guide is shown in Fig. 1.

The guidance suggests several learning activities that could be expected to yield one or more of the short term and/or end-of-grant outcomes listed. Collaborators are, of course, free to propose additional activities with the stipulation that these activities have a direct impact on the outcomes that operationalize the BTE mission.

A typical BTE program receives funding for four years – the first year allows a period of program planning while the subsequent

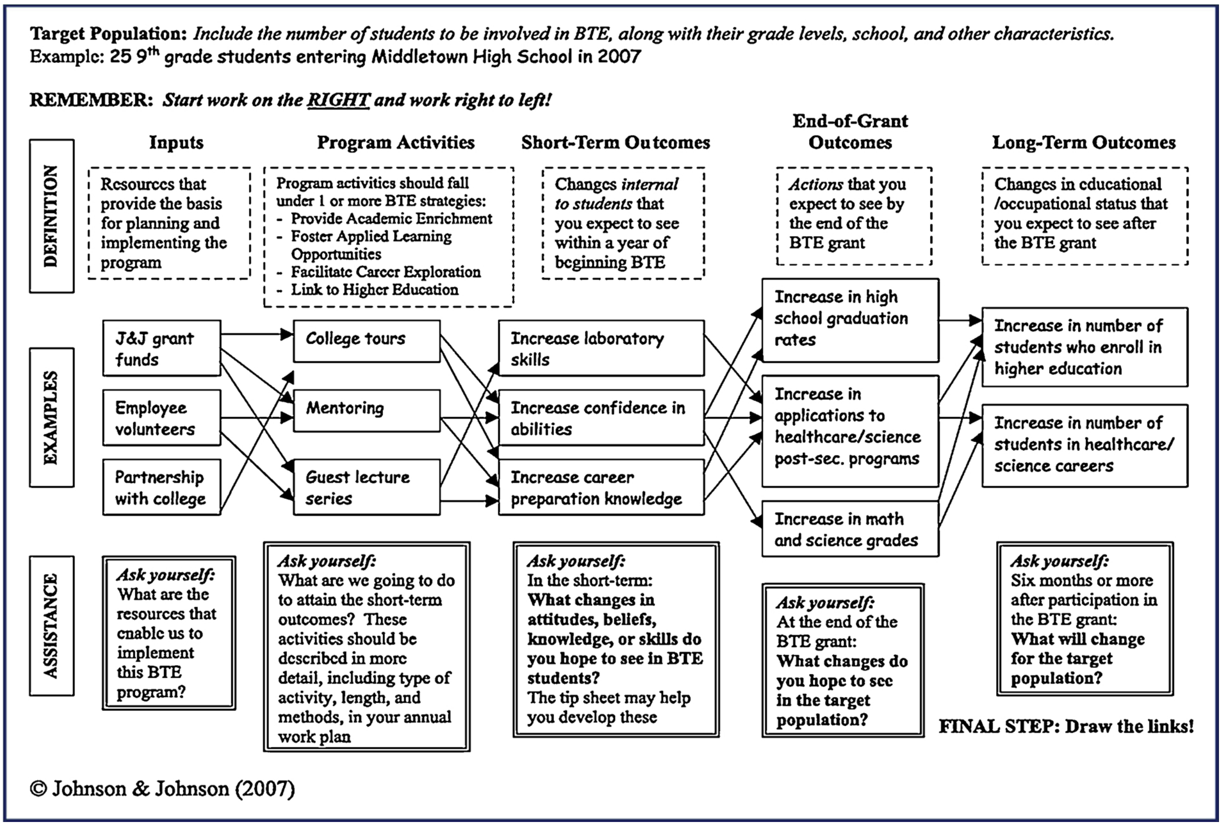

Fig. 1. Bridge to employment program model guide.

years facilitate operations. Students are selected in the 9th grade using an assignment method that the BTE partnership believed would balance the need for a quantitative assessment of impact and that would serve a target group of students who they believed would most benefit from the program. Random assignment was almost never selected; instead, counterfactuals were created by identifying a matched comparison group of students who did not have access to BTE.

Program activities are organized around the three SEP propositions noted earlier. Fig. 1 lists several examples of these propositions in action including guest lectures by health care professionals, mentoring by Johnson & Johnson employees and college tours. Other examples include conducting lab experiments at school and off-campus, math tutoring by college students/health care professionals and science projects that stress fundamental skill development as well as STEM knowledge.

To date quantitative assessments of BTE outcomes have been limited to the two reports outlined above (Detgen, 2017; FHI-360, 2017) and to a series of site-specific assessments, the results of which have been summarized in end-of-year and final reports. The relatively small number of BTE students in each program (thru 2015 this number averaged about 30 in U.S. programs) coupled with issues around BTE student selection has constricted the value of these individual program reports as sources of overall program impact. Absent any type of impact analysis, assessments of a more general BTE effect have relied on qualitative indicators culled from focus groups, mentor surveys, and anecdotal information. (See for example, Brooks, MacAllum, & McMahon, 2005; FHI-360, 2017).

4. Structure of this evaluation

One requirement for the receipt of BTE grant support by a local partnership is the commitment to "use data to continuously improve" program operations (FHI-360, 2017, p.29). Each BTE program site must contract with an outside evaluator (university, consulting firm, etc.), provide the evaluator with a set of specified academic and student opinion data and agree to incorporate evaluation findings into future programing when feasible. For their part, the independent evaluators are required to submit a yearly report that summarizes these data and assess how well the program has met the short term, mid-range, and long term objectives outlined in the partners' logic model.

The evaluation design used by most of the independent evaluators is the nonequivalent comparison group approach (Shadish, Cook, & Campbell, 2002). BTE and comparison group students are observed at the beginning of each academic year on a set of academic, attitudinal, and behavioral measures (O_{T1}, O_{C1}); these same measures are then repeated at the end of each school year (O_{T2}, O_{C2}). Impact is estimated by (O_{T2} - O_{T1}) - (O_{C2} - O_{C1}), i.e., as a simple difference-in-difference. The groups are assumed nonequivalent because of the lack of random assignment. Between 1999 and 2015, Rutgers University was selected to evaluate the impact of 17 local BTE programs. Eleven of these site evaluations yielded three full years of impact data, six did not. Two BTE programs operated in Cincinnati, Ohio were structured as two-year programs, and programs in Trenton, New Jersey and Bound Brook, New Jersey were unable to generate three years of data because of failures to deliver their services in one or more years. Programs in Christiana, Delaware and Kennett Square, Pennsylvania, while conducting three years of programming, were unable to maintain their comparison groups. Specific information on these sites is provided in Appendix A. As independent evaluators, the Rutgers University researchers were not privy to individual student identifiers. The partners at each site, in addition to the assignment of students in BTE and comparison groups, obtained parental consent for student participation and grade release through their Institutional Review Board (IRB) process.

Our focus in this paper is on changes in science, mathematics and language arts grades and overall grade point average (GPA) of 286 BTE and 349 Comparison students from across 10 sites where three years of academic performance data was collected for these annual reports. One site (Bridgewater, New Jersey) was excluded because the program did not provide services to disadvantaged students.

To analyze these data we employ hierarchical linear models (HLM), also known in the literature as latent trajectory models. The HLM is designed to explicitly recognize nested or repeated measures data structures, and permits straightforward examination of both *intra*-unit (within student) change overtime and *inter*-unit (between students) variability in intra-unit change (Curran & Hussong, 2003; Bollen and Curran, 2006; Singer & Willett, 2003). Our trajectory analysis has several advantages over site-specific analysis of academic performance. The pooled data, which are the results of structuring a cross-site analysis improves statistical power and reduces the probability of making Type II statistical errors. Also, modelling the performance trajectories provides a stronger test of any treatment-comparison group difference than do post-intervention only or difference-in-difference analyses inasmuch as intervening period data for students' outcomes are incorporated into analysis and are not simply discarded.

4.1. Sample characteristics and study variables

In Table 1, we provide descriptive data on the BTE and Comparison group students from the ten study sites. For each site, we present the BTE program location, years of operation, size of the BTE and Comparison groups, and information on demographic and academic performance at baseline. All demographic and academic data were obtained from data collection systems maintained in the school for the purposes of producing student report cards and/or reporting student-level information to state-level departments of education. While significant differences between groups on measured demographics do not appear to pose a selection problem, this is not the case for academic performance measures in a number of sites. These differences are almost always in favor of students enrolled in the BTE program who begin the program with higher

science, math and language arts grades and higher overall GPAs.

Table 2 provides the reader with an overall summary of student demographics and academics, employing data pooled from all ten sites. Here we see that White and Black students are more likely to populate BTE groups while Hispanic students are underrepresented. As shown in the previous Table, BTE students demonstrate higher math and science grades and higher GPAs than do Comparison group students. It is clear from Tables 1 and 2 that any attempt to draw inferences regarding BTE impact on academic performance must, at minimum, take into consideration statistically these measured differences at baseline.

Since we are examining student trajectories overtime and are not using simple pre-post, treatment-comparison group difference models, our data are structured as student-year observations. If data were available for all demographic and academic variables in all study periods, i.e., at baseline, end of years one, two and three, for all BTE students (236) and for all Comparison group students (308), we would expect 2176 student-year observations.

Entries in Table 3 show that there is missing data, especially in the case of race which reduces the actual number of student-year observations available for analysis by about 18 percent.

4.2. Analytic approach

We examine trajectories of students' academic outcomes with multilevel models estimated by the method of maximum likelihood using Stata's xtmixed command (Version 15). These models permit straightforward examination of both *intra*-unit (within student) change in outcomes over time and *inter*-unit (between students) variability in intra-unit change. Further, these overtime changes can be conditioned on one or more predictor variables. Here, we estimate two-level models, where the first level investigates *within* student changes overtime in their academic outcomes, i.e., their academic trajectories, and the second level explores if these individual trajectories are altered by participation or non-participation in the BTE program.

We estimate five different specifications, starting with a simple *unconditional means* only model (Model 1), followed by an *unconditional growth* model (Model 2) – these two models provide a useful baseline for comparison with our subsequent models (Models 3–5) that incorporate demographic, treatment group and site predictors. These unconditional models decompose the outcome variability into (a) across students irrespective of time and (b) across both students *and* time, and help establish whether there is predictable variability in the outcome that warrants an investigation and if so, whether this variability exists within or between individuals (Singer & Willett, 2003). The unconditional models are systematically augmented with predictors, with Model 3 introducing BTE participation, Model 4 examining the BTE effect while controlling for student gender and race, and finally Model 5 that looks at any BTE effect while also controlling for site-specific, (fixed effect) time invariant differences.

Model 1 is specified as follows, with a Level 1 equation that models the observed outcome as a function of the individual-specific true mean and its deviation at time *t,* while Level 2 examines how this individual-specific mean varies from the grand mean:

Table 1
Sample characteristics at baseline by program status and BTE site2.

Characteristic	Ambler		Bound Brook		Franklin Township		New Brunswick (1) HSTS		New Brunswick (1) NBHS	
	BTE (N = 50)	Comparison (N = 32)	BTE (N = 19)	Comparison (N = 11)	BTE (N = 21)	Comparison (N = 79)	BTE (N = 12)	Comparison (N = 11)	BTE (N = 12)	Comparison (N = 30)
Demographic %										
Female	60.0	61.2	73.7	72.7	62.0*	39.0	54.5	75.0	41.7*	80.0
White	74.5	74.2	10.5*	36.4	0.0	9.1	9.1	10.0	0.0	0.0
Black	19.1	16.1	5.3	9.1	76.5	63.6	27.3	20.0	30.0	14.3
Hispanic	2.1	3.2	63.2	54.5	17.6	18.2	63.6	40.0	70.0	78.6
Other	4.3	6.5	21.1	0.0	5.9	9.1	0.0*	30.0	0.0	7.1
Academic Outcomes Mean (Std.dev.)										
Math grade	89.4 (7.8)	87.0 (11.4)	71.9 (11.8)	73.5 (13.5)	75.8 (9.7)*	69.7 (8.7)	85.3 (7.6)	82.5 (8.6)	68.1 (27.8)*	79.2 (10.9)
Language Arts grade	87.6 (8.3)	90.4 (6.0)	73.7 (11.6)	74.0 (11.3)	81.9 (11.5)	79.1 (7.7)	85.5 (4.7)	83.6 (5.2)	75.6 (11.8)	80.2 (13.7)
Science grade	87.7 (9.4)	87.1 (9.2)	73.7 (9.0)	74.5 (11.0)	82.0 (8.6)*	74.8 (2.9)	82.3 (2.5)	82.8 (6.1)	81.1 (5.1)*	87.0 (8.9)
GPA	88.2 (7.6)	88.2 (8.5)	73.1 (8.9)	74.0 (10.2)	79.9 (8.9)*	74.5 (4.6)	84.4 (4.4)	83.0 (6.0)	74.9 (11.4)*	82.1 (9.7)

Characteristic	New Brunswick (2) HSTS		New Brunswick (2) NBHS		North Plainfield		Trenton		Wilmington	
	BTE (N = 23)	Comparison (N = 33)	BTE (N = 18)	Comparison (N = 19)	BTE (N = 44)	Comparison (N = 55)	BTE (N = 26)	Comparison (N = 48)	BTE (N = 31)	Comparison (N = 30)
Demographic %										
Female	73.9	64.5	61.1	42.1	68.2	67.9	88.5*	62.5	61.3*	33.3
White	0.0	0.0	0.0	0.0	9.1	5.7	0.0	0.0	3.2	20.0
Black	30.4	15.2	7.7	10.5	34.1	32.1	69.2	66.7	77.4	70.0
Hispanic	69.6	81.8	92.3	89.5	52.3	56.6	19.2	22.9	19.4	10.0
Other	0.0	3.0	0.0	0.0	4.5	5.7	11.5	10.4	0.0	0.0
Academic Outcomes Mean (Std.dev.)										
Math grade	79.6 (9.3)	82.5 (14.4)	85.7 (10.9)	86.1 (8.2)	81.4 (7.4)*	70.1 (28.3)	77.5 (11.6)	79.5 (19.2)	76.3 (10.1)*	70 (11.5)
Language Arts grade	87.7 (4.5)*	85.1 (4.3)	85.9 (6.2)*	79.2 (8.9)	84.7 (6.8)	82.0 (9.2)	84.7 (8.5)	86.5 (12.2)	78.2 (8.8)*	69.2 (7.9)
Science grade	84.3 (6.5)*	79.2 (5.4)	86.3 (9.2)*	76.9 (9.4)	85.8 (8.0)*	80.5 (10.1)	77.7 (9.5)	81.1 (17.5)	76.3 (9.2)*	69.4 (9.6)
GPA	83.9 (5.9)	82.2 (5.8)	86.0 (7.9)*	80.7 (7.3)	84.0 (6.3)*	77.5 (13.8)	80.0 (8.2)	82.4 (15.2)	76.9 (7.7)*	69.5 (6.6)

[1]Maximum N shown - it may vary from variable to variable within each group.
[2]For Bound Brook and Wilmington, the data shown here are from Year 1 because baseline data were not available.
*indicates significant group differences at baseline.

$$\text{Level 1: } Y_{it} = \pi_{0i} + \varepsilon_{it} \tag{1.1}$$

$$\text{Level 2: } \pi_{0i} = \Upsilon_{00} + \zeta_{0i} \tag{1.2}$$

where

Y_{it} represents a particular academic outcome (e.g., school grade in math, language arts, science, and overall GPA) for student *i* at time *t*,
π_{0i} is the individual-specific mean outcome,
ε_{it} is the deviation of the observed outcome from the individual-specific mean,
Υ_{00} is the grand mean, and
ζ_{0i} is the deviation of individual-specific mean from the grand mean.

We assume that the Level 1 and Level 2 residuals (ε_{it} and ζ_{0i}) are normally distributed, both with mean 0, and variance σ_ε^2 and σ_0^2 respectively, so that σ_ε^2 provides an estimate of the variability in the outcome of each individual around his/her own mean, and σ_0^2

Table 2
Sample characteristics at baseline by program status - All sites combined.

Characteristic	BTE (N = 236)	Comparison (N = 308)
Demographic		
%		
Female	65.5	58.5
White*	18.6	12.0
Black*	41.4	31.8
Hispanic*	36.4	48.9
Other	3.6	7.3
Academic Outcomes		
Mean (Std.dev.)		
Math grade*	80.9 (11.9)	78.4 (13.4)
Language Arts grade	84.0 (9.0)	82.9 (9.8)
Science grade*	83.1 (9.7)	80.2 (10.5)
GPA*	82.6 (8.8)	80.5 (9.7)

[1]Maximum N shown - it may vary from variable to variable within each group.
* indicates significant group differences at baseline.

Table 3
Distribution of study variables across student-year observations.

Variable	N	Mean	Standard Deviation	Minimum	Maximum
Female	1,926	0.62	0.49	0	1
Race	1,787				
White		0.18	0.38	0	1
Black		0.35	0.48	0	1
Hispanic		0.41	0.49	0	1
Other		0.06	0.23	0	1
Math grade	1,949	78.6	12.4	20	100
Language Arts grade	1,949	82.3	9.9	24	100
Science grade	1,949	80.3	11.3	25	100
GPA	1,949	80.4	9.5	41.5	100
Site coordination	1,772	7.17	1.47	3	8.5

Observations from each site: Ambler 17.6%, Bound Brook 3.2%, Franklin Township 13.9%, New Brunswick (1) HSTS 5.0%, New Brunswick (1) NBHS 7.5%.
New Brunswick (2) HSTS 10.1%, New Brunswick (2) NBHS 6.7%, North Plainfield 16.9%, Trenton 10.1%, Wilmington 9.0%.

summarizes the variability of individual-specific means around the grand mean.

Since the Level 2 equation cannot be estimated directly because of the structural parameter π_{0i}, we substitute (1.2) into (1.1) to obtain the reduced-form model for the observed responses Y_{it}, with one fixed effect (Υ_{00}) and a composite residual as follows:

$$Y_{it} = \Upsilon_{00} + (\zeta_{0i} + \varepsilon_{it}) \tag{1.3}$$

Crowder and Hand (1990) refer to the fixed component as the “immutable constant of the universe,” to ζ_{0i} as the “lasting characteristic of the individual” and to ε_{it} as the “fleeting aberration of the moment.”

Model 2 estimates an unconditional growth model that introduces the predictor ‘Time’ at Level 1, allowing each student to have a distinct growth rate or trajectory π_{1i}, and enables us to examine whether inter-individual differences emanate from differences in the mean or the growth rate. Level 1, Level 2 and the reduced-form equations are specified as follows:

$$\text{Level 1: } Y_{it} = \pi_{0i} + \pi_{1i}\,\text{Time}_{it} + \varepsilon_{it} \tag{2.1}$$

$$\text{Level 2: } \pi_{0i} = \Upsilon_{00} + \zeta_{0i} \tag{2.2a}$$

$$\pi_{1i} = \Upsilon_{10} + \zeta_{1i} \tag{2.2b}$$

$$\text{Reduced-form: } Y_{it} = (\Upsilon_{00} + \Upsilon_{10}\,\text{Time}_{it}) + (\varepsilon_{it} + \zeta_{0i} + \zeta_{1i}\,\text{Time}_{it}) \tag{2.3}$$

We now have an additional structural parameter π_{1i} and a corresponding Level 2 Eq. (2.2b) that estimates inter-individual differences in the rates of change or growth trajectories. The fixed effects Υ_{00} and Υ_{10} now estimate the mean intercept and mean growth rate, respectively; ζ_{0i} and ζ_{1i} are the deviations of each student from the group mean intercept and group mean growth rate; and the Level 1 residuals ε_{it} now tell us the individual deviation from his/her true growth trajectory. We continue to assume that both the Level 1 and Level 2 residuals have a normal distribution, with ζ_{0i} and ζ_{1i} now bivariate normal with mean 0 and variance σ_0^2 and σ_1^2. In addition, the covariance (σ_{01}) between ζ_{0i} and ζ_{1i} is also estimated in this model.

In Model 2, we have made the assumption that the time and individual-specific values of the outcome (Y_{it}) are completely governed by the underlying trajectory process and any deviations of these values from the trajectory are treated as error. We now extend these models to capture situations in which we do not necessarily anticipate that the growth rates in outcomes are completely determined by the underlying trajectory process; rather they are related only partly to the trajectory process but may also be influenced by their participation in the BTE program. We study the BTE effect in Model 3, and examine how the BTE effect changes when additional predictors are added in Models 4 and 5. In light of our quasi-experimental design, we consider Model 5, which controls for both student demographic characteristics and site-specific factors that remain time invariant, to be our final model.

Level 1, Level 2 and the composite specifications of Model 3 are as follows:

$$\text{Level 1: } Y_{it} = \pi_{0i} + \pi_{1i}\,\text{Time}_{it} + \varepsilon_{it} \tag{3.1}$$

$$\text{Level 2: } \pi_{0i} = \Upsilon_{00} + \Upsilon_{01}\,\text{BTE} + \zeta_{0i} \tag{3.2a}$$

$$\pi_{1i} = \Upsilon_{10} + \Upsilon_{11}\,\text{BTE} + \zeta_{1i} \tag{3.2b}$$

$$\text{Reduced-form: } Y_{it} = (\Upsilon_{00} + \Upsilon_{01}\,\text{BTE} + \Upsilon_{10}\,\text{Time}_{it} + \Upsilon_{11}\,\text{BTE*Time}_{it}) + (\varepsilon_{it} + \zeta_{0i} + \zeta_{1i}\,\text{Time}_{it}) \tag{3.3}$$

Model 3 now includes BTE participation as a predictor of both the initial or baseline outcome levels as well as the growth (change) in the outcomes. The Model contains four fixed effects, Υ_{00}, the level of initial outcome of the average Comparison group student; Υ_{01}, the difference in the initial outcome level between BTE and Comparison students; Υ_{10}, the growth rate of the average comparison student; and finally Υ_{11}, the difference in the growth rate between the BTE and comparison students, which is the coefficient of interest that provides BTE program impact. The random effects parameters are specified as before.

Equations for Models 4 and 5 closely follow the specification used for Model 3, except in Level 2, we add demographic controls in Model 4 and site-specific controls in Model 5. To assess model fit and improvement in model fit across models, we use the likelihood ratio test and the deviance statistic, respectively.

5. Results

The application of our multi-level analysis to BTE academic data indicates these principal findings:

1 While the math grades of comparison group students declined at a rate of 1.11 points per year, the grades of BTE students experienced a decline of 0.24 points per year. This difference is statistically significant.
2 There was no effect of BTE on changes in science grades.
3 While the language arts scores of comparison group students declined at a rate of 0.77 points per year, the grades of BTE students experienced a decline of 0.36 points. This difference is statistically significant.
4 There was no effect of BTE on Grade Point Average (GPA) at the 5 percent significance level.

Table 4 presents results of fitting multi-level trajectory models for math grades. Model 1 (Eqs. (1.1)–(1.3)) shows that the only fixed effects parameter in the model (Υ_{00}), the average math score across all students over all time periods, is 79.46, and is significantly different from zero. The random effects $\sigma_\varepsilon{}^2$ and $\sigma_1{}^2$ provide an estimate of the variability[1] in math grades within and across students, and indicate that there is a significant amount of unexplained variability paving the way for inclusion of predictors. These variance estimates can also be used to calculate an intra-class correlation coefficient, which provides us with an indication of how much variability in math grades is due to differences across students (Singer & Willett, 2003). Model 1 indicates that nearly 50 percent[2] of the variability in math grades is attributable to differences across students.

Model 2 presents the results of the unconditional growth model (Eqs. (2.1)–(2.3)) where the two fixed effects Υ_{00} and Υ_{10} tell us that the estimated average starting point in math grade was 80.26, which was declining over time at a rate of 0.56. The estimated level one residual's standard deviation of 8.3 (σ_ε) shows the amount of average deviation of individual math grades from his/her own linear change trajectory, and when compared to Model 1, indicates that about 7 percent of the within-person variability in math grades (=(8.89–8.3)/8.89) is systematically related to *Time*, with a significant portion of the variability still unexplained. The level two residuals' standard deviations of 10.68 and 2.54 summarize between-individual differences in the starting point and the rates of change and their statistical significance suggests that there is still a substantial amount of unexplained variability in both the starting point and the growth rate and that there is benefit in adding substantive predictors to the model. The Model also estimates that the correlation between the level 2 residuals (σ_{01}) is -0.64, indicating that the relationship between the true starting point and the rate of change in math grades is significant and negative, that is, student scores in math that are higher in the beginning decline less rapidly over time.

In Model 3, we add BTE participation as a substantive predictor in both the initial level of math grades and their growth over time, to assess whether the program served to shift the average trajectory upwards, or if it at least slowed down the decline in math grades.

[1] The Table shows the standard deviation, rather than the variance.

[2] Since the total variability in math grades is the sum of two variance components - within and between variability, we can calculate the intra-class correlation, or that portion of the variability that is due to differences across individuals as: $\sigma_0^2/(\sigma_0^2 + \sigma_\varepsilon{}^2)$.

Table 4
Multi-level regression model for math grade.

Fixed Effects		Parameter	Model 1 Coefficient (Robust Std. Error)	Model 2 Coefficient (Robust Std. Error)	Model 3 Coefficient (Robust Std. Error)	Model 4 Coefficient (Robust Std. Error)	Model 5 Coefficient (Robust Std. Error)
Initial Status (π_{0i})	Intercept	Υ_{00}	79.46 ***	80.26 ***	79.57 ***	82.04 ***	86.35 ***
			(0.44)	(0.59)	(0.86)	(1.17)	(4.28)
	BTE	Υ_{01}			1.69 *	1.57 *	1.48
					(1.18)	(1.16)	(1.12)
Rate of change (π_{1i})	Intercept	Υ_{10}		−0.56 ***	−0.94 ***	−0.98 ***	−1.11 ***
				(0.23)	(0.34)	(0.34)	(0.36)
	BTE	Υ_{11}			0.77 **	0.83 **	0.87 **
					(0.40)	(0.47)	(0.47)
Demographic controls [a]			No	No	No	Yes	Yes
Site fixed effects [b]			No	No	No	No	Yes
Random Effects		**Parameter**	**Model 1 Estimate (Std.Error)**	**Model 2 Estimate (Std.Error)**	**Model 3 Estimate (Std.Error)**	**Model 4 Estimate (Std.Error)**	**Model 5 Estimate (Std.Error)**
Level 1	Within person	σ_ε	8.89 ***	8.30 ***	8.27 ***	8.27 ***	8.28 ***
			(0.30)	(0.33)	(0.33)	(0.33)	(0.33)
Level 2	Initial status	σ_0	8.77 ***	10.68 ***	10.67 ***	10.41 ***	9.96 ***
			(0.46)	(0.59)	(0.59)	(0.61)	(0.66)
	Rate of change	σ_1		2.54 ***	2.61 ***	2.63 ***	2.61 ***
				(0.52)	(0.51)	(0.51)	(0.54)
	Correlation	σ_{01}		−0.64 ***	−0.66 ***	−0.65 ***	−0.63 ***
				(0.07)	(0.07)	(0.07)	(0.09)
Deviance			12,303.70	12,271.77	12,256.13	12,240.30	12,213.36
P (LR Chisquared Test)			0.00	0.00	0.00	0.00	0.00
n			1,608	1,608	1,608	1,608	1,608

One tailed tests; *** p-value < 0.01; ** p-value < 0.05; * p-value < 0.10.
[a] Models 4 and 5 control for race and gender.
[b] Model 5 controls for site specific charecteristics that are time invariant.

The estimated fixed effects for levels of math grade reported in the top panel of the Table show that the average initial math grade for the Comparison group students was 79.57, while for the BTE students it was 1.69 points higher. The estimated growth parameters indicate that while the average Comparison student experienced a significant decline in math grade at a rate of 0.94, the average BTE student had a significantly lower rate of decline per year of 0.14 (= −0.94 + 0.77).

The estimate of the within-variance component (σ_ε) in Model 3 remains similar to that of Model 2 indicating that the model could benefit from the inclusion of time-varying predictors; however, the unavailability of these data preclude us from pursuing this option. Estimates of the Level 2 between-variance components also remain significant and about the same as the previous model suggesting the inclusion of additional predictors for both the level and trajectory in math grades.

Results from Model 4 that includes the students' personal characteristics of gender and race are very similar to that of Model 3, again reinforcing the benefit BTE students accrue from program participation. Model 5 adds site fixed effects, that is, characteristics specific to each BTE site that remain time-invariant. This final model shows that math grades for the Comparison group were declining at an average rate of 1.11 points, but BTE students' scores were declining at a much slower rate of 0.24 (= −1.11 + 0.87). Estimates of both the within- and between-variance components continue to indicate the presence of significant unexplained variance at both levels, and the desirability of including additional predictors, a luxury that our dataset does not permit.

All five models show good fit as indicated by the significant likelihood ratio test. Progressive reductions in the deviance statistic in each model relative to the previous model point to the usefulness of the predictors added. The final model (Model 5) also shows considerable reductions in the within-individual and between-individual error variances relative to the baseline unconditional models (Models 1 and 2), confirming the conclusions indicated by the deviance statistic with respect to improvements in model fit.

In Table 5, we provide the results from our multi-level regression analyses of changes in science grades. In Model 3, it is clear that BTE participation is a significant predictor of the initial level of science grades; the average initial science grade for Comparison group students was 80.55 and for BTE students it was 2.63 points higher. The estimated growth parameter, however, indicates no difference between Comparison and BTE students (−0.24 + 0.23 ≈ 0). The addition of covariates

(Model 4) and site-fixed effects (Model 5) does nothing to change the inference.

When we fit our series of multi-level models for language arts grades (Table 6), we once again find an effect of BTE participation. As in the case of mathematics grades BTE appears to slow down the negative trajectory of grade performance. Model 3, for example, shows that the average initial grade for Comparison group students was 83.26, and for BTE students it was 1.52 higher. Examination of the growth parameters estimated from this model demonstrate a significant decline in language arts scores over time at a rate of 0.77 for the Comparison students, with BTE student scores declining at a significantly lower rate of 0.36 (= −0.74 + 0.43). This difference in trajectory remains significant when covariates and fixed-effects are added.

Table 5
Multi-level regression model for science grade.

Fixed Effects		Parameter	Model 1 Coefficient (Robust Std. Error)	Model 2 Coefficient (Robust Std. Error)	Model 3 Coefficient (Robust Std. Error)	Model 4 Coefficient (Robust Std. Error)	Model 5 Coefficient (Robust Std. Error)
Initial Status (π_{0i})	Intercept	Υ_{00}	81.23 *** (0.42)	81.71 *** (0.50)	80.55 *** (0.71)	83.78 *** (1.07)	79.50 *** (3.71)
	BTE	Υ_{01}			2.63 *** (0.95)	2.53 *** (0.93)	2.48 *** (0.91)
Rate of change (π_{1i})	Intercept	Υ_{10}		−0.34 ** (0.18)	−0.24 (0.27)	−0.28 (0.29)	−0.26 (0.28)
	BTE	Υ_{11}			0.23 (0.36)	0.18 (0.36)	−0.14 (0.36)
Demographic controls [a]			No	No	No	Yes	Yes
Site fixed effects [b]			No	No	No	No	Yes

Random Effects		Parameter	Model 1 Estimate (Std.Error)	Model 2 Estimate (Std.Error)	Model 3 Estimate (Std.Error)	Model 4 Estimate (Std.Error)	Model 5 Estimate (Std.Error)
Level 1	Within person	σ_ε	7.44 *** (0.30)	7.22 *** (0.32)	7.230 *** (0.30)	6.52 *** (0.31)	7.29 *** (0.32)
Level 2	Initial status	σ_0	8.61 *** (0.35)	8.88 *** (0.63)	8.74 *** (0.63)	7.84 *** (0.67)	8.33 *** (0.67)
	Rate of change	σ_1		1.33 ** (0.70)	1.38 ** (0.68)	0.57 *** (0.22)	1.36 ** (0.71)
	Correlation	σ_{01}		−0.21 (0.25)	−0.21 (0.25)	−0.81 *** (0.04)	−0.32 (0.23)
Deviance			11,862.44	11,856.17	11,847.81	11,377.59	11,802.60
P (LR Chisquared Test)			0.00	0.00	0.00	0.00	0.00
n			1,608	1,608	1,608	1,608	1,608

One tailed tests; *** p-value < 0.01; ** p-value < 0.05; * p-value < 0.10.
[a] Models 4 and 5 control for race and gender.
[b] Model 5 controls for site specific charecteristics that are time invariant.

Finally, we see in Table 7 that BTE participation also retards the decline in student GPA. Estimated growth parameters in Model 3 show that students in the Comparison group sustain an annual decline of 0.66 points while BTE students have a slower rate of decline which is about half that of the Comparison group's (−0.66 + 0.33 = −0.33). This difference, while not statistically significant in Model 3 reaches significance in Models 4 and 5 at the 10 percent level.

It is important to note that while the inclusion of covariates and site-specific fixed effects somewhat reduce the estimated within and between variance components shown in Tables 4 thru 7, a great deal of unexplained variance of both kinds still remains. Identification and inclusion of other sources of this variability could substantially increase the predictive validity of the analyses we have undertaken.

6. Discussion and conclusions

Extra-school, SEP programs like BTE attempt to address the serious problem in the US of blacks and Hispanic under-representation in STEM education, degree programs and occupation clusters. Notwithstanding decades of operation in many cases, and new initiatives being funded without intermission (see for example the Green Ribbon Schools program (U. S. Department of Education Green Ribbon Schools, 2018) and the Minority Science and Engineering Program (U.S. Department of Education, 2016), SEPs remain under-investigated (Alvarado & Muniz, 2018; Levine and Zimmerman, 2010).

The Johnson & Johnson Bridge-to-Employment (BTE) program is a SEP design to enhance the STEM and health care science orientations, interests and skill competencies of disadvantaged high school students with the objective of increasing student commitment to pursue higher paying careers and occupations in these labor sectors. In this paper we present some evidence that BTE through its program of STEM exposure, skills enhancement and competency building can impact the academic performance of disadvantaged students. BTE students performed significantly better than their Comparison group counterparts in mathematics and language arts but not in science. While the findings with respect to science are discouraging, the math results are encouraging and run contrary to findings of other SEP evaluations (Berger et al., 2013; James-Burdumy et al., 2005; Kemple & Willner, 2008). The improvement in math performance, however, is not a result of a more positive trajectory in academic grades but rather is a consequence of slower deterioration of performance compared to students who were not exposed to BTE. Our finding that the math, science and language arts grades of samples of (primarily) disadvantaged students decline with age is not unexpected. Similar results have been reported by Hanushek and Rivkin (2009); Alexander, Entwisle, and Olsen, (2007), and Downey, VonHippel, and Broh, (2004).

In their 25-year study of achievement in STEM, Wai et al. (2010), conclude that disparities in science and math performance

Table 6
Multi-level regression model for language arts grade.

Fixed Effects		Parameter	Model 1 Coefficient (Robust Std. Error)	Model 2 Coefficient (Robust Std. Error)	Model 3 Coefficient (Robust Std. Error)	Model 4 Coefficient (Robust Std. Error)	Model 5 Coefficient (Robust Std. Error)
Initial Status (π_{0i})	Intercept	Υ_{00}	83.05 ***	82.89 ***	83.26 ***	83.66 ***	79.58 ***
			(0.36)	(0.34)	(0.65)	(0.94)	(3.26)
	BTE	Υ_{01}			1.52 *	1.40 *	1.56 **
					(0.88)	(0.86)	(0.85)
Rate of change (π_{1i})	Intercept	Υ_{10}		−0.40 ***	−0.77 ***	−0.79 ***	−0.74 ***
				(0.14)	(0.22)	(0.22)	(0.23)
	BTE	Υ_{11}			0.41 **	0.44 **	0.43 *
					(0.24)	(0.27)	(0.29)
Demographic controls [a]			No	No	No	Yes	Yes
Site fixed effects [b]			No	No	No	No	Yes

Random Effects		Parameter	Model 1 Estimate (Std.Error)	Model 2 Estimate (Std.Error)	Model 3 Estimate (Std.Error)	Model 4 Estimate (Std.Error)	Model 5 Estimate (Std.Error)
Level 1	Within person	σ_ε	6.55 ***	6.87 ***	6.50 ***	6.52 ***	6.53 ***
			(0.31)	(0.30)	(0.30)	(0.31)	(0.31)
Level 2	Initial status	σ_0	7.42 ***	7.97 ***	8.10 ***	7.84 ***	7.83 ***
			(0.49)	(0.40)	(0.66)	(0.67)	(0.71)
	Rate of change	σ_1		0.30 ***	0.57 ***	0.57 ***	0.69 ***
				(0.01)	(0.21)	(0.22)	(0.27)
	Correlation	σ_{01}		−0.80 ***	−0.81 ***	−0.81 ***	−0.82 ***
				(0.04)	(0.05)	(0.04)	(0.05)
Deviance			11,434.32	11,416.26	11,400.02	11,377.59	11,363.88
P (LR Chisquared Test)			0.00	0.00	0.00	0.00	0.00
n			1,608	1,608	1,608	1,608	1,608

One tailed tests; *** p-value < 0.01; ** p-value < 0.05; * p-value < 0.10.
[a] Models 4 and 5 control for race and gender.
[b] Model 5 controls for site specific charecteristics that are time invariant.

emerge early in elementary school and worsen over time. These authors note that by the tenth grade, black and Hispanic students are more likely that their white and Asian peers to filter into low education tracks and less likely to pursue STEM courses. Hill (2017) reports that growth curve analyses indicate learning increases more in elementary school than it does in middle school and that this deceleration is most pronounced in poorer school districts.

BTE illustrates an SEP where business/industry guides the career education of disadvantaged students in collaboration with willing partners from public schools and institutions of higher learning. Scaling up this type of partnership to a level that could impact our nation's STEM and healthcare labor force will require a dramatic change in how many people in US government, the labor movement, and the education community view career education and/or job training that is directed from the private sector. In our view rigorous evaluations indicating improvement in student' cognitive skills could dramatically reduce the skepticism and resistance.

Of course, this research has some notable limitations. The design does control for some student characteristics and site level factors that are time invariant, but does little to adjust for unmeasured, time-varying factors at either the student or site level. The selection of students by the partners in more instances than not, resulted in Comparison groups with significantly lower grade performance levels at baseline. It is quite possible that these group differences may signal dissimilarities in resource inputs (in-school, outside school or both) that are correlated with BTE participation and that could also vary over time. As we attempted to make clear in our presentation of results, a large proportion of the variation in grades remains unexplained. The inclusion of human capital inputs from the school and home background factors/resources measured at the student and school levels would very likely reduce this unexplained variation and make our estimates of BTE effect more precise. Family structure and values information would seem indispensable (Coleman et al., 1966; Hill, 2017). Expanded statistical modelling and sensitivity analysis is not the only pathway to more precise BTE estimates. Experimental design with random assignment of students into BTE and control groups would reduce the problem of selection and would decrease the potential influence of covariates (both time invariant and time changing) on any BTE-grade performance relationship.

Finally, our study raises an obvious follow-up research question: Would a more intense BTE that begins earlier in a student's education increase academic performance rather than simply slow down performance declines? We believe efforts to answer this question are important to explore before BTE or other current SEPs are "scaled up" in their current form.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the

Table 7
Multi-level regression model for GPA.

Fixed Effects		Parameter	Model 1 Coefficient (Robust Std. Error)	Model 2 Coefficient (Robust Std. Error)	Model 3 Coefficient (Robust Std. Error)	Model 4 Coefficient (Robust Std. Error)	Model 5 Coefficient (Robust Std. Error)
Initial Status (π_{0i})	Intercept	Υ_{00}	81.32 ***	82.01 ***	81.30 ***	83.01 ***	83.00 ***
			(0.38)	(0.45)	(0.65)	(0.95)	(3.01)
	BTE	Υ_{01}			1.73 **	1.64 **	1.66 **
					(0.90)	(0.88)	(0.86)
Rate of change (π_{1i})	Intercept	Υ_{10}		−0.49 ***	−0.66 ***	−0.67 ***	−0.73 ***
				(0.13)	(0.21)	(0.21)	(0.21)
	BTE	Υ_{11}			0.33	0.35*	0.37 *
					(0.27)	(0.276)	(0.27)
Demographic controls [a]			No	No	No	Yes	Yes
Site fixed effects [b]			No	No	No	No	Yes

Random Effects		Parameter	Model 1 Estimate (Std.Error)	Model 2 Estimate (Std.Error)	Model 3 Estimate (Std.Error)	Model 4 Estimate (Std.Error)	Model 5 Estimate (Std.Error)
Level 1	Within person	σ_ε	5.31 ***	5.07 ***	5.05 ***	5.07 ***	5.08 ***
			(0.17)	(0.22)	(0.22)	(0.22)	(0.22)
Level 2	Initial status	σ_0	8.20 ***	8.99 ***	8.96 ***	8.70 ***	8.54 ***
			(0.34)	(0.39)	(0.39)	(0.41)	(0.44)
	Rate of change	σ_1		1.20 ***	1.26 ***	1.23 ***	1.22 ***
				(0.35)	(0.34)	(0.35)	(0.37)
	Correlation	σ_{01}		−0.54 ***	−0.56 ***	−0.56 ***	−0.54 ***
				(0.10)	(0.09)	(0.10)	(0.12)
Deviance			11,032.94	11,004.21	10,991.92	10,971.86	10,959.82
P (LR Chisquared Test)			0.00	0.00	0.00	0.00	0.00
n			1,608	1,608	1,608	1,608	1,608

One tailed tests; *** p-value < 0.01; ** p-value < 0.05; * p-value < 0.10.
[a] Models 4 and 5 control for race and gender.
[b] Model 5 controls for site specific charecteristics that are time invariant.

institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

Informed consent was obtained from all individual participants included in the study by the individual school districts.

Appendix B. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ijer.2019.01.006.

References

Aakhus, M., & Bzdak, M. (2015). Stakeholder engagement as communication design practice. *Journal of Public Affairs, 15*, 188–200. https://doi.org/10.1002/pa.1569.
Alexander, K. L., Entwisle, D. R., & Olsen, L. S. (2007). Lasting consequences of the summer learning gap. *American Sociological Review, 72*, 167–180.
Alvarado, S. A., & Muniz, P. (2018). Racial and ethnic heterogeneity in the effect of MESA on AP STEM coursework and college STEM major aspirations. *Research in Higher Education, 50*, 933–957.
American College Testing (ACT) (2017). *The condition of college and career readiness – 2017.* www.act.org/condition2017/.
American Institutes for Research (AIR) (2018). *What we know about the impact of the 21st CCLC program. Chicago, IL AIR.* Available at www.air.org/page/afterschool-and-expanded-learning.
Berger, A., Turk-Bicakci, L., & Garett, M. (2013). *Early college, early successes: Early college high school initiative impact study.* Washington, DC: American Institutes for Research.
Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective.* Hoboken, NJ: Wiley.
Brooks, L., MacAllum, K., & McMahon, A. (2005). *The bridge to employment initiative: A decade of promising school-to-career partnerships.* Washington, D.C: AED National Institute of Work and Learning. http://www.bridge2employment.com/act_sendfile.cfm?fileid=97.
Bzdak, M. (2007). The Johnson & Johnson Bridge to Employment initiative: Building sustainable community education partnerships. *Corporate Governance, 7*(4), 486–492.
Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity.* Washington, DC: U.S. Department of Health, Education, and Welfare.
Commission on the Skills of the American Workforce (1990). *America's choice: High skills or low wages?* Rochester: NY: National Center on Education and the Economy.
Committee on Highly Successful Schools or Programs for K-12 STEM Education (2011). *Successful STEM education: A workshop.* Washington DC: National Academic Press.

Crowder, M. J., & Hand, D. J. (1990). *Analysis of repeated measures.* London: Chapman and Hall.
Curran, P. J., & Hussong, A. M. (2003). The use of latent trajectory models in psychopathology research. *Journal of Abnormal Psychology, 112*(4), 526–544.
DeSilver, D. (2017). *U.S. students' academic achievement still lags that of their peers in many other countries.* Retrieved fromPEW Research Centerhttp://www.pewresearch.org./fact-tank/2017.
Detgen, A. (2017). *Bridge-to-Employment from the rearview mirror: Alumni perspectives on a college and career readiness program.* Washington, DC: FHI-360,.
Downey, D. B., VonHippel, & Broh, B. A. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review, 69*, 613–635.
Fayer, S., Lacey, L., & Watson, A. (2017). *STEM occupations: Past, present, future, Retrieved fromU.S. Bureau of Labor Statistics.* https://www.bls.gov/spotlight//STEM.
FHI-360 (2017). *Celebrating the BTE story: 25 years, 25 lessons.* Retrieved fromhttp://www.bridge2employment.org/wp-content/uploads/2017/03/25-Celebrating_the_BTE_Story-25Years-25Lessons.pdf.
Frechtling, J. A. (2007). *Logic modeling methods in program evaluations.* San Francisco, CA: Jossey-Bass.
Graf, N., Fry, R., & Funk, C. (2018). *Seven facts about the STEM workforce.* Retrieved from:PEW Research Centerhttp://www.pewresearch.org/STAFF/.
Hamilton, S. F., & Hamilton, M. A. (1999). *Building strong school-to-work systems: Illustrations of key elements.* Washington, DC: National School-to-work Office.
Hanushek, E. A., & Rivkin, S. G. (2009). Harming the best: How schools affect the black-white achievement gap. *Journal of Policy Analysis and Management, 28*, 366–393.
Hill, H. (2017). The Coleman Report, 50 years on: What do we know about the role of schools in academic inequality? *The Annals of the American Academy of Political and Social Science, 674*, 9–26.
Hughes, K. L., Bailey, T. R., & Mechur, M. J. (2001). *School-to-work: Making a difference in education.* New York: Columbia University, Institute on Education and the Economy.
Jagannathan, R., Camasso, M. J., & Delacalle,M (2018). The effectiveness of the head-heart-hands model for natural and environmental science learning in urban schools. *Evaluation and Program Planning, 66*, 53–62.
James-Burdumy, S., Dynarski, M., Moor, M., Deke, J., & Mansfield, W. (2005). *When Schools Stay Open Late: The National Evaluation of the 21st Century Community Learning Centers Program.* U.S. Department of Education, Institute of Education Services Available at http://www.ed.gov.ies.nces.
Jobs For the Future (2017). *Bridging to a better future: Findings from an evaluation of Bridge-to-College programs for English language learners.* Boston, MA: Jobs For the Future.
Kemple, J., & Willner, C. (2008). *Career Academies: Long term impacts on labor market outcomes, educational attainment and transitions to adulthood.* New York: MDRC.
Larson, E. H., & Vandergrift, J. A. (2000). *Tenth grade student perceptions of career preparation and work experience in Arizona schools. Arizona school-to-work Briefing Paper #18. Tempe, A2: Morrison Institute for Public Policy*.
Levine, P. B., & Zimmerman, D. S. (2010). *Targeting Investments in Children: Fighting Poverty When Resources are Limited.* Chicago: University of Chicago Press.
Maxwell, N. L., & Rubin, V. (2000). *Career academy programs in California: Implementation and student outcomes.* Hayward, CA: Human Investment Research and Education Center.
MESA (Mathematics, Engineering, Science Achievement) (2017). *Program description.* Retrieved fromhttp://mesausa.
National Commission on Excellence in Education (1983). *A Nation at risk: The imperative for educational reform.* Washington, DC: US Department of Education.
Organization for Economic Cooperation and Development (OECD) (2017). *Programme for international student assessment – PISA.* 2015 Retrieved from http://www.oecd.org/PISA/aboutpisa/.
Shadish, W. R., Cook, t. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs.* Boston, MA: Houghton Mifflin.
Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis.* New York: Oxford University Press.
Steinberg, A. (1998). *Real learning: Real work.* New York.
U. S. Department of Education Green Ribbon Schools (2018). *Green ribbon schools overview.* Retrieved fromhttps://www2.ed.gov/programs/green-ribbon-schools/index.html.
U.S. 103rd Congress (1994). *The school-to-Work opportunity act.* Washington, DC.
U.S. Department of Education (2016). *Overview information: Minority science and engineering improvement program (MSEIP). Federal Register, 81.204.* 7291–72797.
Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in science, technology, engineering and mathematics (STEM) and its relation to STEM education dose: A 25-year longitudinal study. *Journal of Educational Psychology, 102*, 860–869.
Webb, M., & Mayka, L. (2011). *Unconventional wisdom: A profile of the graduates of Early College High School.* Boston, MA: Jobs for the Future.