Higher Education Policy (2023) 36:431–456
https://doi.org/10.1057/s41307-022-00263-3

ORIGINAL ARTICLE

Meeting the Moment: Impact of TEACH Grant on US Undergraduate Education Degree Completion in High-Need Content Areas

David J. Peyton[1] · Wilhelmina van Dijk[2] · Loretta Mason-Williams[3]

Published online: 11 February 2022

Abstract

As part of the College Cost Reduction and Access Act (2007), the USA funded the TEACH Grant to incentivize earning a degree in a high-need content area (e.g., STEM fields, language-related areas, and Special Education) and to help meet teacher supply needs in low-income schools. Our analysis investigates the impact TEACH has had on the production of undergraduate education degrees overall and in high-need content areas. Using publicly available datasets and propensity score methods, we compare undergraduate education degree production at institutions of higher education, making comparisons between adopters and non-adopters of TEACH. Our findings suggest the adoption of TEACH had no impact on the overall production of undergraduate education degrees or production of education degrees in STEM, language-related fields, or special education. We situate our findings in the context of unrelenting demand for teachers in the USA.

Keywords Teacher shortages · Fiscal incentives · Teacher preparation

Declining enrollment in teacher preparation programs in the United States (USA) has accelerated the teacher shortage crisis and challenged school districts to adequately address teaching vacancies (Garcia and Weiss 2019). Since 2010, nearly every state suffered a decline in teacher preparation program enrollment and completion, with traditional preparation programs experiencing the greatest decline (Partelow 2019;

✉ David J. Peyton
peytondj@appstate.edu

1 Department of Reading Education & Special Education, Appalachian State University, Boone, USA

2 Department of Special Education and Rehabilitation Counseling, Utah State University, Logan, USA

3 Department of Teaching, Learning, & Educational Leadership College of Community & Public Affairs, Binghamton University, Binghamton, USA

Sutcher et al. 2019). These reductions are acutely felt in high-need content areas such as math and science (collectively referred to as STEM fields), language-related areas, and special education—among others (Cowan et al. 2016). For example, estimates suggest a national decline of STEM and special education degrees by 22% and 14%, respectively (Partelow 2019). Worse yet, declines in enrollment and subsequent staffing challenges disproportionately fell upon disadvantaged school districts with fewer resources and high numbers of students of color (e.g., African American and Latino students, Garcia & Weiss 2019). These school districts report having to hire a greater number of novice and uncertified teachers, as well as lower rates of teacher retention, negatively impacting the educational outcomes of students (Garcia and Weiss 2019).

To address these ongoing challenges, the US Congress prioritized investment in the preparation and adequate supply of teachers in hard-to-staff content areas, including STEM fields, languages, and special education through numerous federal programs (e.g., Teacher Quality Partnership Grants, Robert Noyce Teacher Scholarships). One substantive investment has come from the Teacher Education Assistance for College and Higher Education (TEACH) Grant—focused on the recruitment of candidates who will be fully qualified to teach in these high-need content areas in schools serving students from low-income families. As of 2019–2020, the US federal government, through TEACH, has invested over $900 million at nearly 800 Institutions of Higher Education (IHE; US Department of Education 2020). Despite this long-term investment, the impact on the overall production of undergraduate teacher candidates remains relatively unknown.

Although urgency to reauthorize the Higher Education Act (and potentially remove all public service loan payback programs) stalled with the crisis of COVID-19 in 2020, new plans for workforce development funding to restore the economy are in full force (Dembicki 2020). Simultaneously, the Bureau of Labor Statistics (BLS) reported that voluntary resignations and retirements saw substantial increases in 2020 leaving holes in the workforce as state budgets improved (Aldeman 2021). Therefore, investigating the effectiveness of TEACH to incentivize enrollment into teacher preparation as a way to meet ongoing supply needs is timely and provides policymakers with valuable information for the re-authorization of HEA.

In the current study, we examined ten years of education degree completion data to analyze the impact TEACH has had on the production of undergraduate education degrees in the USA. Specifically, we focused on the completion of teacher candidates in three broad areas: (a) STEM fields (including teachers prepared in math and a variety of science fields), (b) language-related areas (including all languages together, as well as teachers of English as a Second Language), and (c) special education. These areas have consistently been identified in state reports to the US Department of Education as particularly hard-to-staff along with national and state analysis of teacher shortages (Dee and Goldhaber 2017; Goff et al. 2018; Ondrasek et al. 2020; Goldhaber et al. 2015; Sutcher et al. 2019). This study makes several notable contributions. First, we present what we believe is the first empirical analysis of TEACH. Such an investigation is critical in light of current efforts to reauthorize the HEA, potentially removing public service programs, such as TEACH (Strauss 2018). Second, this study expands our understanding of the role of student loan programs to recruit individuals into the teacher pipeline from other academic programs or majors at IHEs. Prior research has focused to great extent on post-baccalaureate career changers, often evaluating the effectiveness of tuition and stipends for individuals to complete graduate programs leading to teacher certification (e.g., Feng and Sass 2017), with limited attention to increasing the supply at the undergraduate level. This gap is considerable given the prominence of teacher preparation as an undergraduate pursuit.

Our path of inquiry proceeds in the following manner. First, we offer a conceptual framework for understanding factors that may influence an IHE to choose to offer the TEACH loan program, which we believe is motivated by a number of local and regional factors related to supply and demand, and provide additional background on TEACH, including potential challenges to its effectiveness in increasing the number of teacher candidates. Next, we describe our analytic approach to examine the impact of TEACH on the production of undergraduate education degrees generally, and in high-need content areas specifically. We then situate our findings within the current research and generate hypotheses about the impact TEACH has had on US teacher supply. We conclude by providing limitations to our approach and future directions that may improve the effectiveness of TEACH to increase the number of completers in high-needs content areas, as the program comes up for reauthorization.

Conceptual Framework

We conceptualize the adoption of TEACH by an IHE within a supply and demand framework. We posit that TEACH is a recruitment policy initiated by Congress to increase the supply of teachers in hard-to-staff content areas in high poverty schools and taken up by IHEs on a case-by-case basis, driven by regional demand. Recruitment policies intended to increase the supply of teachers are thought to target improving the rewards (e.g., monetary compensation and working conditions) of teaching vis-á-vis all other occupations. In this way, TEACH can be considered a monetary incentive designed to offset the cost of college attendance, effectively representing a boost to entry-level compensation by reducing a student's loan burden (Guarino et al. 2006; Lovenheim and Turner 2018).

At the same time, take-up of TEACH by a given IHE may be based on institutional resources and priorities, in addition to local supply and demand conditions. In other words, take-up of TEACH may be driven by variables unique to a given institution and the university mission. For example, regional teacher comprehensive IHEs may produce a larger supply of education graduates than land-grant, research-focused institutions. Therefore, regional IHEs may be more likely to take up TEACH, and as a result have the potential to impact a greater number of students. Additionally, larger universities, with potentially larger student aid offices, may have greater capacity to provide TEACH for students. Thus, we include institutional-level measures that relate to the overall supply of undergraduate enrollment.

In addition, TEACH might be taken-up and marketed by IHEs based on local teacher demand and proximity to eligible schools that would allow recipients to meet

TEACH requirements. Presumably, IHEs intent on serving hard-to-staff schools may be more likely to see TEACH as a vehicle to meet local demand. As such, we consider variables that reflect differences between local supply and demand at the county level, such as the number of schools, county public level expenditures per student, and special support programs. Thus, we adopt a teacher labor market supply and demand framework to both conceptualize TEACH policy and interpret factors that relate to take up by IHEs.

At the same time, participation in TEACH requires an individual to opt-in, potentially requiring a change in major. Unfortunately, little is known about how a student selects an undergraduate major or the factors that may influence the decision-making process, making it difficult to know if TEACH could be influential as a recruitment tool. Stinebrickner and Stinebrickner (2014) conceptualized the decision-making process as being influenced by the beliefs individuals hold about a major at entrance to college, which are then updated during preparation, as influenced by factors such as grade performance and beliefs about future income. This combination of institutional and individual factors may complicate the effectiveness of TEACH as a teacher workforce recruitment tool in increasing the production of education majors within an institution, both generally and in high-need content areas.

Background on the TEACH Grant

Enacted by Congress as part of the College Cost Reduction and Access Act (PL 110-84) of 2007, TEACH helped address the twofold dilemma of teacher candidate shortages in high-need content areas and the inequitable distribution of qualified teachers in schools serving students from low-income families. In essence, TEACH provided a mechanism to increase the number of teacher graduates through levers intended to promote recruitment (i.e., fiscal incentives) and encourage retention (i.e., required length of service) in high-needs content areas in schools serving low-income families.

Across the USA, not all institutions offer TEACH to students as a way to offset the costs of attendance. For a student to participate in TEACH, the financial aid offices at IHEs must elect to participate. Additionally, the IHE must offer TEACH-eligible preparation programs that allow students to become highly qualified in high-need content areas. Among states, the designation of a high-need content may vary as the designation comes from state reports to the USDOE, as published annually in the Nationwide Teacher Shortage List. According to these reports, STEM fields, language-related areas, and special education are consistently identified as high-needs content areas.

Even if available at the IHE, not all undergraduates are able to take advantage of TEACH. First, a student must meet basic eligibility requirements for the federal student aid programs and complete the free application for federal student aid (FAFSA). Next, a student must elect to major in a TEACH-eligible content area and demonstrate high academic achievement (i.e., score above the 75th percentile on one or more portions of a college admissions test or maintain a minimum 3.25 grade point average). Based on cost of attendance, qualified undergraduate students may be eligible for an award up to $4,000 annually (with a $16,000 limit; graduate students have an $8,000 limit). On average, TEACH participants have received an annual disbursement of approximately $2,900 (US Department of Education 2020).

As a public service loan program, recipients agree to several conditions in exchange for the award. First, following graduation, TEACH participants must serve as a full-time, highly qualified teacher for at least four of the first eight academic years of receiving the award in a high-needs content area (presumably the content area in which they completed preparation). Additionally, the teaching position must be at a low-income school or education-based service agency, as listed in the Teacher Cancellation Low-Income Directory, a listing of all schools or districts with at least 30% of student enrollment qualifying for Title I services (https://studentaid.gov/app/tcli.action). To verify meeting this commitment, TEACH recipients annually certify their intent to teach or verify a current teaching position at a qualifying school. TEACH recipients who fail to meet these conditions have the grant converted to an unsubsidized loan and are required to pay the grant back in full plus interest.

Concerns Raised Concerning the Effectiveness of TEACH

To date, no research studies have examined the impact of TEACH on increasing degree production in teacher preparation programs in high-need content areas. However, two US federal reports suggested significant challenges to the success of TEACH related to the (a) marketing, (b) implementation, and (c) award amount of TEACH. First, Nowicki (2015), in an evaluation conducted by the US Government Accountability Office, found there to be an overall lack of awareness of TEACH by financial aid officers and members of teacher preparation programs. Moreover, Nowicki found that IHEs were unlikely to adopt TEACH due to a perception of graduates' limited likelihood of meeting TEACH requirements, overly burdensome paperwork, and the potential for unfavorable conversion rates.

Second, a 2018 US Department of Education report conducted by Barkowski and colleagues provided insight on (a) recipients' perceptions of the requirements of the grant, (b) their ability to meet requirements, and (c) administration of the grant by IHEs. Barkowski et al, reported only half of respondents representing IHEs stated they used the grant to encourage or recruit undergraduates and graduate students into teaching. Interview data revealed the process, persons, and manner in which details of the grant were introduced to students varied between and within institutions. Similarly, Denning and Turley (2017) investigated the comparably designed US Science and Mathematics Access to Retain Talent Grant (SMART) and noted that the information gap likely played a role in take-up. Only 6.8% of federal aid recipients—the target population of the program—demonstrated existing knowledge of the SMART Grant. Evans (2016) concluded that institutions were left to promote the program as they wished, introducing variability by which students were made aware of the program.

The TEACH grant relies heavily on the influence an IHE has on an undergraduate's selection of a major, despite that major can be a significant decision with considerable lifelong impact. And yet, very little research is available on factors influencing the decision to switch (Denice 2021). Generally, researchers conceptualize the process as including beliefs about a major an undergraduate held at entrance to the university, beliefs that may be updated during school, as well as other determinants such as performance in the major, access to programs, and perceived employability (Stinebrickner and Stinebrickner 2014). For TEACH to impact the teacher workforce, IHEs need to persuade undergraduates to select certain majors leading to certain content, as well as a willingness by the undergraduate to accept teaching positions in certain locations.

Finally, Podolsky et al. (2016) found financial awards that fell between $1,000 and $3,000 were ineffective at either attracting teachers into high-need districts or enticing individuals into majoring in science education. Further, Evans (2016) and Denning and Turley (2017) research on SMART grants, with an annual award amount of $4000, found similar outcomes. Their analysis found no significant impact of the SMART grant program on undergraduate enrollment in STEM majors. There were, however, mixed results with regard to the impact of the SMART Grant on student persistence. Evans (2016) found that SMART recipients did not persist at any higher rate than non-recipients. In contrast, Denning and Turley (2017) found slightly different impacts of SMART grants at Texas public universities and Brigham Young University, with STEM majors persistence resulting in increases by 3% and 10%, respectively.

With what is known about the limited effectiveness of fiscal incentives to attract and retain individuals into high-need content areas (Carver-Thomas 2018; Denning and Turley 2017; Evans 2016), and prior research on implementation of TEACH (Barkowski et al. 2018; Nowicki 2015), we posit that TEACH likely has had a limited impact on the enrollment and completion of undergraduate education programs, generally, and high-need content areas, specifically. To test these assumptions, we pose the following research questions.

Research Question #1 What effect has the adoption of TEACH had on the production of undergraduate education degrees, generally?
Research Question #2 What effect has the implementation of TEACH had on the production of undergraduate education degree majors in high-need content areas (i.e., STEM, language-related areas, and special education)?

Methods

Sample

To estimate the impact of TEACH on undergraduate degree production, we relied on publicly available data from the National Center for Educational Statistics (NCES) databases, specifically, the Integrated Postsecondary Education Data System (IPEDS), the Federal Student Aid (FSA) Data System, the Elementary and Secondary Information System (ELSI), and the Bureau of Labor Statistics (BLS). Our decision to focus on undergraduate degree production was purposeful as in the USA, the vast majority of new teachers enter the workforce after completing preparation at the undergraduate level. Moreover, the complexities (e.g., professional influences, availability, and complexity of alternative route programs) of estimating the policy effects on graduate degree enrollment and degree completion would only further complicate our analysis as the motivations of these populations are substantively different and likely requires a separate analysis (Bedard and Herman 2008).

First, descriptive data for all institutions were derived from the IPEDS database (https://nces.ed.gov/ipeds/). IPEDS is a robust set of surveys collected annually by the US Department of Education across all IHEs involved in federal student financial aid programs. Key to our analysis, for example, was data gathered on enrollment, classification of instructional programs codes (CIP) to determine eligible majors as identified in the TEACH statute, and degrees and certificates conferred. Next, to determine participation by institution, we examined the FSA database. FSA reports the quarterly financial disbursement of TEACH funds to participating institutions beginning in 2008, the first year of grant eligibility. Finally, ELSI and BLS databases were used to generate data at the state, county, and institutional level. Together, this created a rich dataset of attributes for the IHEs and their local context for investigation.

Based on data availability, our initial sample consisted of all public and private IHEs in existence between 2000 and 2014–2238 distinct institutions. We removed from the sample: (a) any institution that was classified as a two-year community college, (b) public or private four-year institutions that awarded more than 25% of total degrees as associates' degrees or certificates, and (c) institutions that enrolled less than 10 undergraduate students. However, our sample does not distinguish between institutions that adopted TEACH at the undergraduate or graduate level only (i.e., some institutions may allow TEACH participation only among undergraduates or vice versa), a limitation of the analysis. Our final institution-level sample included 1467 distinct institutions.

The institutions sampled represent the diversity found among colleges and universities in the USA. Approximately one-third of IHEs were private institutions (n = 525). More private institutions adopted TEACH than did not adopt (347 vs. 178), whereas the proportion was more equal among public institutions (471 vs 437). Fewer than one in five IHEs were research/doctoral institutions (18.89%, n = 277), and the majority of those were adopters (68%, n = 188). On average, the IHEs produced 48.06 BA degrees for every 100 students (graduate and undergraduate) enrolled. Adopting and non-adopting IHEs did not differ in this number (t [1465] = 1.647, $p = 0.10$).

Table 1 provides an overview of the annual average degree production in teacher preparation-related programs throughout our sample from 2000 to 2016. TEACH-eligible degree programs (including STEM fields, language-related areas, and special education) constituted 12.7% of education degrees produced annually. Using CIP codes available within the dataset, we combined multiple majors into our three categories. For instance, STEM fields include those teacher candidates majoring in biology teacher education, physics teacher education, mathematics education, and a number of other content areas often designed as "STEM fields." Similarly, we included undergraduate programs coded as Teachers of English as a Second

Table 1 Average annual education degree production in public and private four-year colleges (2000–2016)

	All four-year		Public four-year		Private four-year	
	N	% of Total	*N*	% of Total	*N*	% of Total
Total	104,677.4	100	72,522.88	69.3%	30,407.29	29.1%
Percent of education degrees across enrollment		1.33%		1.34%		1.36%
Content Areas						
High-Needs	13,289.38	12.7%	9,167	68.9%	3,546	26.7%
Not High-Needs	91,388.06	87.3%	66,402	72.7%	27,353	29.9%
Race/ethnicity						
White	82,422.88	78.8%	61,592	74.7%	25,096	30.4%
Students of Color	21,563.25	20.6%	13,538	62.8%	6,107	28.3%

High-needs content areas in this analysis included degrees in STEM fields, world languages, special education. Percentages do not total 100 due to 34 institutions with missing values on the variable indicating private/public.

Language, Bilingual Education, Spanish Teacher Education, and a number of other areas in the category of "language-related fields." A complete listing is available in the Supplemental Tables.

Public four-year institutions produced a significant majority of education bachelor's degrees (71%), as well as education degrees in high-need content areas. About 10% of all education-related degrees were awarded to undergraduates in high-need content areas. More specifically, each year approximately 3% of all undergraduates were awarded a degree in a STEM field, 1% a degree in a language-related area, and 6% a degree in special education. About 10% of undergraduates completing a bachelor's degree in education were TEACH recipients. Of all TEACH recipients, 33% pursued a degree in STEM fields, 10% in language-related areas, and 52% in special education.

Measures

Independent Variable

Our independent variable measured the production of undergraduates who participated in TEACH by IHEs anually. We identified institutional participation in TEACH by examining annual financial disbursement of funds to eligible institutions through the FSA database.

Covariates

We inspected several variables at the institution, county, and state levels. At the institution level, we employed IPEDS data to examine (a) total undergraduate enrollment, (b) the percentage of institutional enrollment engaged in graduate studies, (c) instructional expenditures per full-time employee (FTE), (d) in-state tuition and fees, (e) total education bachelor's degrees produced per FTE, and (f) status as a Historically Black College or University (HBCU, Chen and Wiederspan 2014; DesJardins et al. 2003). At the county-level, we gathered data from ELSI on (a) total number of public elementary and secondary schools operating in the county, (b) proportion of students served in special support programs (i.e., free or reduced-price lunch, special education, and English Language Learners), (c) racial/ethnic composition of the student body, (d) county-level public educational expenditures per student, and (e) the proportion of instructional expenditures on salaries and benefits, a proxy for average teacher salaries (Hillman et al. 2015). At the state-level, we accessed data from the BLS to estimate (a) total state college-aged (18–24) population, (b) state median per capita income, and (c) state unemployment rate (Doyle 2009; Humphreys 2000; McLendon et al. 2009; Weerts and Ronca 2012).

Outcome Variables

Several dependent (outcome) variables allowed us to test the efficacy of TEACH to increase enrollment in education degree areas, especially in high-needs content. We were interested in: (a) total number of bachelor's degrees in education, and (b) total number of bachelor's degrees in education within TEACH-eligible high-needs content areas (i.e., STEM, language-related areas, and special education).

Table 2 provides an overview of the annual average education degree production among IHEs included in the sample in 2007, the year prior to implementation of TEACH. In addition, Table 2 provides means and standard deviations for each of the control variables at the institution-, county-, and state-level used in the propensity scoring model used in the analysis.

Data Analytic Approach

To determine the effectiveness of TEACH to increase the number of undergraduates prepared to teach high-need content areas, we conducted a quasi-experimental design analysis comparing IHEs adopting TEACH (n = 821) to propensity score matched IHEs that did not adopt TEACH (n = 646). Propensity score methods have the potential to reduce bias in non-randomized samples by using a model of potential confounders to identify groups of similar cases based on their conditional probability for being assigned to the treatment given the observed covariates (Leite 2016; Rosenbaum and Rubin 1983). Previous research has shown effect estimates from studies using propensity scores to be equally accurate as estimates from randomized controlled trial studies (Fortson et al. 2012), as well as to analyze outcomes at the institutional level (e.g., Rodríguez and Galdeano 2015).

Propensity Score Estimation

To estimate the propensity scores, we used a multilevel logistic regression model, keeping the clustered nature of the data intact (i.e., IHEs clustered in counties,

Table 2 Descriptive statistics of adopters and non-adopters in 2007, prior to the Implementation of TEACH

Covariates	All IHEs (n = 1,467)		Adopters (n = 821)		Non-Adopters (n = 646)	
	Mean	SD	Mean	SD	Mean	SD
Institution-level						
Undergraduate enrollment (#)	5,179.98	6,393.31	6,320.81	7,182.71	3,730.10	4,852.97
Proportion graduate students (%)	16.58	16.58	17.98	15.49	14.79	17.73
Instructional expenditures per FTE ($)	7,104.93	6,457.57	6,512.47	4,850.97	7,857.88	7,990.49
In-state tuition and fees ($)	17,479.87	11,001.93	16,006.43	10,123.44	19,456.55	11,803.31
Total education BA degrees per 100 FTE (#)	1.86	1.79	1.91	1.64	1.78	2.02
HBCU status (#)	82		52		30	
County-level						
K-12 public schools (#)	16.89	51.39	16.18	47.43	17.87	56.51
Free/reduced lunch (%)	41.24	16.23	41.65	15.81	40.72	16.77
Special education (%)	14.39	8.88	13.98	7.02	14.93	10.82
ELL (%)	4.45	5.52	4.57	5.82	4.30	5.10
White student enrollment (%)	61.13	26.97	60.72	26.22	61.66	27.93
Per pupil expenditures ($)	15,043.24	22,042.07	14,483.15	16,161.28	15,824.85	28,272.17
Instructional expenditures (%)	66.49	14.59	65.94	14.73	67.25	14.38
State-level						
College-aged population per1,000	1,027.05	904.71	1,030.16	913.47	1,022.90	893.60
Per capita income ($)	45,203.00	6,794.60	44,710.83	6,516.86	45,827.94	7,087.68
Unemployment rate (%)	4.59	0.80	4.61	0.86	4.58	0.72

IHE, Institution of Higher Education; *FTE*, full-time equivalent; *ELL*, English Language Learners

clustered in states). Specifically, all IHEs adopting TEACH were coded as one, and all other IHEs as zero on a treatment indicator. To account for the period leading up to the decision to participate in TEACH, we used data for each of the IHEs from 2007, that is, one year before TEACH eligibility of IHE's and two years before students taking advantage of TEACH would graduate. Since the treatment assignment (i.e., adopting TEACH, yes or no) happened at the IHE-level, we used the institution-, county-, and state-level covariates identified as potential confounders, similar to other analyses (Chen and Wiederspan 2014; Des Jardins et al. 2003; Doyle 2009; Hillman et al. 2015; Humphreys 2000; Kelcey 2011; Leite et al. 2015; Li et al. 2013; McLendon et al. 2009; Weerts and Ronca 2012).

We estimated the propensity scores marginally across clusters, recommended when cluster sizes are small, accounting for the clustering effect in both the propensity score model, and the outcome model (Leite 2016), while ignoring the variation of covariate effects across clusters by not including random slopes. This method is adequate to remove bias even if clusters differ substantially (Gurel and Leite 2014; Kelcey 2011; Leite et al. 2015). We estimated the propensity score model in *R* (R Core Team 2019) with the *lme4* package (Bates et al. 2015, see online supplement). To evaluate if the propensity scores had common support across both adopters and non-adopters, we visually inspected box and whisker plots and histograms to check if the distribution of the adopters was contained within that of the non-adopters.

Propensity Score Implementation

We then used two methods to implement the propensity score to obtain the average treatment effect on the treated (ATT): weighting-by-the-odds and full matching. In weighting-by-the-odds, all treated cases receive a weight of one. The untreated cases received a weight proportional to their propensity score divided by one minus their propensity score. The calculated weights were then used to adjust for over-selection of IHEs with certain characteristics to either the treated or untreated groups, and the weighted sample is representative of a pseudo-population with similar covariate distributions across the two groups (Leite 2016). Full matching connects each treated case to at least one untreated case based on their propensity score, and vice versa, producing several matched sets of cases (Harder et al. 2010). This matching method minimizes global propensity score distance, without having overlap in matched sets or discarding cases (Leite 2016). After matching, each case receives a weight. Similar to the weighting-by-the-odds method, all treated cases receive a weight of one, while untreated cases receive a weight that is proportional to the number of treated cases in the matched set divided by the number of untreated cases in that same set (Harder et al. 2010). We used the *matchIt* (Ho et al. 2011) and *optmatch* (Hansen & Klopfer 2006) packages in R to implement the full matching algorithm.

We checked for covariate balance for both propensity score implementation methods by comparing the standardized difference between the weighted means of both groups. Balance is considered to be achieved if the standardized mean difference on a covariate is lower than 0.1 (Austin 2011). To control for the lack of balance in covariates with standardized mean differences between 0.1 and 0.25, we also included those as covariates in the outcome estimation models (Leite 2016). We proceeded to estimate the effect of TEACH with the method which yielded the best balance.

Estimation of Treatment Effect

To estimate the effect of adoption of TEACH on the production of bachelor's degrees in education at the undergraduate level while considering the overdispersion of the data, we estimated two-piece mixed effects negative binomial growth models using *glmmTMB* (Brooks et al. 2017) package in R. We then compared the growth rate during the adoption period for treated and untreated IHEs. The outcome variables were calculated as the percentage of undergraduate degrees awarded at each institution (i.e., the number of degrees divided by the number of undergraduates). All unbalanced covariates (i.e., with a standardized mean difference between .1 and .25) were entered as control variables, centered at the grand mean. Both slope indicators were coded in such a way as to have the intercept represent the outcome in 2009, the year in which the first eligible recipients would graduate, signaling completion of a degree in teaching. In this way, our models can be considered causal, with the propensity scores estimated on variables that may have influenced adoption decisions in 2007 and the analysis reflecting change in degree production of adopting and non-adopting institutions in 2009.

Results

The purpose of this investigation was to examine the impact TEACH has had on the production of undergraduate education degrees and undergraduate education degrees in high-need content areas. Propensity score methods in a quasi-experimental design allowed us to establish baseline equivalent groups through estimating their conditional probability for being assigned to the treatment given a set of covariates (Leite 2016; Rosenbaum and Rubin 1983). Any differences between groups on outcome variables can then be attributed to the treatment (i.e., implementation of TEACH).

Treatment Effects

Table 3 provides an overview of the annual average education degree production per 1,000 undergraduates throughout our sample from 2001 to 2016. Regarding overall undergraduate education degrees, the average number of undergraduates who earned an education degree ranged from 10.30 in 2012 to 8.78 in 2016. Undergraduate education degrees in STEM degrees ranged from 1.78 to 2.47, peaking in 2012 and at the lowest point in 2016. In addition, undergraduate degrees in language-related areas ranged from .58 per 1,000 undergraduates in 2002 to 1.06 in 2009. Finally, the average number of undergraduate education degrees in special education ranged from 5.35 in 2005 to 6.47 in 2013.

Propensity Score Evaluation

The propensity scores generated by the three-level logistic model were adequate. Visual inspection of Figures 1 and 2 illustrate that the range of propensity scores of adopting IHEs is contained within that of non-adopting IHEs. Propensity scores ranged from 0.12 to 0.98, having good common support. For each of the propensity score implementation methods (i.e., weighting-by-the-odds and full matching), we evaluated equivalence across groups through covariate balance estimation. Table 4 shows the equivalence statistics for each implementation method. Using the weighting-by-the-odds method, one variable remained unbalanced (i.e., total education bachelor's degrees produced per FTE [smd − 0.37] and one variable exceeded the ideal value of .10 [i.e., counties' salary expenditures]). Using the full matching method, all variables reached adequate balance, with four variables exceeding the 0.1 ideal value (i.e., number of undergraduates, tuition and fees, counties salary expenditures, and state's average income). Therefore, we only estimated the

Table 3 Average annual degree production per 1000 undergraduates in education programs overall and by high-needs content areas

	Pre-Adoption of TEACH								Post TEACH Adoption							
	2001 (*SD*)	2002 (*SD*)	2003 (*SD*)	2004 (*SD*)	2005 (*SD*)	2006 (*SD*)	2007 (*SD*)	2008 (*SD*)	2009 (*SD*)	2010 (*SD*)	2011 (*SD*)	2012 (*SD*)	2013 (*SD*)	2014 (*SD*)	2015 (*SD*)	2016 (*SD*)
All Education	9.41 (25.93)	9.24 (25.64)	8.84 (23.96)	8.95 (24.54)	8.91 (22.83)	8.91 (22.67)	9.33 (23.81	8.96 (23.29)	9.21 (24.37)	9.60 (26.60)	9.68 (26.51)	10.30 (28.11)	10.27 (28.16)	9.91 (27.61)	9.18 (25.26)	8.78 (24.07)
STEM fields	1.94 (8.11)	2.01 (8.27)	1.94 (8.00)	2.12 (9.00)	2.11 (8.42)	2.19 (8.14)	2.30 (8.84)	2.21 (8.28)	2.11 (7.79)	2.25 (7.79)	2.26 (7.95)	2.47 (8.33)	2.44 (8.83)	2.24 (8.78)	2.02 (7.93)	1.78 (7.14)
World languages	0.60 (2.71)	0.58 (2.6)	0.59 (2.46)	0.68 (2.97)	0.82 (3.8)	0.84 (3.75)	0.85 (4.29)	0.86 (4.00)	1.06 (6.82)	1.02 (6.56)	1.03 (6.39)	1.04 (6.55)	0.91 (4.85)	0.87 (3.94)	0.75 (3.63)	0.68 (2.95)
Special education	6.23 (18.81)	6.02 (18.3)	5.67 (16.98)	5.40 (15.45)	5.38 (14.93)	5.31 (14.45)	5.61 (14.98)	5.39 (14.98)	5.55 (15.67)	5.82 (18.46)	5.86 (17.77)	6.30 (19.23)	6.47 (19.54)	6.40 (19.33)	6.01 (18.34)	5.91 (17.98)

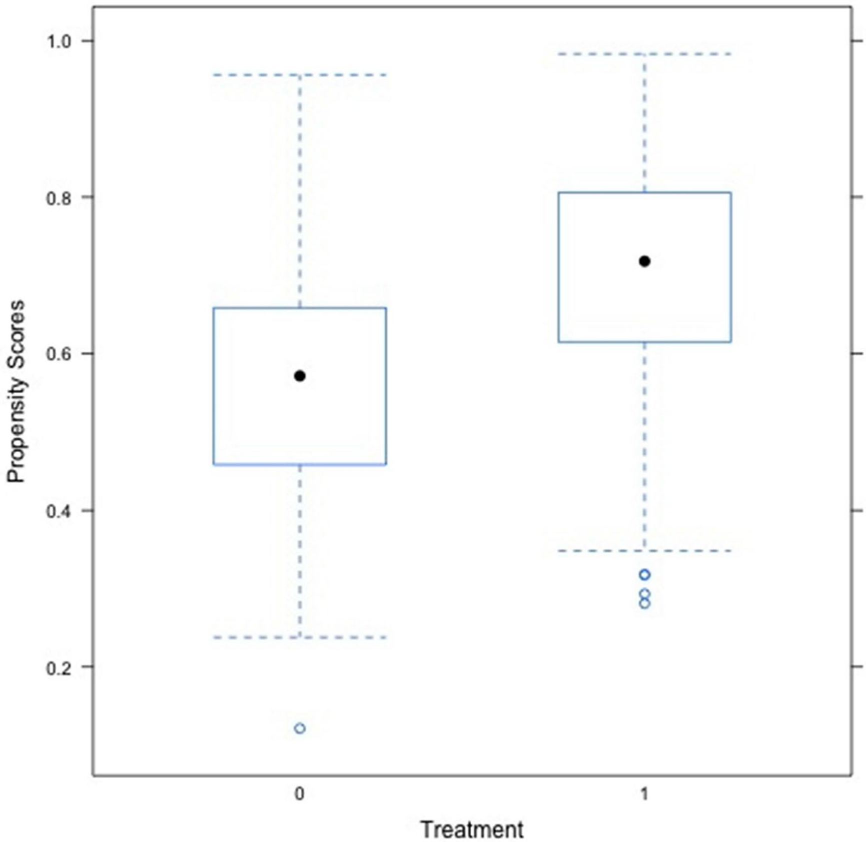

Fig. 1 Box and whisker plot of propensity scores

treatment effect using the matching method and included the four unbalanced variables in the effect estimation models.

The results of our first research question, the number of undergraduate degrees in education, showed there was almost no difference in growth between participating and non-participating schools across all students after adopting TEACH (log (0.015), SE = .068, OR = 1.02, p = .075). Schools with higher undergraduate enrollment and a higher per capita income in the state were associated with minimally lower rates of education degree production (i.e., log (< -0.001), SE < .001, OR = 0.999, p < .001 for both variables). For TEACH adopters, per capita income had a statistically significant, but very small, effect on degree production (log (0.001), SE = <.001, OR = 1.000017, p = .002), both pre- and post-adoption. Note that the pre-adoption slope had a statistically significant interaction with the treatment indicator, signaling that education degree awards differed for adopting and non-adopting institutions before TEACH. Stated differently, adopting institutions education degree production was declining at a faster rate

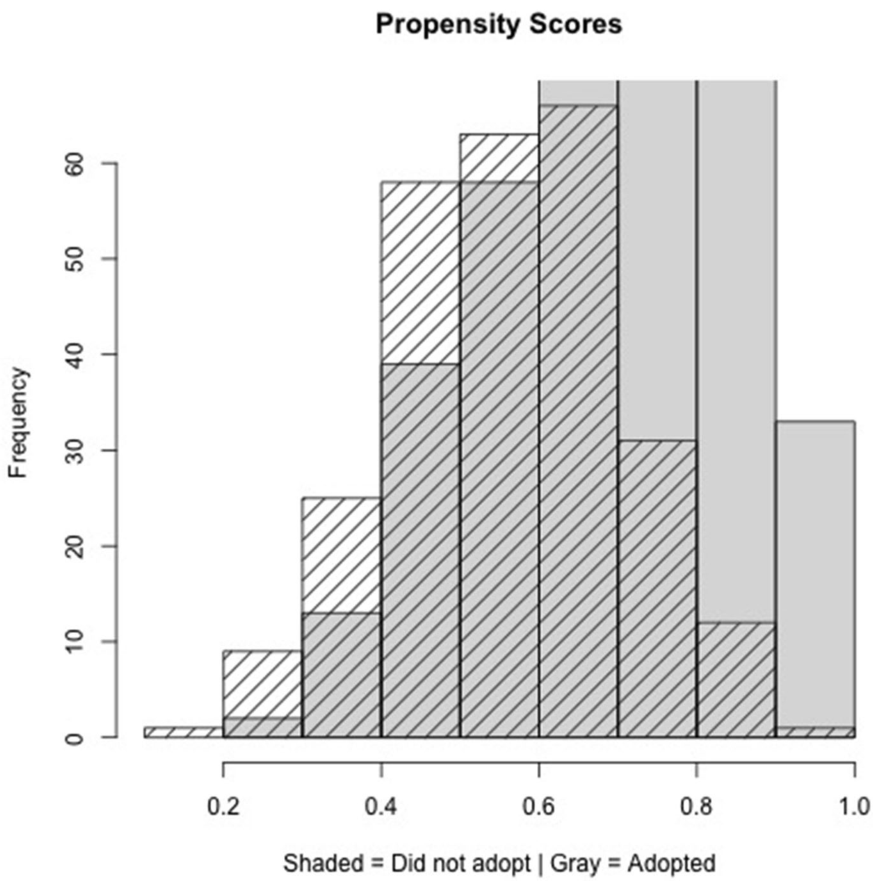

Fig. 2 Histogram of common support of propensity scores

compared to non-adopting institutions prior to TEACH. All parameter estimates for total undergraduate degree production are presented in Table 5.

For research question 2, we investigated more closely the impact TEACH availability had on degree production in three high-need content areas: STEM fields, language-related areas, and special education. Among STEM fields, there was a significant interaction for post-adoption slope by treatment. This indicates that adopting institutions had a steeper rate of growth of degrees awarded to undergraduates than non-adopting institutions (log (0.095), SE = 0.040, OR = 1.099, p = .018). This interaction effect is shown in Fig. 3. Similar to the overall degree model, the pre-adoption slope had a statistically significant interaction with the treatment indicator, indicating that the trajectories of adopting and non-adopting institutions differed before TEACH. None of the other variables were significant predictors in the STEM subject models. To evaluate the robustness of this effect, we ran *post hoc* analyses for two additional placebo treatment years (i.e., a pre-decision year, 2005,

Table 4 Equivalence statistics for propensity score implementation methods

Covariates	Weighting-by-the-odds	Full matching
Institution-level		
Undergraduate enrollment	– 0.007	**– 0.239**
Proportion graduate students	0.066	0.030
Instructional expenditures per FTE	0.033	0.082
In-state tuition and fees	0.013	**0.105**
Total education BA degrees per 100 FTE	**– 0.349**	– 0.041
HBCU status	< 0.001	0.079
County-level		
K-12 public schools	< – 0.001	0.063
Free/reduced lunch	– 0.023	– 0.063
Special education	– 0.018	< 0.001
ELL	0.048	0.027
White student enrollment	– 0.049	– 0.004
Per pupil expenditures	0.035	0.078
Instructional expenditures per $10,000	**– 0.135**	**– 0.156**
State-level		
College-aged population per 10,000	0.033	0.075
Per capita income per $1,000	0.067	**0.180**
Unemployment rate	0.026	0.097

Estimates in bold face denote insufficiently balanced covariates. FTE= full-time employed; ELL= English Language Learners

a pre-adoption year, 2008, and an additional post-adoption year, 2012). The effect of TEACH adoption on STEM undergraduate degrees was present in all of these models, eliminating any confidence in the results. Supplemental Table 1 provides the results of these additional robustness models.

In language-related areas, our analyses demonstrated no statistically significant interactions for slope by treatment at pre- or post-adoption periods. Similarly, the models for special education also did not show a significant slope by treatment interaction pre- or post-adoption. Similar to all education degrees, undergraduate enrollment and state per capita income influenced degree production negatively across the time span. In addition, the institution's tuition and fee rate also had a small negative impact on special education degree production (log (– 0006), SE = 0.003, OR = 0.994, p < .001). All parameter estimates for high-needs areas are presented in Table 6.

Discussion

TEACH represents a US federal incentive program aimed at increasing production of education degrees in high-need content areas and distributing highly qualified teachers to schools serving students from low-income families. Using propensity

Table 5 Parameter estimates for education degrees overall

	All Undergraduates		
	Log coef	SE	*p*
Intercept	5.012	0.068	< .001***
Pre-adoption slope	– 0.003	0.004	.443
Post-adoption slope	– 0.019	0.007	.006**
County salary expenditures[a]	< 0.001	< 0.001	.632
State per capita income[b]	< – 0.001	< 0.001	< .001***
Undergraduate enrollment	< – 0.001	< 0.001	< .001***
Tuition and fees	< 0.001	< 0.001	.489
Treatment indicator	0.139	0.063	.028*
Pre-adoption slope by treatment	– 0.029	0.004	< .001***
Post-adoption slope by treatment	0.015	0.008	.075
County salary expenditures[a] by treatment	< 0.001	< 0.001	.616
State per capita income[b] by treatment	< 0.001	< 0.001	.002**
Undergraduate enrollment by treatment	< – 0.001	< 0.001	.121
Tuition and fees by treatment	< – 0.001	< 0.001	.234

Intercept is set at 2009. All covariates are grand mean centered.

[a] Per $10,000, [b]: Per $1,000

**p* < .05; ***p* < .01; ****p* < .001

score methods, we analyzed the impact of TEACH on the overall production of undergraduate education degrees and production in high-need content areas (STEM, language-related areas, and special education). Results of our analyses suggest IHEs saw no difference in the overall number of education degrees among adopting and non-adopting institutions. With regard to our inquiry into high-need content areas, our analysis suggests increases in the number of STEM education degrees produced, but robustness analyses show this is likely a persistence effect. As found in previous studies of the SMART grant, funding did not increase enrollment, but allowed students to persist in completing their degrees (Denning and Turley 2017; Evans 2016). Finally, we found no increases in education degrees produced in language-related areas and special education.

In contrast with our conceptual framework, the variables representing institutional, local, and state factors included did not appear to influence education degree production. One exception was per capita income, which had a small impact on overall production and production of language-related areas. This may be indicative of perceived “town-gown” divides, where IHEs operate in isolation of pressing workforce needs. Regardless, as a recruitment tool participation in TEACH did little to help increase the production of undergraduate education majors. Although changing a major is common among undergraduates, can be disruptive and have major consequences for GPAs, time to obtain degree, and perceived lifetime earnings (Denice 2021). It may be that the award TEACH offers is insufficient to make such a disruption seem worthwhile.

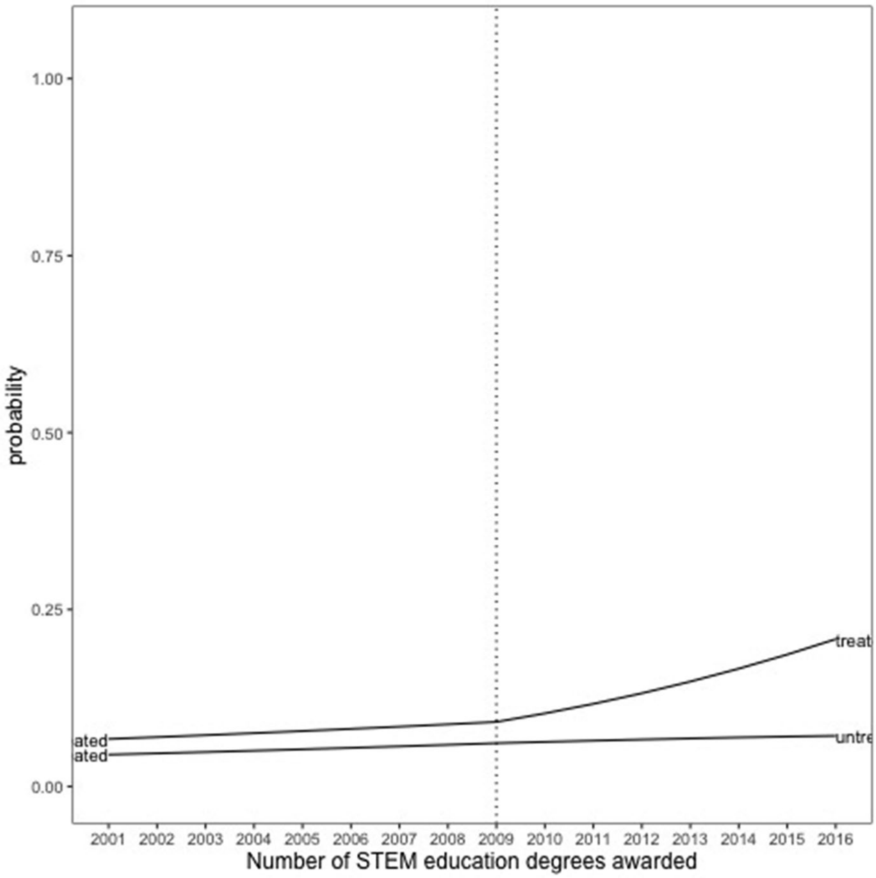

Fig. 3 Probability plots for STEM degrees. The plot shows the predicted probability of an increase in degree production per year as a function of treatment status

The lack of impact of TEACH on increasing the number of overall education degrees is neither entirely disappointing nor surprising, given previous research on the impact of fiscal incentives (Barkowski et al. 2018; Denning and Turley 2017; Evans 2016; Nowicki 2015). We posit that these outcomes are likely related, but not limited, to both the design and implementation of TEACH (e.g., total award amount, disbursement practices, & marketing). If changes were made to increase the attractiveness of the TEACH program leading to even a minor increase in the uptake, it could have a considerable impact on the teaching supply for hard-to-staff schools. For instance, a 10–25% increase in undergraduates participating in TEACH could lead to more than 2,000 new teachers. At the same time, a major increase in participation would do even more to stave off shortages. For instance, if the TEACH program participation increased by 50% in special education in 2016, approximately 3,000 new, fully certified special educators would be ready to enter high-need

Table 6 Parameter etimates for degrees in high-need areas

	STEM			World Languages			Special education		
	Log coef	SE	p	Log coef	SE	p	Log coef	SE	p
Intercept	− 2.698	0.461	< .001***	− 8.566	0.545	< .001***	− 5.988	0.661	< .001***
Pre-adoption slope	0.042	0.017	.014*	0.029	0.023	.200	− 0.037	0.019	.048*
Post-adoption slope	0.042	0.034	.209	0.097	0.051	.057	0.044	0.035	.214
County salary expenditures[a]	− 0.002	0.002	.296	− 0.005	0.003	.113	0.004	0.003	.112
State per capita income[b]	< − 0.001	< 0.001	.355	< 0.001	< 0.001	.225	< − 0.001	< 0.001	< .001***
Undergraduate enrollment	< − 0.001	< 0.001	.265	< 0.001	< 0.001	.247	< 0.001	< 0.001	.003**
Tuition and fees	− 0.006	0.003	.109	< 0.001	0.005	.937	− 0.023	0.005	< .001***
Treatment indicator	0.399	0.412	.355	0.210	0.530	.691	3.437	0.519	< .001***
Pre-adoption slope by treatment	− 0.043	0.021	.039*	0.023	0.028	.426	− 0.023	0.021	.290
Post-adoption slope by treatment	0.095	0.040	.018*	− 0.061	0.060	.316	0.037	0.040	.361
County salary expenditures[a] by treatment	0.002	0.003	.454	0.006	0.004	.119	− 0.007	0.003	.028*
State per capita income[b] by treatment	< 0.001	< 0.001	.700	< − 0.001	< 0.001	.232	< 0.001	< 0.001	.012*
Undergraduate enrollment by treatment	< − 0.001	< 0.001	.332	< − 0.001	< 0.001	.528	< − 0.001	0.006	.021*
Tuition and fees by treatment	− 0.004	< 0.001	.396	− 0.005	0.006	.361	0.009	0.006	.121

Intercept is set at 2009. All covariates are grand mean centered

[a] Per \$10,000, [b]: Per \$1,000

*$p < .05$; **$p < .01$; ***$p < .001$

classrooms. While perhaps not enough to meet all supply needs, the impact would be significant.

In investigating the impact of TEACH on high-need content education degrees, we found the lack of an impact in STEM to be consistent with previous findings from the SMART grant, intended to increase the number of STEM majors. Even when presented with funding and a higher earning potential in a STEM field, there was little evidence of increased degree production. Similarly, findings in language-related areas are consistent with overall decline in enrollment in language-related courses since 2006 in the USA (for a detailed account see Looney and Lusin 2018). In addition, the lack of measured impact on special education degrees may be a function of when individuals seek degrees in special education. Considerable US recruitment efforts in special education are focused on post-baccalaureate and graduate degrees (e.g., Boe 2014; Galloway 2020), neither of which was the focus of the current analysis and as such, requires consideration in future research.

Policy Implications

For TEACH to have greater efficacy, policymakers might consider several structural changes. First, the full award value ($4,000), along with the reported average annual disbursement (~ $2,900), falls below the threshold believed to influence the consequential decision of selecting or changing a major (Denice 2021; Stinebrickner and Stinebrickner 2014). Previous reviews and empirical research on fiscal incentives suggest the dollar amount must be greater than the annual ceiling of $4,000 offered by TEACH. For example, a recent analysis by Feng (2020) found approximately $7000 is necessary to attract and retain teachers in high-need schools (i.e., schools serving largely students of color and low-income families). In addition, work by (Peyton et al., 2021) analyzed differences between states with shortages of qualified special education teachers, finding that states with lower shortages tended to pay teachers approximately $7,700 more than states with higher shortages. Even with an ever-increasing average annual debt for college undergraduates (>$26,500; Hershbein and Hollenbeck 2014), the trade-off of selecting into a teaching major does not appear to be one US undergraduates are willing to make. Therefore, for TEACH to be an effective recruitment tool, policymakers must consider increasing the dollar amount awarded.

Next, in the process of our research we found it troubling that the annual average award amount fell short of the $4000 available for TEACH recipients. That is, most TEACH participants received awards less than the amount available (on average, $2,900). We were able to verify through government documentation TEACH is considered to be a *first dollar*, rather than a *last dollar,* grant. A first dollar grant, as explained by the US Association of Community College Trustees (https://www.acct.org/page/first-dollar-vs-last-dollar-promise-models), allows for the full fiscal funds to be applied to student need before any other grant or support (e.g., Pell, state, or institutional aide), does not take into account student eligibility for additional funding, and has the potential to reduce associated costs of a student (e.g., transportation or child care).

In contrast, a last-dollar grant requires students to draw upon other public funding before an award is given, the total amount varies based upon student eligibility for other sources of funding, and does not allow for use with other associated student costs. The fluctuating dollar disbursement suggests that TEACH is indeed a last dollar grant, meaning, the fiscal award for each student could vary a great deal. If, for example, the annual award was < $1000 for a given student, it would seem unlikely that such a small dollar amount would entice undergraduates to take up a high-need content area that was not previously of interest to them. Moreover, the small dollar amount is unlikely to keep recipients in hard-to-staff school amidst troubling working conditions, as the work of Feng (2020) and others suggests. As a step toward increasing the award for recipients, policymakers must also consider the process in which TEACH is applied within the loan package.

If the expectation is that TEACH is meant to attract individuals into high-need education majors, there should be maximum flexibility in how the funds can be used in the calculation of college costs. Moreover, the fiscal needs of undergraduates may extend beyond tuition costs and can include transportation and childcare costs, for example, as well as the high costs of state-required certification tests (e.g., the Praxis or the edTPA). These are substantive barriers to education degree attainment that TEACH funds could assist with, and require further examination. Additionally, as education majors spend considerable time in classrooms during preparation, providing opportunities for that time to count toward the service loan payback period may be a critical change. It may be that many potential recipients see the time required for student teaching and other classroom-based experiences as prohibitive. Allowing these experiences to "count" toward payback of the loan provides a no-cost opportunity to increase the value of TEACH.

Finally, in the event monetary allocations can be increased to better represent research on effective fiscal incentives, there is still the ongoing implementation challenge of TEACH. Previous research on the implementation of TEACH, and the similarly designed SMART grant, suggests that the variance in the ways in which these initiatives are marketed may play a role in their effectiveness. In interviews with IHE staff about the implementation of TEACH, for example, Barkowski and colleagues (2018) found that some institutions reported a high degree of collaboration between the financial aid office and the education department in targeting the appropriate programs. In contrast, other IHEs reported that the financial aid office was solely responsible and there was little to no coordination with the education department. Clearly, the degree to which there is synergy between these departments is likely to affect initial and repeated exposure of the requirements and benefits of TEACH to interested parties. The overall and mixed outcomes of TEACH might also reflect program marketing, student financial need, or a general lack of interest in taking up teaching as a profession.

Limitations and Future Steps

Our analysis provides an initial foray into the impact of TEACH on undergraduate educator supply. Our work comes with several limitations, but also directions in need of further inquiry. First, the analysis examined TEACH from a national perspective, potentially missing nuances that may exist within geographic regions and across different types of institutions. Future research should describe the number and amount of annual disbursement across geographical regions, types of intuitions, (e.g., public, private, for-profit) of higher education, and sub-groups of students (e.g., race and gender) Moreover, further research should examine at what dollar amount or percentage of tuition, TEACH becomes a more attractive option for undergraduates across sub-populations.

A second limitation is our focus on TEACH participation solely among undergraduates, eliminating individuals who enter teaching at the post-baccalaureate or graduate levels. We acknowledge that this blunts the impact of our analysis; however, we believe there to be distinct differences in the undergraduate and graduate populations and require close contextual examination along with analysis of TEACH outcomes. For instance, graduate students tend to enter into education from different pathways, with varied fiscal needs (e.g., reduced federal funding at the graduate level and lesser financial support from family members); therefore, response to TEACH may vary in comparison with undergraduates in important ways. That being said, further study of the availability of TEACH grants to influence teacher production at the graduate level is much needed.

The third limitation is an extension of the second. While our focus was on undergraduate degree production, not all IHEs included in our sample may have provided undergraduate degree programs in education, generally, or in some of the high-needs content areas, more specifically. Some of these institutions may have been small liberal arts colleges without their own education programs, but may offer joint programs with neighboring IHEs. Other institutions may offer education programs at the graduate but not the undergraduate level or may not offer degree programs in certain areas, such as special education. Although we note that as a limitation, it is likely such institutions are dispersed over both the adopting and non-adopting IHEs.

Lastly, a set of limitations relates to our propensity score model. Determining causal inference from observational studies, such as the present study, is inherently different from causal inference from intervention studies (Shadish 2010). The aim of a propensity score model is to equate groups on covariates as to resemble the expectation of similarity in randomized designs. While we included many known potential confounders to equalize the two groups of IHEs in our propensity score model, it is possible the model missed other influential covariates. Especially since several pre-adoption slopes by treatment interactions were statistically significant. Additionally, the adopting and non-adopting IHEs were just balanced (SMD = .249) on the overall number of undergraduates enrolled, and this variable remained a significant predictor of total degrees and special education degrees awarded. Finally, we calculated the propensity scores based on numbers two years before students receiving TEACH support graduated (i.e., 2007), but it is conceivable trends from a longer period influenced the decision of IHEs to adopt TEACH; for example, declining enrollment in their teacher preparation programs.[1]

[1] On recommendation of a reviewer, we evaluated this limitation by conducting additional analyses including percent decline in education degrees conferred between 2001 and 2007 (i.e., the first year of data until the decision year) as a covariate in the propensity score model. Outcomes are presented in supplemental Tables 2 and 3. Weighting methods did not lead to acceptable balance. Matching reached

Conclusion

Fiscal incentives for teacher candidates to earn a degree in a high-need content area (i.e., STEM, language-related areas, and special education) and to teach low-income schools are a necessary policy tool to meet demand. Using publicly available data, we analyzed the limited impact TEACH has had on the production of undergraduate education degrees overall and in high-need content areas since its adoption in 2009. The propensity score modeling did not find an overall impact on education degrees produced between institutions, and no impact on the recruitment of undergraduates into STEM, language-related areas, or special education teaching fields. Complicating factors such as the recognition, availability, and administration of TEACH grants have made the impact of this policy lever negligible in meeting supply needs in the US teacher labor market for high demand positions.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1057/s41307-022-00263-3.

References

Aldeman, C. (2021). *During the pandemic, 'lost' education jobs aren't what they seem.* Brookings Institute. https://www.brookings.edu/blog/brown-center-chalkboard/2021/03/02/during-the-pandemic-lost-education-jobs-arent-what-they-seem/

Austin, P.C. (2011) An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research 46*(3): 399–424. https://doi.org/10.1080/00273171.2011.568786

Barkowski, E., Nielsen, E., Noel, H., Dodson, M., Sonnenfeld, K., Ye, C., DeMonte, E., Monahan, B., and Eccleston, M. (2018). *Study of the teacher education assistance for college and higher education (TEACH) grant program.* https://www2.ed.gov/rschstat/eval/highered/teach-grant/final-report.pdf

Bates, D., Mächler, M., Bolker, B. and Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software 67*(1): 1–48. https://doi.org/10.18637/jss.v067.i01

Bedard, K. and Herman, D.A. (2008) Who goes to graduate/professional school? The importance of economic fluctuations, undergraduate field, and ability. *Economics of Education Review 27*(2): 197–210. https://doi.org/10.1016/j.econedurev.2006.09.007

Boe, E. (2014) 'Teacher demand, supply, and shortage in special education: A national perspective', in P. Sindelar, E. McCray, M.T. Brownell and B. Lignugaris-Kraft (eds.) *Handbook of research on special education teacher preparation*New York, NY: Routledge, pp. 67–93.

Brooks, M.E., Kristensen, K., van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., et al. (2017) GlmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal 9*(2): 378–400

Footnote 1 (continued)
acceptable quality of propensity scores and matches, with 6 covariates remaining about the 0.2 threshold. Even with percent decline balanced across adopting and non-adopting institutions, the pre-adoption slope differed significantly across both groups indicating there were still unknown influential covariates. Parameter estimates for post-adoption slope by treatment were statistically significant overall and for foreign languages, a change from our non-significant findings. The results from our original analyses should thus be taking with caution.

Carver-Thomas, D. (2018). *Diversifying the teaching profession: How to recruit and retain teachers of color.* https://learningpolicyinstitute.org/sites/default/files/product-files/Diversifying_Teaching_Profession_REPORT_0.pdf

Chen, R. and Wiederspan, M. (2014) Understanding the determinants of debt burden among college graduates. *The Journal of Higher Education 85*(4): 565–598

Cowan, J., Goldhaber, D., Hayes, K. and Theobald, R. (2016) Missing elements in the discussion of teacher shortages. *Educational Researcher 45*(8): 460–462. https://doi.org/10.3102/0013189X16679145

Dee, T. S. & Goldhaber, D. (2017). *Understanding and addressing teacher shortages in the United States.* https://www.brookings.edu/research/understanding-and-addressing-teacher-shortages-in-the-united-states/"

Dembicki, M. (2020, May 19). Restoring the workforce to revive the economy. *Community College Daily.* https://www.ccdaily.com/2020/05/restoring-the-workforce-to-revive-the-economy/

Denice, P.A. (2021) Choosing and Changing Course: Postsecondary Students and the Process of Selecting a Major Field of Study. *Sociological Perspectives 64*(1): 82–108

Denning, J.T. and Turley, P. (2017) Was that SMART? Institutional financial incentives and field of study. *Journal of Human Resources 52*(1): 152–186. https://doi.org/10.3368/jhr.52.1.0414-6340R1

DesJardins, S.L., Kim, D.O. and Rzonca, C.S. (2003) A nested analysis of factors affecting bachelor's degree completion. *Journal of College Student Retention: Research, Theory & Practice 4*(4): 407–435

Doyle, W. (2009) The effect of community college enrollment on bachelor's degree completion. *Economics of Education Review 28*(2): 199–206. https://doi.org/10.1016/j.econedurev.2008.01.006

Evans, B. (2016) SMART Money: Do financial incentives encourage college students to study science? *Education Finance and Policy*. https://doi.org/10.1162/EDFP_a_00

Feng, L. (2020). Compensating differentials in teacher labor markets. In S. Bradley & C. Green (Eds.), *The economics of education: A comprehensive overview* (2nd ed., pp. 415-430). Elsevier. https://doi.org/10.1016/C2017-0-2304-2

Feng, L. and Sass, T. (2017) The impact of incentives to recruit and retain teachers in "hard-to-staff" subjects. *Journal of Policy Analysis and Management 37*(1): 112–135. https://doi.org/10.1002/pam.22037

Fortson, K., Verbitsky-Savitz, N., Kopa, E., and Gleason, P. (2012). *Using an experimental evaluation of charter schools to test whether nonexperimental comparison group methods can replicate experimental impact estimates.* Washington, DC: US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://files.eric.ed.gov/fulltext/ED531481.pdf

Galloway, M. (2020). Human resources special education teacher recruitment: Utilizing grow your own. *Journal of Transformative Leadership & Policy Studies, 8* (2), 65-87. https://doi.org/10.36851/jtlps.v8i2.2227

Garcia, E. and Weiss, E. (2019). *U.S. schools stage to hire and retain teachers.* Economic Policy Institute. https://www.epi.org/publication/u-s-schools-struggle-to-hire-and-retain-teachers-the-second-report-in-the-perfect-storm-in-the-teacher-labor-market-series/

Guarino, C.M., Santibanez, L. and Daley, G.A. (2006) Teacher recruitment and retention: A review of the recent empirical literature. *Review of Educational Research 76*(2): 173–208. https://doi.org/10.3102/00346543076002173

Goff, P., Carl, B., and Yang, M. (2018). *Supply and demand for public school teachers in Wisconsin* (WCER Working Paper No 2018-2). https://wcer.wisc.edu/docs/working-papers/Working_Paper_No_2018_02.pdf

Goldhaber, D., Krieg, J., Theobald, R. and Brown, N. (2015) Refueling the STEM and special education teacher pipelines. *Phi Delta Kappan 97*(4): 56–62

Gurel, S., and Leite, W. L. (2014, April). *Evaluation of propensity score strategies with multilevel data when treatment assignment mechanism varies between clusters.* Paper presented at the annual meeting of the American Educational Research Association, Philadelphia, PA.

Hansen, B.B. and Klopfer, S.O. (2006) Optimal full matching and related designs via network flows. *Journal of Computational and Graphical Statistics 15*(3): 609–627

Harder, V.S., Stuart, E.A. and Anthony, J.C. (2010) Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychological Methods 15*(3): 234

Hershbein, B. J., and Hollenbeck, K. (2014). *College costs: Students can't afford not to know.* W.E. Upjohn Institute for Employment Research. https://research.upjohn.org/cgi/viewcontent.cgi?article=1069&context=externalpapers

Hillman, N.W., Tandberg, D.A. and Fryar, A.H. (2015) Evaluating the impacts of "new" performance funding in higher education. *Educational Evaluation and Policy Analysis 37*(4): 501–519

Ho, D.E., Imai, K., King, G. and Stuart, E.A. (2011) MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software 42*(8): 1–28

Humphreys, B.R. (2000) Do business cycles affect state appropriations to higher education? *Southern Economic Journal 67*(2): 398–413. https://doi.org/10.2307/1061477

Kelcey, B. (2011) Assessing the effects of teachers' reading knowledge on students' achievement using multilevel propensity score stratification. *Educational Evaluation and Policy Analysis 33*(4): 458–482

Leite, W. (2016) *Practical propensity score methods using R*, Sage Publications: Sage Publications.

Leite, W.L., Jimenez, F., Kaya, Y., Stapleton, L.M., MacInnes, J.W. and Sandbach, R. (2015) An evaluation of weighting methods based on propensity scores to reduce selection bias in multilevel observational studies. *Multivariate Behavioral Research 50*(3): 265–284

Li, F., Zaslavsky, A.M. and Landrum, M.B. (2013) Propensity score weighting with multilevel data. *Statistics in Medicine 32*(19): 3373–3387

Looney, D. and Lusin, N. (2018). *Enrollments in languages other than English in United States institutions of higher education, Summer 2016 and Fall 2016: Preliminary Report*. Modern Language Association. Retrieved from https://www.mla.org/content/download/ 83540/2197676/2016-Enrollments-Short-Report.pdf

Lovenheim, M. and Turner, S. (2018) *Economics of education*, New York, NY: Worth Publishers.

McLendon, M.K., Hearn, J.C. and Mokher, C.G. (2009) Partisans, professionals, and power: The role of political factors in state higher education funding. *The Journal of Higher Education 80*(6): 686–713

Nowicki, J. M. (2015). *Better Management of federal grant and loan forgiveness programs for teachers needed to improve participant outcomes.* United Sates Government Accountability Office. https://www.gao.gov/assets/670/668634.pdf

Ondrasek, N., Carver-Thomas, D., Scott, C., and Darling-Hammond, L. (2020). California's special education teacher shortage. Learning Policy Institute. https://edpolicyinca.org/sites/default/files/2020-02/r_ondrasek_feb20.pdf

Partelow, L. (2019). *What to make of declining enrollment in teacher preparation programs?* Center for American Progress. https://www.americanprogress.org/issues/education-k-12/reports/2019/12/03/477311/make-declining-enrollment-teacher-preparation-programs/

Peyton, D. J., Acosta, K., Harvey, A., Pua, D. J., Sindelar, P. T., Mason-Williams, L., Dewey, J., Fisher, T. L.,& Crews, E. (2021). Special education teacher shortage: Differences between high and low shortage states. *Teacher Education and Special Education*, *44*(1), 5–23. https://doi.org/10.1177/0888406420906618

Podolsky, A., Kini, T., Bishop, J., Darling – Hammond, L. (2016). *Solving the teacher shortage: How to attract and retain excellent educators.* Learning Policy Institute. https://learningpolicyinstitute.org/sites/default/files/product-files/Solving_Teacher_Shortage_Attract_Retain_Educators_REPORT.pdf

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rodríguez, A., & Galdeano, E. C. (2015). Do Hispanic-Serving Institutions really underperform? Using propensity score matching to compare outcomes of Hispanic-Serving and non–Hispanic-Serving Institutions. In Hispanic-Serving Institutions (pp. 210-230). Routledge

Rosenbaum, P.R. and Rubin, D.B. (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1): 41–55

Shadish, W.R. (2010) Campbell and Rubin: A primer and comparison of their approaches to causal inference in field settings. *Psychological Methods 15*(1): 3–17. https://doi.org/10.1037/a0015916

Strauss, V. (2018). Is congress about to gut funding for teacher preparation programs. *Washington Post.* https://www.washingtonpost.com/news/answer-sheet/wp/2018/05/22/is-congress-about-to-gut-funding-for-teacher-preparation-programs/

Stinebrickner, R. and Stinebrickner, T.R. (2014) A major in science? Initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies 81*(1): 426–472

Sutcher, L., Darling-Hammond, L., and Carver-Thomas, D. (2019). Understanding teacher shortages: An analysis of teacher supply and demand in the United States. *Education Policy Analysis Archives*. https://doi.org/10.14507/epaa.27.3696

U.S. Department of Education, Federal Student Aid. (2020). *Title IV program volume report, grant volume*. https://studentaid.gov/data-center/student/title-iv

Weerts, D.J. and Ronca, J.M. (2012) Understanding differences in state support for higher education across states, sectors, and institutions: A longitudinal study. *The Journal of Higher Education 83*(2): 155–185