Student Learning and Teacher Retention for Graduates of Texas Noyce Programs[1]

Michael Marder
The University of Texas at Austin

Catherine Horn
University of Houston

Sarah Stephens
University of Wisconsin-Milwaukee
&
Annelies Rhodes
The University of Texas at Austin
United States

Abstract: The shortage of secondary STEM teachers has led to periodic calls over the past four decades for federal intervention. For more than 15 years, the National Science Foundation

Robert Noyce Teacher Scholarship Program has constituted the most important response at the national level to these calls. The main activity is direct support to undergraduate students and degree holders to become K-12 STEM teachers. Here we examine the Noyce program by looking beyond the qualities of particular university programs and measuring long-term outcomes on a statewide level. We accomplished this through a collaboration of eight Texas universities that pooled data on Noyce scholars supported as far back as 2003. Making use of a state longitudinal dataset, we examined whether recipients of Noyce Scholarships are more likely than other math and science teachers to remain in teaching, determined where Noyce scholars taught, and estimated student learning in classes taught by Noyce Scholars. Noyce Scholars are more likely than other STEM teachers from the same universities to teach marginalized students, and students of Noyce Scholars obtained higher value-added scores in mathematics than students of other teachers at the same schools. These are some of the most important intended outcomes. However, Noyce Scholars did not stay in teaching as long as other STEM teachers from their universities, and they were more likely to switch away from schools with low-income students. The Noyce program currently operates far below the scale needed to start reducing national STEM teacher shortages. We conclude with prospects for expansion and modification of the Noyce program.

Keywords: teacher preparation; Noyce scholarships; hierarchical linear models; STEM education

Aprendizaje de estudiantes y retención de maestros para graduados de los programas de Texas Noyce

Resumen: La escasez de maestros STEM de secundaria ha llevado a llamadas periódicas durante las últimas cuatro décadas para la intervención federal. Durante más de 15 años, el Programa de Becas para Maestros Robert Noyce de la Fundación Nacional de Ciencias ha constituido la respuesta más importante a nivel nacional a estas convocatorias. La actividad principal es el apoyo directo a estudiantes de pregrado y titulados para que se conviertan en maestros STEM K-12. Aquí examinamos el programa Noyce mirando más allá de las cualidades de los programas universitarios particulares y midiendo los resultados a largo plazo a nivel estatal. Logramos esto a través de una colaboración de ocho universidades de Texas que recopilaron datos sobre los becarios de Noyce que recibieron apoyo desde 2003. Haciendo uso de un conjunto de datos longitudinales estatales, examinamos si los beneficiarios de las Becas de Noyce tienen más probabilidades que otros maestros de matemáticas y ciencias de permanecer en enseñanza, determinó dónde enseñaron los becarios de Noyce y estimó el aprendizaje de los estudiantes en las clases impartidas por los becarios de Noyce. Los Noyce Scholars tienen más probabilidades que otros docentes de STEM de las mismas universidades de enseñar a estudiantes marginados, y los estudiantes de Noyce Scholars obtuvieron puntajes de valor agregado más altos en matemáticas que los estudiantes de otros docentes en las mismas escuelas. Estos son algunos de los resultados esperados más importantes. Sin embargo, los becarios de Noyce no permanecieron en la enseñanza tanto tiempo como otros docentes de STEM de sus universidades, y era más probable que abandonaran las escuelas con estudiantes de bajos ingresos. El programa Noyce actualmente opera muy por debajo de la escala necesaria para comenzar a reducir la escasez nacional de docentes STEM. Concluimos con perspectivas de ampliación y modificación del programa Noyce.

Palabras-clave: formación docente; becas Noyce; modelos lineales jerárquicos; educación STEM

Aprendizagem do aluno e retenção de professores para graduados dos programas Texas Noyce

Resumo: A escassez de professores STEM secundários levou a chamadas periódicas nas últimas quatro décadas para intervenção federal. Por mais de 15 anos, o Programa de Bolsas de Estudo para Professores Robert Noyce da National Science Foundation constituiu a resposta mais importante em nível nacional a essas chamadas. A principal atividade é o apoio direto a estudantes de graduação e diplomados para se tornarem professores K-12 STEM. Aqui examinamos o programa Noyce olhando além das qualidades de programas universitários específicos e medindo os resultados de longo prazo em nível estadual. Conseguimos isso por meio de uma colaboração de oito universidades do Texas que reuniram dados sobre acadêmicos Noyce apoiados desde 2003. Usando um conjunto de dados longitudinal estadual, examinamos se os beneficiários das bolsas Noyce são mais propensos do que outros professores de matemática e ciências a permanecer em ensino, determinado onde os estudiosos Noyce ensinavam e aprendizado estimado dos alunos em aulas ministradas por Noyce Scholars. Os Noyce Scholars são mais propensos do que outros professores de STEM das mesmas universidades a ensinar alunos marginalizados, e os alunos de Noyce Scholars obtiveram pontuações de valor agregado mais altas em matemática do que os alunos de outros professores nas mesmas escolas. Estes são alguns dos resultados pretendidos mais importantes. No entanto, os Noyce Scholars não permaneceram no ensino por tanto tempo quanto outros professores STEM de suas universidades, e eram mais propensos a se afastar das escolas com alunos de baixa renda. O programa Noyce atualmente opera muito abaixo da escala necessária para começar a reduzir a escassez nacional de professores STEM. Concluímos com as perspectivas de expansão e modificação do programa Noyce.

Palavras-chave: preparação de professores; Bolsas Noyce; modelos lineares hierárquicos; educação STEM

Student Learning and Teacher Retention for Graduates of Texas Noyce Programs

The shortage of secondary STEM teachers has constituted a persistent national concern at least since *A Nation at Risk* (National Commission on Excellence in Education, 1983), and teacher shortages remain severe today (Sutcher et al., 2019). Districts and teacher preparation programs consistently report considerable shortages of STEM teachers, particularly physics, chemistry, and mathematics (American Association for Employment in Education, 2020). Other than some categories of special education teachers, teachers in these subject areas are the most difficult to hire. The shortages of STEM teachers explain and exacerbate low levels of STEM course availability in high school.

What policies are necessary to reduce teacher shortages? According to Darling-Hammond and Podolsky (2019, p. 9), countries that successfully address teacher shortages have the following in common:

- Equitable funding of schools
- Teacher compensation competitive with other college-educated professions
- High-quality preparation available at little or no cost to entering teachers
- Careful recruitment of candidates with the commitment and dispositions for teaching, as well as academic background
- Readily available support from trained mentors for beginning teachers

- Ongoing time and support for professional learning and collaboration.

The United States has never implemented a coordinated national effort with all these ingredients. However, the US does have a national program that is focused on the third and fourth items on the list: teacher preparation and recruitment. The Robert Noyce Scholarship program, which has been in operation for more than 15 years, is a national effort to attract students into STEM teaching by providing them financial support during preparation. STEM teacher shortages have not been eliminated during the time the Noyce program has been active, but the effort loses little of its significance even if it has not yet fully achieved its aims. Our purpose here is to take stock of the program's accomplishments, compare outcomes to aims, and then return to the implications for national teacher shortages.

Origins and Characteristics of the Noyce Scholarship Program

Origins

The Noyce Scholarship program is the product of a view of education reform that was particularly impactful in the early 2000's. The narrative was forcefully articulated in *The World is Flat* (T. Friedman, 2005), which warned that the United States could lose its economic competitiveness if our students did not improve their performance in science and mathematics. The National Academies commissioned a study, *Rising Above the Gathering Storm* (National Research Council, 2005), that revisited the argument in terms of specific legislative recommendations. Among proposals to shore up the economic future of the United States, the top recommendation was to prepare "ten thousand teachers for ten million minds." We quote from the original report to highlight the tenor, scale, and details of the proposal:

> Action A-1: Annually recruit 10,000 science and mathematics teachers by awarding 4-year scholarships and thereby educating 10 million minds. Our public education system must attract at least 10,000 of our best college graduates to the teaching profession each year. A competitive federal scholarship program would allow bright, motivated students to earn bachelors' degrees in science, engineering, and mathematics with concurrent certification as K–12 mathematics and science teachers.
>
> Students could enter the program at any of several points and would receive annual scholarships of up to $20,000 per year in the program for tuition and qualified educational expenses. Awards would be given on the basis of academic merit. Each scholarship would carry a 5-year postgraduate commitment to teach in a public school. The annual investment in such scholarships at steady state would be $400 million to $800 million. (National Research Council, 2005 p. 115)

Toward this end, since 2002, Congress has appropriated funds for the Robert F. Noyce Scholarship Program, administered through the National Science Foundation (NSF). Figure 1 sketches a theory of action for the Scholarship Program, displaying some of the rationales for program choices, inputs assumed to be available, main activities carried out, indicators used to monitor success, and desired outcomes.

Figure 1

Theory of Action for the Noyce Scholarship Program

Rationales	Inputs	Activities	Indicators	Outcomes
Project leadership must include STEM faculty as well as Education faculty to ensure engagement of STEM college and majors	College of Natural Sciences faculty committed to STEM teacher preparation	Provide reports on program completers to check on teaching obligation	Numbers of students enrolling in STEM educator preparation program	Improved opportunity for students to take STEM courses in high-need schools.
Funding will enable creation or improvement of STEM teacher preparation pathways	College of Education faculty collaborating with Natural Sciences faculty on STEM teacher preparation	Create STEM teaching preparation program or new component of STEM teaching preparation program	Numbers of students completing STEM educator preparation program	Improved quality of STEM coursework in high-need schools
Payback provisions in support will increase the likelihood of program completers teaching in high-need schools	STEM majors, either undergraduates or degree holders, willing to consider teaching as a profession	Award and administer Noyce scholarships	Numbers of program completers who teach in high-needs schools	Equity in education
Student financial support will increase the quantity and quality of STEM students and degree holders who become teachers	NSF Noyce Program funding for Scholarships and Stipends	Improve quality of STEM teaching preparation program	Length of time program completers teach in high-needs schools	Competitiveness in economics

Expenditures

How much has been spent on the Noyce Scholarship Program? Congress authorized $20 million for it as part of the National Science Foundation Authorization Act of 2002 (107th Congress, 2002). On the NSF side, the Noyce program appears for the first time in the request to Congress for the 2003 fiscal year, when $4 million was requested, 20% of $20 million Congress had authorized (National Science Foundation, 2003), and 1% of the minimum advocated by *Gathering Storm*. The Noyce program was authorized again in the 2007 COMPETES Act (*COMPETES Act*, 2007) with $89.8 million authorized to be appropriated per year. Funding never reached the authorized level, but it increased to around $50 million/year by 2010, $58 million/year by 2019, and $67 million/year by 2021.

The total amount spent on Noyce scholarships is not completely straightforward to estimate. If one adds up the total amount in Noyce Requests for Proposals through 2019, the sum comes to $608 million. Alternatively, if one adds the total amount in committed awards to programs the NSF codes as Noyce alone through 2019, the sum comes to $716 million, with another $467 million committed to awards with multiple funding sources including Noyce. Our inspection of the latter award abstracts indicates they do not typically include the scholarships on which we focus here, so we suggest the amount spent on Noyce Scholarship programs through 2019 was around $700 million.

The reason we describe Noyce expenditures through 2019 is that we know that between 2002 and Fall 2019, the Noyce program supported approximately 12,000 STEM teacher candidates (Hobson Hargraves, 2022). Comparison of this number with the roughly $700 million expended

gives some sense of the cost effectiveness of the effort: $60,000 per teacher. We acknowledge that our estimate is imprecise, for we do not know how many of the supported candidates taught, or for long, or in which schools, nor what fractions of the grants went for scholarships or for program building.

Structure

The Noyce Scholarship program has had many tracks with different detailed requirements over the years, but there is a consistent theme and purpose underlying them. We quote a recent NSF description of the Noyce program:

> The Robert Noyce Teacher Scholarship Program responds to the critical need for highly effective K-12 mathematics, science (including computer science), and engineering teachers. The program seeks to encourage institutions of higher education to develop and sustain a culture where undergraduate STEM majors and STEM professionals, especially those of the highest achievement and ability who might otherwise not have considered a career in K-12 teaching, are encouraged and supported to become teachers in high-need local educational agencies (National Science Foundation, 2017).

Here we will focus on tracks providing Scholarships and Stipends. These were the earliest forms of the program and are the types of grants mainly received by the participants in this study. Educator Preparation Programs apply for them and if successful receive roughly half a million to 1.5 million dollars over five years to support scholarships and program improvement. Noyce Scholarship and Stipend grants must include scholarships for students pursuing STEM teacher certification. For the Scholarship program track (up to three years of support for undergraduates) and Stipend program track (up to one year of support for degree holders), 60% of the money must be spent on student support. Degree holders are allowed to receive a year of support, while undergraduates pursuing certification can receive up to three years. It is permitted to provide scholarships as large as the full cost of attendance, which is around $25,000/year for undergraduates paying in-state tuition, although many institutions offer smaller scholarships to increase the number of students they are able to support. Thus, because of the desire to spread the funds among many students, the programs in this study gave awards of around $10,000 to $15,000 per year per student, which we believe is typical of how universities across the country have behaved. Funding not spent on student scholarships is spent on program development.

Noyce Scholarships and Stipends come with a payback provision, which is dictated by the authorizing legislation. The recipient must teach in a high-need Local Education Agency (LEA) two years for each year of financial support, or else pay back the scholarships, pro-rated, to the Treasury. An LEA is typically a school district, and a district can be high-need even if not all schools in it are high need. Thus, the Noyce program addresses teacher retention by requiring a commitment from scholarship recipients as *Gathering Storm* had recommended and backing it with financial consequences.

How Receiving Noyce Grants Changes University Programs and Influences Students

The NSF awards Noyce grants to "develop and implement exemplary STEM education programs." The grants direct funding towards certain types of activities. At least 60% of Noyce Scholarship and Stipend grants must be spent on scholarships and stipends, while the remainder can be used for program enhancements and improvements. This remainder rarely if ever suffices to create a genuinely new program. The eight universities involved this study used Noyce grants for program enhancement, not program creation. One may speak of students who went through the

Noyce program but the main difference between Noyce Scholars and other students at their universities is that the Noyce Scholars get Noyce Scholarships. Otherwise, they take the same coursework and receive the same services as their peers. In addition, Noyce program enhancements are typically available to all students in the teacher preparation program. Therefore, when comparing Noyce Scholars to other teacher candidates from the same universities, one should not be imagining that coursework or services led them to be different.

This does not mean one should expect Noyce Scholars to be the same. The NSF directs universities to select students of "highest achievement and ability." Typically, this requirement is met by choosing students already enrolled in the teacher preparation program with high grade-point averages and who have written a good essay. The NSF also asks for candidates "who might otherwise not have considered a career in K-12 teaching," but this is harder to determine. In any event, differences between Noyce Scholars and other STEM teachers from the same universities can be attributed to a mixture of selection of students who would have pursued teaching anyway and recruitment of students who would not.

The NSF Noyce program solicitations support both creating new high-quality programs or program components, and searching for excellent programs that already exist and providing scholarships to increase their student numbers. The institutions in this study mainly used the funds for student scholarships, with some funds leading to program enhancements for all students, but not major programmatic changes.

Noyce grants can have major programmatic implications because of the challenging environment surrounding teacher preparation in the United States. Eminent universities including Cornell have eliminated their teacher preparation programs completely (Lang, 2010). One of the programs participating in this collaborative project at The University of Texas at Arlington was officially closed in a round of budget cuts and then brought back following public outcry (Brewer, 2019). The substantial NSF funding the program had received and was receiving (Table 1) played a role in the rescue.

A final influence one should expect from Noyce grants concerns where the graduates teach. Because of written commitments they make when they accept funds, one should expect Noyce Scholarship and Stipend recipients to be more likely to teach in high-need LEAs than other graduates of the educator preparation programs at their universities.

Critiques of the Narrative of Competitiveness and Studies of Noyce Programs

Competitiveness

The academic literature directly concerning the Noyce program is sparse, which is perhaps because of a larger body of literature that questions the values and assertions that led to the creation and funding of the program (Hoeg & Bencze, 2017; Sharma, 2016; Sharma & Hudson, 2021; Zeidler, 2016). The Noyce program was originally created and funded to enhance US economic competitiveness (*COMPETES Act*, 2007; *America COMPETES Reauthorization*, 2010), and this framing remains current: the Biden administration claimed Build Back Better (White House, 2022) would "lead to lifelong educational and economic benefits for children and parents, and is a transformational investment in America's future economic competitiveness."

The main objection to competitiveness as a framework for education is that "[t]he discourse related to citizenship embedded in much STEM policy from the United States appears to prioritise augmentation of STEM workforce numbers and economic gains for corporate networks driving science reform, rather than developing democratically grounded citizenship." (Hoeg & Bencze, 2017, p. 857) Specific challenges to the rationale for the Noyce program include arguments that there is no shortage of STEM workers in the US, there is no skills gap affecting US students'

prospects, and that students already take enough STEM classes (Cappelli, 2015). Cappelli is particularly effective in dissecting arguments originating in the business community about the difficulties of hiring workers. He shows that repeated claims that poor school performance put the US at economic risk are grounded in self-interest of businesses and unsupported by workforce data. He says, "Overall, the available evidence does not support the idea that serious skill gaps or skill shortages exist in the U.S. labor force. The prevailing situation in the U.S. labor market, as in most developed economies, continues to be skill mismatches in which the average worker and job candidate have more education than their current job requires [p. 281]". Cappelli's arguments apply with particular force to vocational training and to preparation for highly specific jobs; when he brings up the matter of foundational skills, he says "The proposals in the various employer-led reports that we need to increase student academic achievement, reduce drop-out rates, and generally improve the quality of secondary school education are difficult to object to, but these goals are already accepted [p. 281]."

Tougher criticism comes from Sharma (2016), who argues that "educators and educator groups working to disarticulate STEM and rearticulate an alternative progressive vision of school science should benefit from coming together on a common platform to form a global movement against STEM [pp. 47-48]." The NSF is part of the "big and powerful actor-network that exists to support the hegemonic practices, institutions and positive knowledges associated with STEM education [p. 48]." If we accept Sharma's arguments, starting with Cappelli's data, then government support to increase the number of STEM teachers through programs like Noyce is a corporate subsidy and we should oppose it.

Figure 2

Total STEM Certificates, 2010-11 through 2018-19, linear projection through 2029

Secondary STEM Certificates, selected by Area
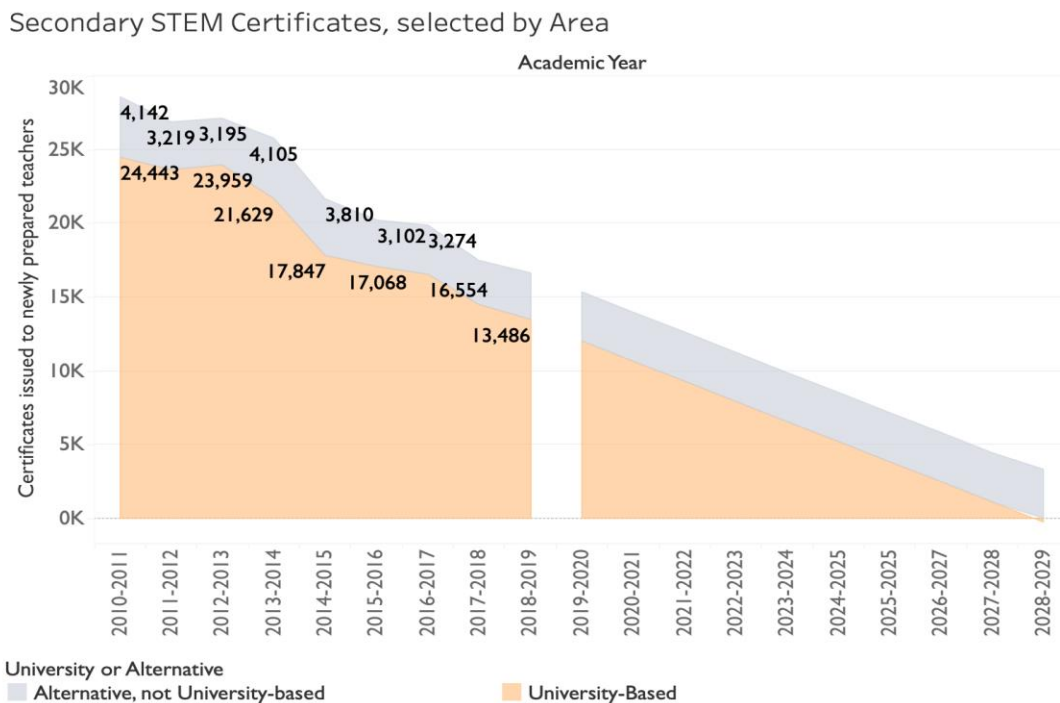
Source: Data from Title II (US Department of Education, 2021)

This point of view overlooks developments in the United States that will undermine broad scientific literacy, impact all students negatively, and particularly hurt those from marginalized groups. Our analysis of national data (US Department of Education, 2021) finds that teacher preparation in the United States is declining at an alarming rate (Partelow, 2019). While 85,386 individuals completed traditional university-based teacher education programs in 2008-2009, by 2018-2019 the number had fallen to 50,692, a 40% drop. In Figure 2 we focus on STEM certificates. The number of STEM certificates issued by universities to new teachers in 2010-2011 was 24,443, while in 2018-2019 it had fallen 45% to 13,486. If the drop in production continues at the pace of the last eight years, Figure 2 shows that in a decade U.S. universities will not be producing secondary STEM teachers at all.

Figure 3

Probability students attend schools without physics and calculus teachers, and probabilities students from different groups take high school calculus, chemistry, and physics
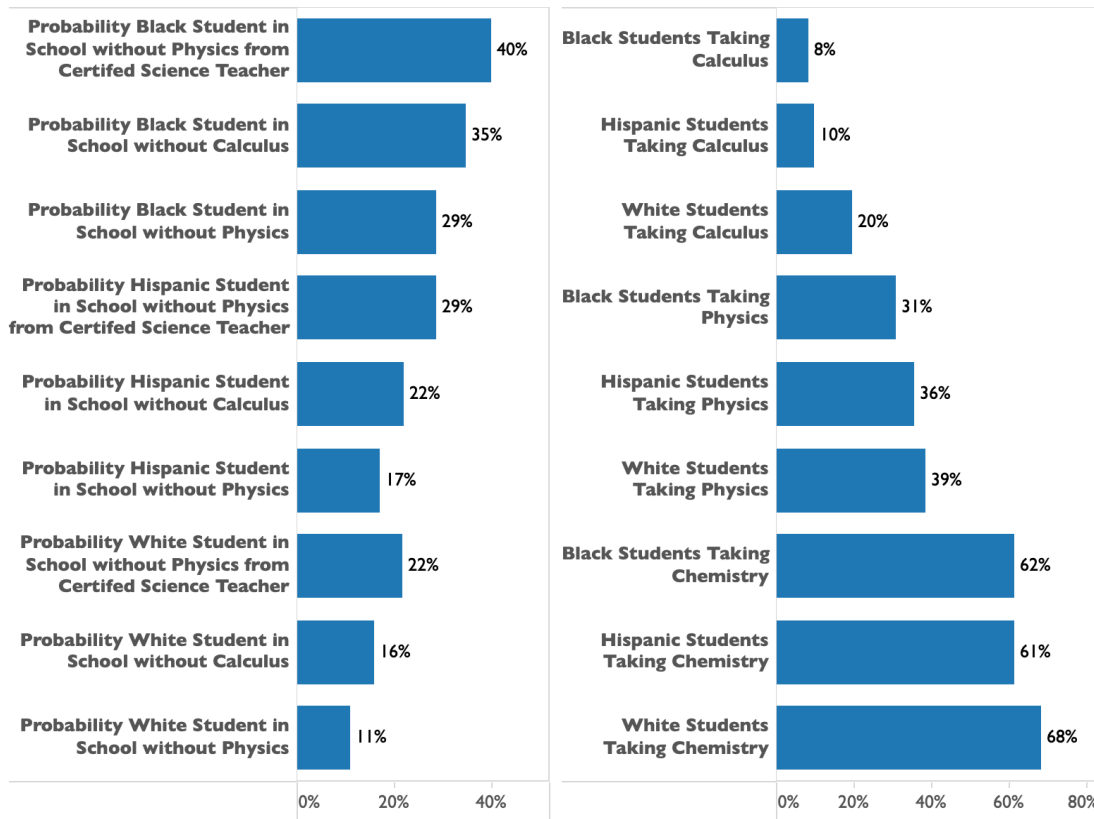
Source: Data from 2017-2018 Civil Rights Data Collection (US Department of Education, 2020)

STEM teacher shortages already exist. School districts report great difficulty in hiring STEM teachers (American Association for Employment in Education, 2020). Shortages of computer science teachers are well documented (Code.org Advocacy Coalition et al., 2021). Specific evidence comes from school-by-school Civil Rights data describing the demographics of students taking coursework (US Department of Education, 2020). As shown in Figure 3, Black and Hispanic students are much more likely than their White counterparts to attend schools completely lacking

basic STEM courses such as physics or calculus. Overall, they take calculus, physics, and chemistry at much lower rates than White students. Thus, a large fraction of the country's Black and Hispanic students are effectively deprived any realistic opportunity of pursuing STEM careers even before they leave high school.

The disparate outcomes for marginalized groups are even more evident when one examines STEM degrees from universities. While Black and Hispanic Americans constitute 14% and 18% of the US population respectively, their share of bachelor's and doctoral degrees awarded in STEM is much smaller than this, as shown in Figure 4. It is these observations, more than the worry that the US cannot meet workforce needs, that motivate us to advocate for more STEM teachers.

Figure 4

Percentage of bachelor's and doctoral degrees in selected STEM areas awarded to various groups in 2018-2019. Total numbers of degrees awarded indicated as labels
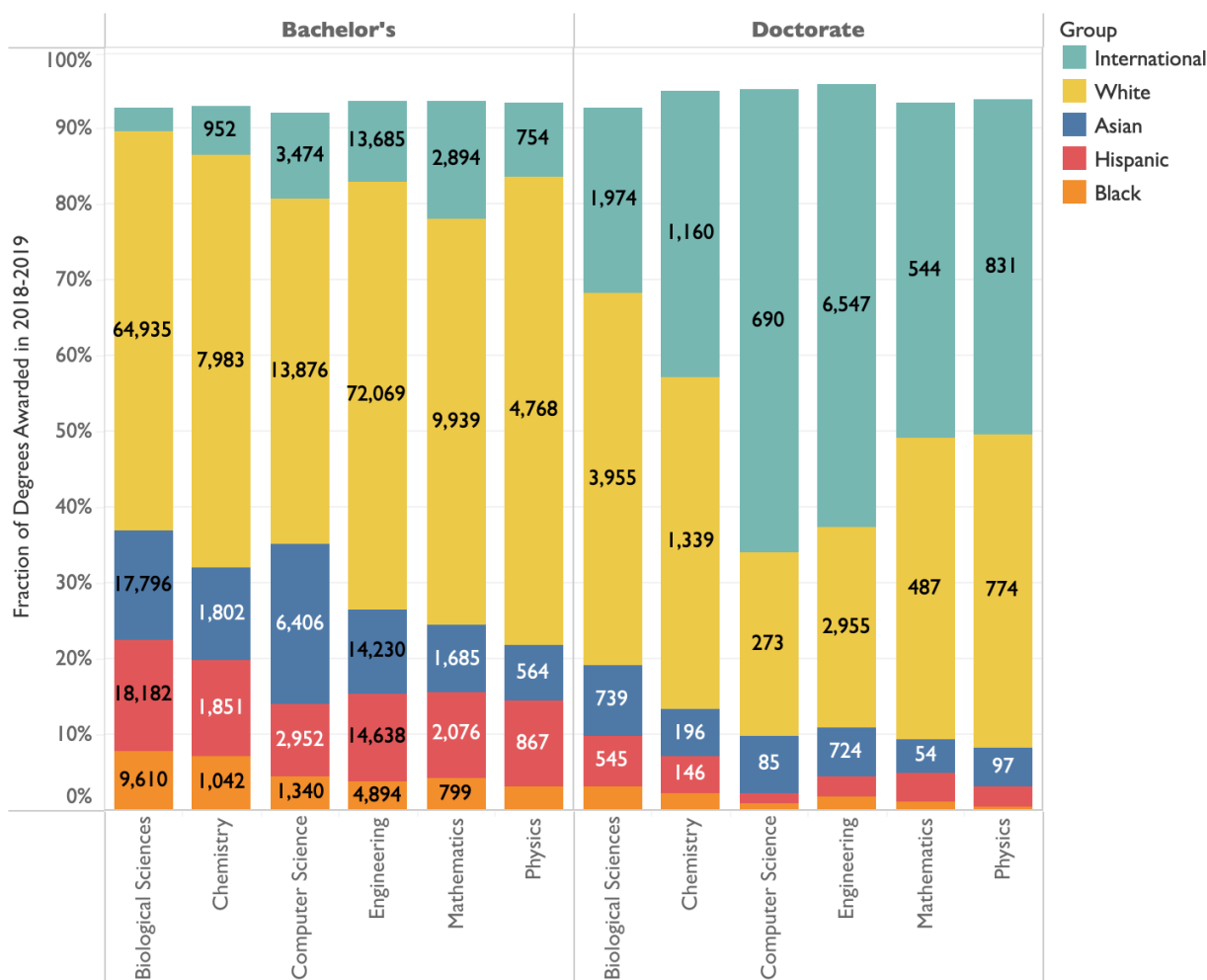
Source: Data for 2018-2019 from the Integrated Postsecondary Education Data System(National Center for Education Statistics, 2019)

There may well not be a great current shortage of STEM workers in the United States (Cappelli, 2015), but if so it is because, as illustrated for doctoral degrees in Figure 4, the US freely

imports STEM talent from around the world. The opportunities available to people growing up in the US are limited and inequitable. This must change, starting with opportunities and enrollment in STEM courses in high school. The US requires more great secondary STEM teachers than we have now, rather than the declining numbers shown in Figure 2. There is no high school science, progressive or otherwise, without mathematics, chemistry, physics, biology, and computer science teachers, and there is no federal program to increase their numbers but Noyce.

Studies of Noyce

Only a few peer-reviewed studies empirically assess the effectiveness of the Noyce Programs or their graduates, and most are more than a decade old.Kirchhoff & Lawrenz (2011) used an inductive grounded theory strategy to understand the role of teacher education on the career paths of 38 Noyce scholarship recipients and concluded that the role of the program in providing support was influential in retaining teachers in high-need settings after the program requirements lapsed. In a survey of 555 scholarship recipients, Liou, Kirchhoff, & Lawrenz (2010) found that Noyce awards increased the commitment of recipients to complete their certification and to teach in high-need schools. They asked whether Noyce recipients were influenced to become teachers by the scholarships and found that 57% said the scholarships were not influential and only 17% said they were very influential. Liou, Desjardins, & Lawrenz (2010) determined that a scholarship recipient's race influenced their commitment to teach and teach in a high-need school, and that the scholarships had little effect on students who entered college committed to teach. Liou & Lawrenz (2011) analysed factors that influence students' pathways to teaching in high-need schools through scholarship and loan forgiveness programs. They determined that the most important factors in these pathways were the scholars' race, the path into teaching, perceptions of preparation for teaching in a high-need school, and the amount of funding.

Liou, Kirchhoff, et al. (2010, p. 467) note that "[a]dditional data need to be collected and future analyses need to be conducted to determine if the scholarship actually affects scholars' commitments to teaching in high need schools and retention in high need schools, not just their perceptions of how the scholarship affected them." One response came from Zahner et al. (2019), who compared persistence in the classroom of 47 Noyce Scholars from Boston University with 47 Teach for America participants in Boston over the period 2009-2015. They found through multiple forms of analysis that Noyce Scholars were much more likely to remain in teaching than participants in Teach for America.

The payback provision of the Noyce program raises the question of whether commitments to teach backed by loan forgiveness are effective in promoting teacher retention. The most authoritative information on this question we have found is Allen (2003), who presented no conclusions and said that we should "undertake more research on and rigorous evaluations of early recruitment efforts, loan forgiveness programs and the many other specific kinds of recruitment strategies that have been employed [p. 122]".

Individual Noyce programs have provided reports on their participants. Scott, Milam, Stuessy, Blount, & Bentz (2006) sought to understand the Math and Science Scholars (MASS) Program, which was partially supported through Noyce Scholarships. They found that the program's efforts to streamline certification along with the supports through tuition remission, scholarships, and other academic experiences resulted in increases in the numbers of undergraduate majors in mathematics and science considering teaching as a career path. Bischoff, French, & Schaumloffel (2014) found in a study of 22 Noyce Scholar science education majors that those committed to teaching in urban New York City schools viewed themselves and the opportunity positively relative to those who were inclined toward rural high-need teaching placements.

Other studies discuss topics such as the effectiveness of a summer intern program in recruiting students to teaching (Mundy & Ratcliff, 2021), professional development for participants in a Master Teacher track (Alemdar et al., 2018), a case study of propensity to teach (Ticknor et al., 2017), the impact of scholarships on STEM teacher recruitment (Morrell & Salomone, 2017), and a comparison of recruitment into regular and alternative preparation programs (Bowe et al., 2011). There has been some attention directed to the quality or characteristics of teaching of the recipients of Noyce Scholarships and Stipends. The teaching practices of Noyce scholars from the UT Austin UTeach program were compared with non-Noyce UTeach graduates and with graduates of other programs using an observational instrument (Walkington & Marder, 2015). The small sample size made it impossible to discern differences between the Noyce and non-Noyce UTeach graduates, although the UTeach graduates overall had practices that could be distinguished from those of other programs.

In summary, the main effort to study the Noyce program has been done by surveying participants while they were still in school (Liou, Desjardins, et al., 2010; Liou et al., 2009; Liou, Kirchhoff, et al., 2010). The surveys provided excellent information on Noyce Scholars' motivation to teach and intention to teach. However, these studies did not address the question of where the Noyce scholars actually taught, how well they taught, how long they stayed in teaching, and whether they taught students in low-income schools. We will address these questions.

Research Questions

Our study arose from a collaborative research grant among eight universities[2] in Texas that received Noyce scholarship support. Data to address these types of research questions are aggregated at the state level in longitudinal data systems, and this is what made it practical to conduct the study in Texas. We had access to all student and teacher data in Texas since 2012, merged with identities of Noyce graduates from multiple universities spanning a period of 15 years. Our specific research questions were:

R1. How did the likelihood of remaining in teaching vary between teachers who received Noyce support and other STEM teachers who graduated from the same universities?

R2. How did the likelihood of remaining in teaching vary between teachers who received Noyce support and other STEM teachers in the same schools?

R3. How did the likelihood of moving from high-need to lower-need schools vary between teachers who received Noyce support and other STEM teachers who graduated from the same universities or other STEM teachers in the same schools?

R4. How did the student demographics in schools of teachers who received Noyce support compare with student demographics in schools employing other STEM teachers from the same universities?

R5. How did student test score changes in Algebra I and Biology vary between teachers who received Noyce support and other teachers in the same high schools?

---

[2] The institutions in the research collaborative were The University of Houston, Stephen F. Austin State University, The University of Houston-Clear Lake, The University of Houston-Downtown, The University of Texas at Austin, The University of Texas at Arlington, Texas State University, and Texas A&M University-Kingsville. The first five of these contributed data to this study.

R6. How did student test score changes in Algebra I and Biology vary between teachers who received Noyce support and other teachers who graduated from the same universities?

Before presenting our results, we describe the theoretical framework for one of our main technical tools, value-added modelling.

The Complexity of Measuring the Value of Teaching from Student Outcomes

Particularly since the 2001 passage of the No Child Left Behind Act, increasing attention has been placed on the connections between teaching and K-12 student learning outcomes. The reliability of these methods is a legitimate subject of debate, in part because of their analytical complexity, in part because of the difficulty of obtaining necessary data, and in part because of disagreements over whether to trust causal conclusions from anything but an experiment with randomized controls (Campbell & Stanley, 1963; Imbens & Rubin, 2015; Pearl, 2009).

Value-added models analyze changes over time in student test scores, controlling for many demographic, classroom, and school variables. Using these variables, one computes an expectation for the score changes of a classroom of students. When the students in a classroom have scores that deviate from these expectations by a significant amount, one attributes the deviation to the qualities of the teacher. These methods were introduced by Hanushek (1971) and Sanders & Rivers (1996), and have been under continuous development ever since (Koedel, Mihaly, et al., 2015).

Many studies, particularly in the last decade, have begun to offer robust findings that inform the continuing debate (Chetty et al., 2011; Metzler & Woessmann, 2012). Some studies (Harris & Sass, 2011; Kane et al., 2008; Rivkin et al., 2005) suggest that traditional indicators of teacher quality (e.g., certification status, content preparation) account for very little of the variation in student outcomes. In contrast, other studies (Backes et al., 2018; Chetty et al., 2014b; Hanushek, 2011) find substantial associated student-level outcomes (e.g., increased likelihood to attend college; increased academic performance). These reflect questions of measurement of teacher quality and the relation of these indicators to student outcomes that are not fully answered, suggesting the urgent need for additional rigorous work to inform the conversation.

Value-added methods are controversial when they are used to make decisions about tenure or merit of individual teachers (Amrein-Beardsley, 2009; Darling-Hammond et al., 2012; Hill et al., 2011). In particular, as Pivovarova et al. (2016) note, there is a "number of issues with VAMs that should be taken into account when using VAMs in educational context, especially for high-stakes decision-making purposes (e.g., teacher merit pay, tenure, termination) [pp. 1-2]". Here we employ value-added scores to study categories of teachers rather than individuals. This means that uncertainties are reduced, and the possibility of negative consequences for individuals is removed.

Value-added models have previously been employed in studies in Texas. Two first studies of this sort compared student learning as a function of preparation program of the teacher and concluded that the value-added measures were so noisy no differences between programs could reliably be detected (Mellor et al., 2008; von Hippel et al., 2016). A third study focused on graduates of the UTeach program and found positive learning gains both in math and science for their students in comparison with peers whose teachers came from other programs (Backes et al., 2018). Another study (Marder et al., 2020) found that value-added student gains in Algebra I, and to a lesser extent in Biology, were higher for teachers from standard programs than they were for alternatively certified teachers.

Data Sample[3]

We examined outcomes of Texas Noyce programs making use of longitudinal administrative data from the Texas Education Research Center. These data made it possible for us to examine student learning gains in Algebra I and Biology, teacher retention, and propensity to teach in high-need schools. Student identifiers for these Noyce recipients were transmitted in secure fashion to the Texas Education Agency, which merged the identifiers to link the Noyce recipients to certification information, school assignments, student membership in their classes, and other information. The Noyce Scholars are a small subset of the overall teacher population from our universities. Thus, over the 15-year period of interest and eight funded universities, we began with a limited population of 248 recipients of Noyce Scholarships. From these, the Texas Education Agency successfully found a match on 202 of the 248. Of the Noyce recipients found in the dataset, however, only a subset taught in Texas schools, and a smaller subset taught tested high school subjects (Biology and Algebra I) during the time period 2012-2018 where student-teacher links were available in the state data. Table 1 describes the funding history and number of individuals in our sample from Noyce programs at eight universities. These universities received Noyce grants at many different times, with the earliest grant in 2003.

Table 1

Funding history of Scholarship and Stipend or Master Teacher Noyce grants and number of Noyce Scholars contributed to sample from each University in Collaborative Research project. Noyce Recipients labelled "-" indicate that data were unavailable, not that there were no graduates

Year of Start Date	Univ. of Houston	Univ. of Texas at Arlington	Stephen F. Austin State Univ.	Univ. of Houston - Downtown	Univ. of Texas at Austin	Texas State Univ. - San Marcos	Texas A&M Univ. - Kingsville	Univ. of Houston - Clear Lake
2003					$500K			
2006					$400K			
2008	$900K							
2009		$899K	$1.50M	$889K				
2010		$1.45M						
2011			$1.45M	$1.10M				
2012	$985K							
2013						$350K		$1.12M
2014							$1.20M	
2015		$960K		$1.45M				
2016	$1.74M	$300K	$1.05M		$800K	$1.26M		
2018	$2.83M	$1.45M						
2020	$1.00M		$1.01M					
Grand Total	**$7.45M**	**$5.06M**	**$5.00M**	**$3.41M**	**$1.70M**	**$1.60M**	**$1.20M**	**$1.12M**
Noyce Recipients in Sample	66	-	20	48	106	-	-	6

For UT Austin, which contributed the largest number of teachers to the study, we are able to provide an overview of the grants. A total of 127 preservice teachers were funded for whom outcomes are known (several more are still in the midst of service obligations). Of these, 92 met their teaching obligations, 20 completed the program but did not complete their teaching obligation in high-need schools and 15 left the program without completing certification. From Table 1 we see that $106-92 = 14$ of those who have not completed the obligation to teach in high-need schools have taught in Texas and appear in the TEA dataset.

Retention in Teaching of Texas Noyce Scholars

We now address the first three of our research questions on how long Noyce Scholars in our sample remained in teaching, and whether they moved away from high-need schools. We compare Noyce Scholars with two groups we use throughout this study: other STEM teachers at the same school, and other STEM teachers who came from the same universities. To form these comparisons, we found all the schools in which the Noyce Scholars taught, and then all the teachers who taught between 2003 and 2020 at the same middle and high schools with first teaching certificate in a STEM field. Similarly, we found all teachers whose first certificate was from the five universities in Table 1 contributing data, who taught in Texas between 2003 and 2020, and whose first teaching certificate was in a STEM field.

Research Questions R1, R2 & R3

To measure retention in teaching we look at the percentage of teachers in each group who taught for at least five years in a seven-year window. Teachers for whom it is too soon to tell are removed from consideration. For example, a teacher who first taught in 2015-2016 and taught four of the years through 2020-2021 would be removed from consideration because it would still be possible for them to complete a fifth year of teaching in 2021-2022.

We see in Table 2 that 63% of the Noyce Scholars completed five years of teaching by this definition. However, 74% of the other STEM teachers in their schools completed five years of teaching, as did 77% of the other STEM teachers prepared by their universities. Thus, the Noyce Scholars leave teaching at a higher rate by the five-year mark than the other groups.

Table 2

Five-year retention in teaching and mobility from high-need schools, comparing Noyce Scholars with other STEM teachers at same schools and other STEM teachers from same universities (R1, R2 & R3)

	Noyce Scholars	Other STEM teachers at same schools	Other STEM teachers from same universities
Total number taught in Texas	169	9812	3058
Taught 5 years (in first 7 years)	95	6546	2071
Did not teach 5 years (in first 7 years)	56	2307	633
Percent taught 5 years (in first 7 years)	63%	74%	77%
Average school poverty concentration	54%	53%	51%
Years switchers taught in first school	2.94 ± 0.24	3.08 ± 0.03	3.40 ± 0.06
Poverty concentration change on switch	$-6.4\% \pm 3.1\%$	$-0.2\% \pm 0.3\%$	$0.1\% \pm 0.5\%$

It is also interesting to investigate if Noyce Scholars are likely to leave high-need schools for lower-need schools once their teaching obligation is complete. Table 2 provides information on teachers who changed schools. The last two rows concern teachers who taught in more than one school between 2003 and 2020. The Noyce Scholars who switched stayed in the first school a shorter time (2.94 years) than the other teachers from their same universities (3.40 years). The schools to which they switched had lower poverty concentration (an average drop of 6.4%), while the schools to which teachers from the other groups switched did not. The differences between the Noyce Scholars and the other two groups are significant ($p<0.05$).

Noyce Scholars Teaching Tested Subjects

We next turn to a subset of the Noyce Scholars whose class assignments were in tested high school subjects and address research questions R4−R6. We were limited to those recipients who taught the test subjects of Algebra I and Biology during the academic years 2011/12−2017/18 for which we could match them with students. During this period, 24 Noyce teachers taught Algebra I in 26 campuses. We compared these teachers with 549 other Algebra I teachers in the same schools and with 594 other teachers who came from the same postsecondary institutions teaching at 416 other campuses. The comparison teachers teaching Algebra I in the same schools were restricted to teachers who obtained their first certification in Texas and were certified within the last 20 years.

For Biology, 17 Noyce teachers taught biology in periods where we could link them to classrooms and students in 25 campuses. We compared these biology Noyce scholars with 494 other biology teachers in the same schools and with 225 other teachers from the same postsecondary institutions teaching at 174 other campuses. The comparison teachers of biology in the same schools are also restricted to teachers who obtained their first certification in Texas and were certified within the last 20 years.

For this subset of Noyce Scholars, we investigated more deeply the students they were serving. The Noyce requirement of teaching in a high-need local education agency does not necessarily turn out as one might think. A district can have a large fraction of low-income students overall and therefore be categorized as a high-need LEA, but the Noyce recipient could end up teaching in a school in the district serving mostly non-low-income students. Or, in a school with a majority low-income population, a graduate could end up teaching classes where the students are non-low-income. For this reason, we examined the actual demographic characteristics of students in Noyce Scholars' classes.

Research Question R4

In Table 3 we provide information on the demographics of students taught by the subset of Noyce Scholars in the tested high school subjects Algebra I and Biology for the academic years 2011/12 through 2017/18. This is the group we will examine next with value-added models, and it is helpful to have the demographic background for comparison.

There are nine columns in the table. For each student population, the third column provides the percentage of students in the Noyce Scholars' classes that belonged to that group. The fourth column displays the difference in percentage compared with the classes of non-Noyce teachers from the same universities. For example, the Algebra I classes of Noyce Scholars had 54.5% economically disadvantaged students (column 3). This is 13.3% more than the 41.2% of economically disadvantaged students in the classes of the non-Noyce teachers (column 4).

Overall, we see in Table 3 that in Algebra I, Noyce Scholars compared with other teachers graduating from the same universities taught a higher percentage of economically disadvantaged, gifted, and Hispanic students, but a lower percentage of Special Education, Black, and White

students. In Biology the situation is the same, except there is also a larger percentage of Black students in the Noyce Scholars' classrooms.

Table 3
Student demographics of Noyce Scholar classrooms compared with classrooms of other teachers from the same university (R4). Numbers in parentheses are one standard uncertainty. Final column from Eq. (1)

Student Category	Discipline	Noyce classrooms	Difference between Noyce classrooms and those of other teachers from same university			Difference between Noyce classrooms and those of other teachers from same university, controlling for school demographics		
Eco Dis	Algebra I	54.5%	13.3%	(0.6%)	***	−0.2%	(0.7%)	
Gifted	Algebra I	3.4%	3.1%	(0.2%)	***	3.0%	(0.3%)	***
SpecEd	Algebra I	5.0%	−3.3%	(0.2%)	***	−3.5%	(0.3%)	***
LEP	Algebra I	8.8%	−0.2%	(0.3%)		−5.0%	(0.5%)	***
Asian	Algebra I	2.9%	−0.3%	(0.2%)		1.0%	(0.3%)	***
Black	Algebra I	17.7%	−6.0%	(0.4%)	***	−1.0%	(0.6%)	
Hispanic	Algebra I	49.6%	22.7%	(0.6%)	***	−0.3%	(0.7%)	
White	Algebra I	27.6%	−15.4%	(0.5%)	***	0.7%	(0.6%)	
EcoDis	Biology	44.9%	8.1%	(0.6%)	***	2.8%	(0.7%)	***
Gifted	Biology	13.0%	1.1%	(0.4%)	**	−3.6%	(0.5%)	***
SpecEd	Biology	4.0%	−0.8%	(0.2%)	***	−0.4%	(0.3%)	
LEP	Biology	6.5%	−0.1%	(0.3%)		−1.0%	(0.4%)	**
Asian	Biology	6.3%	−1.7%	(0.3%)	***	0.2%	(0.4%)	
Black	Biology	12.8%	3.4%	(0.4%)	***	2.5%	(0.5%)	***
Hispanic	Biology	48.3%	8.1%	(0.6%)	***	1.7%	(0.7%)	**
White	Biology	30.0%	−9.4%	(0.5%)	***	−3.7%	(0.7%)	***

* p<.05, ** p<.01, *** p<.001

Are these differences due to how the Noyce Scholars were assigned within their schools or due to the schools in which they taught? Inspection of the final columns in Table 3 shows that the differences are *between* schools, not *within* schools. We checked this by computing hierarchical linear models (Bolker et al., 2016) of the form

$$X_i = C_j + \text{Noyce} + \epsilon_{ij}, \tag{1}$$

where X_i is a demographic characteristic of student i, C_j their campus, Noyce is a flag for whether the student's teacher is a Noyce scholar, and ϵ_{ij} is an error term. We modelled the campus contribution C_j as a random effect in a hierarchical linear model.

This model controls for the demographics of all the Algebra I or Biology classes in each school before comparing the classrooms of the teachers. For example, while the Algebra I classrooms of Noyce Scholars have 13% more economically disadvantaged students than the Algebra I classrooms of other teachers from the same universities, after controlling for school demographics, the Noyce Scholars have 0.2% fewer economically disadvantaged and 0.3% fewer Hispanic students in Algebra I. These differences are not statistically significant. That is, the Noyce Scholars have more economically disadvantaged students in their classrooms because they are teaching in schools with higher poverty concentration, not because they are preferentially picking up

these students within the schools. Some significant differences for economically disadvantaged and Hispanic students remain for the biology teachers, but they are still only a few percentages. A thorough examination of Table 3 finds some interesting significant differences in the final column. Noyce Scholars are less likely to teach White and Gifted students in their schools, even after controlling for school population. Overall, however, the effect sizes are much smaller once one accounts for school demographics. Therefore, it appears the Noyce programs in Texas are achieving their goal of incentivizing recipients to teach in schools with a high percentage of economically disadvantaged and marginalized students, particularly Hispanic students. Within these schools serving marginalized students, the Noyce scholars are as likely as other teachers to teach them.

For our value-added models we are not only comparing teachers who came from the same universities as the Noyce Scholars, but also teachers who taught the same subjects at the same schools but were not Noyce Scholars. For reference, we report on the demographic characteristics for the second comparison group in Table 4. Some of the same patterns repeat, with Noyce Scholars a bit more likely to teach economically disadvantaged students than other teachers in the same school, even after controlling for school demographics using Equation (1). However, the effects smaller, and they do not have as much direct significance for the Noyce program as the results in the previous table.

Table 4

Student demographics of Noyce Scholar classrooms compared with classrooms of other teachers who taught in the same schools (R4). Numbers in parentheses are one standard uncertainty. Final column from Eq. (1)

Student Category	Discipline	Noyce	Difference between Noyce classrooms and those of other teachers from same school			Difference between Noyce classrooms and those of other teachers from same school, controlling for school demographics		
EcoDis	Algebra I	48.8%	4.2%	(0.6%)	***	1.3%	(0.6%)	*
Gifted	Algebra I	10.6%	3.4%	(0.4%)	***	−2.0%	(0.4%)	***
SpecEd	Algebra I	4.3%	−1.1%	(0.2%)	***	−1.0%	(0.2%)	***
LEP	Algebra I	5.1%	1.2%	(0.3%)	***	0.4%	(0.3%)	
Asian	Algebra I	5.8%	−1.2%	(0.3%)	***	−0.4%	(0.3%)	
Black	Algebra I	15.4%	0.7%	(0.4%)		1.6%	(0.4%)	***
Hispanic	Algebra I	54.1%	2.4%	(0.6%)	***	1.8%	(0.6%)	***
White	Algebra I	22.3%	−1.7%	(0.5%)	***	−2.5%	(0.5%)	***
EcoDis	Biology	49.3%	5.2%	(0.7%)	***	3.6%	(0.6%)	***
Gifted	Biology	10.3%	2.9%	(0.4%)	***	−3.3%	(0.4%)	***
SpecEd	Biology	3.3%	−0.1%	(0.2%)	***	−0.5%	(0.2%)	***
LEP	Biology	4.0%	1.9%	(0.3%)	***	1.0%	(0.3%)	**
Asian	Biology	6.3%	−2.5%	(0.3%)	***	−0.8%	(0.3%)	***
Black	Biology	16.7%	1.0%	(0.5%)	*	1.8%	(0.4%)	***
Hispanic	Biology	52.7%	4.6%	(0.7%)	***	3.5%	(0.6%)	***
White	Biology	21.8%	−2.6%	(0.6%)	***	−3.7%	(0.5%)	***

* p<.05, ** p<.01, *** p<.001

Methods: Student Learning in Classrooms of Texas Noyce Graduates

Now we turn to the question of student learning in the classrooms of Noyce Scholars. We computed value-added estimates using lmer in R (Bolker et al., 2016), as shown in Equation (2). At

the top level, A_{it} is the score of student i in some year and $A_{i,t-1}$ is the same student's score on the exam in the same subject the previous year $t-1$ in a cubic polynomial. When inserted into the model, each score is always normalized by the maximum score, so it lies in the interval [0,1]. We use normalized raw score rather than a state-supplied scaled score for reasons described in (Bendinelli & Marder, 2012). Teacher j of student i contributes through the random intercept T_j. By modeling the teacher in this way, each teacher should contribute equally to the estimate of the effect of their pathway to teaching (Koedel, Parsons et al., 2015). The campus k contributes random intercept C_k as does the class n through Class_n. Coefficients for student-level demographic factors X_i range over Gifted, racial and ethnic groups, Limited English Proficiency (LEP), Free/Reduced Lunch Eligibility (EcoDis), and Special Education. Classroom characteristics Z_n are classroom averages of these same quantities (Chetty et al., 2014a; J. N. Friedman et al., 2014).

We modeled the influence of tracking, as recommended by Jackson (2014). The most important form of student tracking in Texas is placing students in Algebra I in eighth-grade. To control for this, we removed students enrolled in Algebra 1 during their eighth-grade year from our study of mathematics and created a flag for them when modeling biology. We include a variable E to control for teacher years of experience, assuming binned values of 0-4, 4-10, and 10-20 years of experience.

We employed some additional variables with information about the teachers. A variable StdCert marks standard as opposed to alternative teacher certification. We had available the undergraduate Grade Point Average (GPA) of the teachers as well as their major.

The main item of interest, certification pathway Noyce for teacher j of student i, enters as a fixed effect. The computation includes data for years t from 2012 through 2018. Finally, the second level of the model has random intercepts for teacher T, campus C, and class section Class.

$$A_{it} = \sum_{\alpha=1}^{3} \lambda_{\alpha t} A_{i,t-1}^{\alpha} + T_{jt} + \text{Noyce} + \text{Class}_{nt} + \beta_1 X_{it} + \epsilon_{ijnt} \tag{2a}$$

$$A_{it} = \sum_{\alpha=1}^{3} \lambda_{\alpha t} A_{i,t-1}^{\alpha} + T_{jt} + \text{StdCert} + \text{Noyce} + C_{kt} + \text{Class}_{nt} + \beta_1 X_{it} + \epsilon_{ijknt} \tag{2b}$$

$$A_{it} = \sum_{\alpha=1}^{3} \lambda_{\alpha t} A_{i,t-1}^{\alpha} + T_{jt} + \text{StdCert} + \text{Noyce} + E_{jt} + C_{kt} + \text{Class}_{nt} + \beta_1 X_{it} + \epsilon_{ijknt} \tag{2c}$$

$$A_{it} = \sum_{\alpha=1}^{3} \lambda_{\alpha t} A_{i,t-1}^{\alpha} + T_{jt} + \text{Noyce} + E_{jt} + C_{kt} + \text{Class}_{nt} + \beta_1 X_{it} + \beta_2 Z_{nt} + \epsilon_{ijknt} \tag{2d}$$

We computed the estimates while progressively adding and subtracting terms described above. This produced 4 models we describe in Equations 2a – 2d. We provide detailed information from each these four models in Tables 5 through 8. The four Tables correspond to the four combinations produced by the two disciplines, Algebra I and Biology, and the two comparison groups, other teachers teaching the same course in the same school, and other teachers teaching the same course from the same university. Tables 9 and 10 contain again a comparison of Noyce Scholars teaching Algebra I with other Algebra I teachers in the same school; however, the GPA of the teachers is added as a factor to the model.

Table 5

Algebra I, compare to teachers in same schools (R5). Standard deviation units. Standard uncertainty in parentheses. We compare 24 Noyce teachers with 549 others at 26 campuses in 6898 classes with 132138 total students

Random Effects	2a			2b			2c			2d		
Campus SD				0.22			0.22			0.20		
Teacher SD	0.27			0.20			0.20			0.20		
Class SD	0.17			0.17			0.17			0.17		
Fixed Effects												
Noyce	0.17	(0.06)	**	0.10	(0.05)	*	0.10	(0.05)	*	0.11	(0.05)	*
StdCert				0.04	(0.02)		0.03	(0.02)				
EcoDis	−0.05	(0.00)	***	−0.05	(0.00)	***	−0.05	(0.00)	***	−0.05	(0.00)	***
Gifted	0.22	(0.01)	***	0.22	(0.01)	***	0.22	(0.01)	***	0.21	(0.01)	***
SpecEd	−0.28	(0.01)	***	−0.29	(0.01)	***	−0.29	(0.01)	***	−0.25	(0.01)	***
LEP	−0.18	(0.01)	***	−0.18	(0.01)	***	−0.18	(0.01)	***	−0.17	(0.01)	***
Asian	0.20	(0.02)	***	0.19	(0.02)	***	0.19	(0.02)	***	0.18	(0.02)	***
Black	−0.08	(0.02)	***	−0.08	(0.02)	***	−0.08	(0.02)	***	−0.07	(0.02)	***
Hispanic	−0.05	(0.02)	**	−0.05	(0.02)	**	−0.05	(0.02)	**	−0.05	(0.02)	**
White	−0.03	(0.02)		−0.03	(0.02)		−0.03	(0.02)		−0.02	(0.02)	
AveEcoDis										−0.01	(0.03)	
AveGifted										0.26	(0.06)	***
AveSpecEd										−0.11	(0.02)	***
AveLEP										−0.05	(0.02)	**
AveAsian										0.40	(0.09)	***
AveBlack										−0.05	(0.05)	
AveHispanic										−0.01	(0.04)	
AveWhite												
Tracked												
Yrs: 4-10							0.03	(0.01)	*	0.03	(0.01)	*
Yrs: 11-20							0.04	(0.02)				
$\lambda_1(2012)$	2.19	(0.34)	***	2.20	(0.34)	***	2.23	(0.34)	***	2.15	(0.34)	***
$\lambda_1(2013)$	5.04	(0.35)	***	5.06	(0.35)	***	5.09	(0.35)	***	4.99	(0.35)	***
$\lambda_1(2014)$	3.34	(0.33)	***	3.34	(0.33)	***	3.35	(0.33)	***	3.27	(0.33)	***
$\lambda_1(2015)$	2.67	(0.33)	***	2.67	(0.33)	***	2.67	(0.33)	***	2.62	(0.33)	***
$\lambda_1(2016)$	3.74	(0.35)	***	3.74	(0.35)	***	3.72	(0.35)	***	3.67	(0.35)	***
$\lambda_1(2017)$	3.12	(0.33)	***	3.10	(0.33)	***	3.08	(0.33)	***	3.08	(0.33)	***
$\lambda_1(2018)$	3.19	(0.32)	***	3.15	(0.32)	***	3.12	(0.32)	***	3.17	(0.32)	***
$\lambda_2(2012)$	−1.36	(0.72)		−1.41	(0.72)	*	−1.45	(0.72)	*	−1.32	(0.72)	
$\lambda_2(2013)$	−5.09	(0.74)	***	−5.13	(0.74)	***	−5.19	(0.74)	***	−5.01	(0.74)	***
$\lambda_2(2014)$	0.40	(0.70)		0.40	(0.70)		0.39	(0.70)		0.50	(0.70)	
$\lambda_2(2015)$	2.02	(0.68)	**	2.03	(0.68)	**	2.02	(0.68)	**	2.07	(0.68)	**
$\lambda_2(2016)$	−0.14	(0.72)		−0.12	(0.72)		−0.09	(0.72)		−0.04	(0.72)	
$\lambda_2(2017)$	2.35	(0.70)	***	2.37	(0.70)	***	2.42	(0.70)	***	2.35	(0.70)	***
$\lambda_2(2018)$	1.17	(0.65)		1.24	(0.65)		1.28	(0.65)	*	1.10	(0.65)	
$\lambda_3(2012)$	2.42	(0.44)	***	2.45	(0.44)	***	2.46	(0.44)	***	2.40	(0.44)	***
$\lambda_3(2013)$	4.17	(0.49)	***	4.19	(0.49)	***	4.22	(0.49)	***	4.12	(0.49)	***
$\lambda_3(2014)$	0.13	(0.47)		0.14	(0.47)		0.14	(0.47)		0.09	(0.47)	
$\lambda_3(2015)$	−0.78	(0.44)		−0.78	(0.44)		−0.78	(0.44)		−0.80	(0.44)	
$\lambda_3(2016)$	0.73	(0.47)		0.71	(0.47)		0.69	(0.47)		0.66	(0.47)	
$\lambda_3(2017)$	−1.40	(0.46)	**	−1.42	(0.46)	***	−1.45	(0.46)	***	−1.39	(0.46)	**
$\lambda_3(2018)$	−0.43	(0.40)		−0.47	(0.40)		−0.50	(0.40)		−0.38	(0.40)	

* p<.05, ** p<.01, *** p<.001

For Table 5, we report all the model coefficients, including all the parameters for the expected value for student test scores in year t as a function of scores in year $t-1$. These coefficients are extremely important for the modeling but not very informative to read, and therefore we refrain from including them in the subsequent tables. In Table 5, and all following tables, the columns 2a through 2d correspond to Equations 2a through 2d above.

In all the tables, we report the coefficients in standard deviation units. To communicate these results more effectively, it is useful to observe that in the course of a nine-month school year, students typically improve by a quarter of a standard deviation in high school math or science (Lipsey et al., 2012). Thus an improvement of 3% in standard deviation units can also be expressed as one more month of schooling in a year. For Algebra I and Biology, a standard deviation is close to 16% of the maximum score in all cases.

Findings

Research Question R5

Table 6
Algebra I, compared to teachers from same universities (R6). Standard deviation units. Standard uncertainty in parentheses. We compare 24 Noyce teachers with 594 others at 416 campuses in 2367 classes with 145680 total students

Random Effects	2a			2b			2c			2d		
Campus SD				0.24			0.24			0.23		
Teacher SD	0.30			0.22			0.22			0.22		
Class SD	0.19			0.18			0.18			0.17		
Fixed Effects												
Noyce	0.17	(0.07)	**	0.11	(0.06)		0.11	(0.06)		0.10	(0.06)	
StdCert				0.07	(0.05)		0.07	(0.05)				
EcoDis	−0.07	(0.00)	***	−0.07	(0.00)	***	−0.07	(0.00)	***	−0.06	(0.00)	***
Gifted	0.20	(0.01)	***	0.20	(0.01)	***	0.20	(0.01)	***	0.18	(0.01)	***
SpecEd	−0.29	(0.01)	***	−0.30	(0.01)	***	−0.30	(0.01)	***	−0.26	(0.01)	***
LEP	−0.17	(0.01)	***	−0.17	(0.01)	***	−0.17	(0.01)	***	−0.17	(0.01)	***
Asian	0.20	(0.02)	***	0.19	(0.02)	***	0.19	(0.02)	***	0.18	(0.02)	***
Black	−0.03	(0.01)	**	−0.03	(0.01)	*	−0.03	(0.01)	*	−0.03	(0.01)	*
Hispanic	0.00	(0.01)		0.00	(0.01)		0.00	(0.01)		0.00	(0.01)	
White	0.01	(0.01)		0.01	(0.01)		0.01	(0.01)		0.02	(0.01)	
AveEcoDis										−0.11	(0.03)	***
AveGifted										0.45	(0.05)	***
AveSpecEd										−0.14	(0.02)	***
AveLEP										−0.01	(0.02)	
AveAsian										0.45	(0.07)	***
AveBlack										0.01	(0.04)	
AveHispanic										0.02	(0.03)	
AveWhite												
Tracked												
Yrs: 4-10							0.00	(0.02)		0.00	(0.02)	
Yrs: 11-20							0.00	(0.03)		0.00	(0.03)	

* p<.05, ** p<.01, *** p<.001

For each model, the coefficient of ultimate interest is Noyce, which is a flag indicating whether the teacher was a Noyce Scholar or not. In Algebra I, students of Noyce Scholars gained 10% to 17% (3.5 to 6.1 months of schooling) more than students of other teachers in the same schools (Table 5). This result is robust and persists across every model we examined, except ones including GPA which we will discuss below. The largest estimate, 17% (6 months of schooling) gained per year in Algebra I, shows up for models where the Campus term is missing. This means that these models do not correct for the overall level of student accomplishment in each school. This choice of model is appropriate if one believes that the main reason some schools are better is mainly due to their teachers, so if students in a school are doing particularly well overall, one should not subtract this from all the teachers' scores.

Table 7

Biology, compared to teachers in same schools (R5). Standard deviation units. Standard uncertainty in parentheses. We compare 17 Noyce teachers with 494 others at 25 campuses in 7564 classes with 158079 total students

Random Effects	2a			2b			2c			2d		
Campus SD				0.20			0.20			0.16		
Teacher SD	0.25			0.19			0.19			0.16		
Class SD	0.16			0.15			0.15			0.13		
Fixed Effects												
Noyce	0.05	(0.07)		0.08	(0.05)		0.08	(0.05)		0.08	(0.05)	
StdCert				0.01	(0.02)		0.01	(0.02)				
EcoDis	−0.08	(0.00)	***	−0.08	(0.00)	***	−0.08	(0.00)	***	−0.07	(0.00)	***
Gifted	0.23	(0.01)	***	0.23	(0.01)	***	0.23	(0.01)	***	0.19	(0.01)	***
SpecEd	−0.30	(0.01)	***	−0.30	(0.01)	***	−0.30	(0.01)	***	−0.22	(0.01)	***
LEP	−0.28	(0.01)	***	−0.28	(0.01)	***	−0.28	(0.01)	***	−0.26	(0.01)	***
Asian	0.09	(0.01)	***	0.09	(0.01)	***	0.09	(0.01)	***	0.08	(0.01)	***
Black	−0.08	(0.01)	***	−0.08	(0.01)	***	−0.08	(0.01)	***	−0.07	(0.01)	***
Hispanic	−0.08	(0.01)	***	−0.08	(0.01)	***	−0.08	(0.01)	***	−0.07	(0.01)	***
White	0.03	(0.01)	**	0.03	(0.01)	**	0.03	(0.01)	**	0.03	(0.01)	**
AveEcoDis										−0.19	(0.02)	***
AveGifted										0.21	(0.03)	***
AveSpecEd										−0.24	(0.02)	***
AveLEP										−0.1	(0.02)	***
AveAsian										0.22	(0.05)	***
AveBlack										−0.22	(0.03)	***
AveHispanic										−0.18	(0.03)	***
AveWhite												
Tracked										0.14	(0.00)	***
Yrs: 4-10							0.02	(0.01)		0.02	(0.01)	
Yrs: 11-20							0.06	(0.02)	***	0.03	(0.02)	*

* p<.05, ** p<.01, *** p<.001

Research Question R6

When comparing Algebra I teachers with other Algebra I teachers who came from the same universities, the only model that estimates students of Noyce scholars to have significantly greater learning is the one without the Campus intercept. In all other cases the uncertainty is too large to draw a conclusion. In the case of Biology, for all models the uncertainties are too great to draw conclusions about whether the students of Noyce Scholars got higher scores than their peers. This is

true whether the comparison group is students of other teachers in the same schools, or students of other graduates of the same universities.

Table 8

Biology, compare to teachers from same universities (R6). Standard deviation units. Standard uncertainty in parentheses. We compare 17 Noyce teachers with 225 others at 174 campuses in 4285 classes with 91021 total students

Random Effects	2a			2b			2c			2d		
Campus SD				0.23			0.23			0.17		
Teacher SD	0.27			0.19			0.19			0.18		
Class SD	0.17			0.16			0.16			0.14		
Fixed Effects												
Noyce	0.02	(0.07)		0.02	(0.06)		0.02	(0.06)		0.02	(0.06)	
StdCert				0.16	(0.08)		0.17	(0.08)	*			
EcoDis	−0.08	(0.01)	***	−0.08	(0.01)	***	−0.08	(0.01)	***	−0.07	(0.01)	***
Gifted	0.20	(0.01)	***	0.20	(0.01)	***	0.20	(0.01)	***	0.17	(0.01)	***
SpecEd	−0.27	(0.01)	***	−0.28	(0.01)	***	−0.28	(0.01)	***	−0.20	(0.01)	***
LEP	−0.28	(0.01)	***	−0.28	(0.01)	***	−0.28	(0.01)	***	−0.25	(0.01)	***
Asian	0.06	(0.02)	***	0.06	(0.02)	***	0.06	(0.02)	***	0.05	(0.02)	***
Black	−0.10	(0.01)	***	−0.09	(0.01)	***	−0.09	(0.01)	***	−0.08	(0.01)	***
Hispanic	−0.10	(0.01)	***	−0.09	(0.01)	***	−0.09	(0.01)	***	−0.09	(0.01)	***
White	−0.01	(0.01)		−0.01	(0.01)		−0.01	(0.01)		−0.01	(0.01)	
AveEcoDis										−0.18	(0.03)	***
AveGifted										0.32	(0.03)	***
AveSpecEd										−0.30	(0.03)	***
AveLEP										−0.10	(0.03)	***
AveAsian										−0.02	(0.06)	
AveBlack										−0.25	(0.05)	***
AveHispanic										−0.19	(0.04)	***
AveWhite												
Tracked										0.13	(0.01)	***
Yrs: 4-10							−0.02	(0.02)		−0.04	(0.02)	*
Yrs: 11-20							0.01	(0.03)		−0.05	(0.03)	

*p<.05, ** p<.01, *** p<.001*

We make selected comments about other random and fixed effects in the models. First, we make general remarks that apply to all models and both disciplines (Tables 5 – 8). The Campus random effect is 16% to 24% (5.7 to 8.7 months of schooling). This is the scale of variation from one campus to another, even after taking into account detailed demographic information about all students in the campus. The Teacher random effect is even bigger, 18% to 30% (5.7 to 10.7 months of schooling). Thus, student scores vary almost exactly as much from teacher to teacher as they do school to school. Only slightly smaller is the Class random effect at 14% to 19% (4.7 to 6.9 months of schooling). This means that when teachers teach more than one class section, either during one year, or over many years, the variation from one section to another is almost as big as the typical variation from one school to another or one teacher to another. This is typical of the high levels of variation that render value-added models unsuitable for judging individual teachers(von Hippel et al., 2016; von Hippel & Bellows, 2018).

Most of the fixed effects are highly significant. The main exception is Standard Certification. This factor is significant in models without Noyce (Marder et al., 2020) but since all Noyce Scholars have standard certification, and other teachers in the same schools so frequently come from alternative certification programs (247/549 for math, 224/494 for science), Standard Certification as a factor has nothing to add to any of the models.

All of the other fixed effects are strongly correlated with student outcomes. It is particularly striking that classroom average effects are larger than individual student effects. For example, in Table 7 for Biology, when a student is economically disadvantaged, they gain −8% (2.8 fewer months of schooling) than other students, but in a class that is 100% economically disadvantaged the students gain −19% (6.7 months less schooling) than a class where none of the students is economically disadvantaged.   There are large positive effects for Gifted and Asian students and large negative effects for Economically Disadvantaged, Limited English Proficiency, Special Ed, Hispanic, and African American students. In all models the score of each student is adjusted according to their group assignments, and in model 2d there is additional adjustment for classroom averages. Model 2d also adjusts for the teacher's years of experience, and in Biology for a flag that indicates whether a student has been tracked into Algebra I in eighth-grade. The eighth-grade Algebra I students are excluded from the Algebra I models altogether.

Table 9
Algebra I, compare with other teachers in the same schools (R5); include GPA. Units are standard deviation. Standard uncertainty in parentheses. Because of the inclusion of GPA, the sample size drops from 573 to 96 teachers and this is the main reason for loss of significance

Random Effects	2a			2b			2c			2d		
Campus SD				0.22			0.23			0.20		
Teacher SD	0.26			0.18			0.19			0.19		
Class SD	0.16			0.16			0.16			0.16		
Fixed Effects												
Noyce	0.06	(0.09)		0.03	(0.07)		0.03	(0.07)		0.03	(0.07)	
StdCert				0.01	(0.06)		0.00	(0.06)				
GPA	0.13	(0.07)		0.14	(0.06)	**	0.14	(0.06)	**	0.14	(0.06)	**
EcoDis	−0.03	(0.01)	**	−0.03	(0.01)	**	−0.03	(0.01)	**	−0.03	(0.01)	**
Gifted	0.15	(0.03)	***	0.14	(0.03)	***	0.14	(0.03)	***	0.12	(0.03)	***
SpecEd	−0.28	(0.02)	***	−0.28	(0.02)	***	−0.28	(0.02)	***	−0.26	(0.03)	***
LEP	−0.14	(0.02)	***	−0.14	(0.02)	***	−0.14	(0.02)	***	−0.14	(0.02)	***
Asian	0.06	(0.06)		0.05	(0.06)		0.05	(0.06)		0.04	(0.06)	
Black	−0.10	(0.04)	*	−0.10	(0.04)	*	−0.10	(0.04)	*	−0.10	(0.05)	*
Hispanic	−0.08	(0.04)		−0.08	(0.04)		−0.07	(0.04)		−0.07	(0.04)	
White	−0.01	(0.04)		−0.01	(0.04)		−0.01	(0.04)		−0.01	(0.05)	
AveEcoDis										0.00	(0.06)	
AveGifted										0.31	(0.13)	**
AveSpecEd										−0.07	(0.05)	
AveLEP										0.02	(0.04)	
AveAsian										0.40	(0.21)	
AveBlack										0.15	(0.12)	
AveHispanic										0.06	(0.10)	
Yrs: 4-10							0.07	(0.06)		0.06	(0.06)	
Yrs: 11-20							0.04	(0.06)		0.03	(0.06)	

* p<.05, ** p<.01, *** p<.001

We ran a series of models in which we included the undergraduate GPA of the teachers. These models have small sample sizes. Student transcripts are only available in the Texas dataset after 2012, and therefore the only teachers for which we can establish a GPA are a small number of recent graduates. Of the 573 teachers whose students provide data for Table 5, only 96 are left with GPA for Tables 9 and 10. Not only is the sample size greatly reduced, but one should remember that Noyce Scholars are chosen on the basis of academic qualifications such as GPA. This means that GPA is correlated with the probability of being a Noyce Scholar. Specifically, for both Algebra I and Biology teachers, each unit change in GPA leads to a factor of 1.2 increase in the odds ratio for being a Noyce scholar. Therefore, both GPA and Noyce have elevated levels of uncertainty because of the correlation. Once GPA is included, the flag for Noyce Scholars stops being significant. In Table 9, which compares with other teachers in the same schools, teacher GPA is a significant predictor of student success in most models, despite the reduced sample size, with students gaining 13% (4.8 months of schooling) for each increase in teacher GPA on a 4-point scale. However, in Table 10, which compares with other teachers prepared by the same universities, GPA is not a significant predictor.

Table 10
Algebra I, compare with other teachers from the same universities (R6); include GPA. Units are standard deviation. Standard uncertainty in parentheses. Because of the inclusion of GPA, the sample size drops from 573 to 96 teachers and this is the main reason for loss of significance

Random Effects	2a			2b			2c			2d		
Campus SD				0.23			0.22			0.21		
Teacher SD	0.27			0.2			0.2			0.2		
Class SD	0.2			0.18			0.18			0.18		
Fixed Effects												
Noyce	0.16	(0.08)	*	0.08	(0.08)		0.07	(0.08)		0.05	(0.08)	
StdCert				-0.12	(0.14)		-0.11	(0.14)				
GPA	0.07	(0.05)		0.09	(0.05)		0.09	(0.05)		0.08	(0.05)	
EcoDis	-0.03	(0.01)	***	-0.03	(0.01)	***	-0.03	(0.01)	***	-0.03	(0.01)	***
Gifted	0.19	(0.02)	***	0.19	(0.02)	***	0.19	(0.02)	***	0.17	(0.02)	***
SpecEd	-0.26	(0.02)	***	-0.26	(0.02)	***	-0.27	(0.02)	***	-0.2	(0.02)	***
LEP	-0.16	(0.01)	***	-0.16	(0.01)	***	-0.16	(0.01)	***	-0.16	(0.01)	***
Asian	0.17	(0.03)	***	0.16	(0.03)	***	0.16	(0.03)	***	0.16	(0.03)	***
Black	-0.07	(0.03)	**	-0.07	(0.03)	**	-0.07	(0.03)	**	-0.06	(0.03)	*
Hispanic	-0.03	(0.03)		-0.03	(0.03)		-0.03	(0.03)		-0.02	(0.03)	
White	-0.03	(0.03)		-0.03	(0.03)		-0.03	(0.03)		-0.02	(0.03)	
AveEcoDis										-0.08	(0.05)	
AveGifted										0.44	(0.10)	***
AveSpecEd										-0.22	(0.04)	***
AveLEP										-0.04	(0.04)	
AveAsian										0.44	(0.12)	***
AveBlack										0.17	(0.07)	**
AveHispanic										0.17	(0.06)	**
Yrs: 4-10												
Yrs: 11-20												

* p<.05, ** p<.01, *** p<.001

Limitations and Implications

Limitations

The Noyce program has been in operation for more than 15 years and has served more than 12,000 preservice teachers. Previous studies have either resulted from surveys of participants while they were still in school, or from examinations of individual programs over short periods of time. To the best of our knowledge, we are presenting the first study that aggregates data from multiple programs with multiple grants and examines a variety of quantitative outcomes.

In some respects, our sample is large. We had data from seven years, thousands of classrooms, and hundreds of thousands of students. However, in another respect our sample is small; for the value-added models we had 24 math teachers and 17 science teachers in the treatment group. These small numbers were somewhat counterbalanced by the fact that we had observations over multiple years. In the end we rely upon statistical tests to tell us if the sample was large enough to draw conclusions, and the tests tell us that results in Algebra I are fairly robust, while in Biology there is no evidence students are getting higher test scores in Noyce Scholar classrooms.

One might also be worried that over the course of 15 years there could have been major changes in the teacher preparation programs we studied that would have made it inappropriate to group together teachers coming out at different times. There were certainly major changes in school systems and statewide educational policy. However, our concern on this score is limited because the programs contributing teachers to this study did not make major changes to their curriculum or support structures during the study period, even taking into account the receipt of the multiple Noyce grants.

We acknowledge that studies with administrative data are less reliable than random-controlled trials for removing unobserved causal factors. However, when random-controlled trials have been conducted in similar circumstances to ours (Kane & Staiger, 2008; Turner et al., 2012) the randomization has been done by assigning students to classrooms within a school. Our descriptive data provide no suggestion that students were not assigned randomly within schools, but abundant evidence that teachers were not assigned randomly to schools. Therefore, the randomized experiment one would really need to do would be to randomly assign teachers to schools. Such an experiment has not been done and would probably be impossible.

Implications

The theory of action (Figure 1) behind the Noyce program is that scholarships can motivate strong STEM majors to become teachers, that financial penalties will motivate them to teach low-income and other marginalized groups of students, and that K-12 students will learn more as a result than they would with other teachers.

Liou, Kirchhoff, et al. (2010) established that the Noyce scholarships are of limited effectiveness in persuading students to teach. Around half the Noyce Scholars in their survey said the scholarships were not very influential in their decision to become a teacher or complete teaching certification. However, there were hints that the scholarships could affect where they taught. Nearly three quarters said the scholarships influenced them to teach in a high-need school and to stay for the term of the commitment, but only half said the scholarship would influence them to stay in a high-need school past the term of the commitment. Since these results were based on student intentions rather than actions they deserved to be followed up, which was the aim of this study.

We find that Noyce Scholars in Texas do indeed work in schools with a high percentage of marginalized students. The Noyce Scholars are more likely than other graduates from their universities to have low-income, Black, and Hispanic students in their classrooms, but no more or

less likely than other teachers in the school where they teach. Thus, the Noyce Scholars are preferentially seeking schools with marginalized populations, and at those schools they are teaching representative populations of students. This result should not be taken for granted. The Noyce grants require recipients to teach in high-need LEAs rather than in high-need schools. Thus, Noyce Scholars would be able to evade the spirit of the Noyce requirements by seeking out mostly non-high-need schools in mostly high-need districts. On average, however, the recipients do not and were found to be teaching a large fraction of low-income students.

Our evidence on student learning in the classrooms of Noyce Scholars is mixed. In mathematics (Algebra I) we have robust evidence that the students of Noyce Scholars learn more than they do from other teachers in the same school after controlling for a wide array of factors in many different models. When we compare the Noyce Scholars with other teachers coming from the same universities, some models say their students do better, while other models find no difference. In Biology, the value-added methods find no difference between student learning in classrooms of Noyce Scholars and classrooms of other teachers.

Typically, Noyce Scholars are selected for being strong students within their programs, and are chosen to motivate them through scholarship funds to teach in low-income schools. It is not clear that Noyce Scholarships are in fact going to the future teachers who will ultimately have the best records. If conventional ways of defining top students, such as GPA, were strong predictors of great teaching, then GPA should be an important factor in Table 10, where the teaching accomplishments of Noyce Scholars are compared with those of others who graduated from the same university programs. But GPA has no significance. It is true the sample size is small, but this did not prevent GPA from being significant in Table 9, where the Noyce Scholars were compared with other teachers in the same schools. This could be taken as an encouraging result, for if Noyce Scholarships were given to a larger and more diverse set of applicants, there is no evidence the recipients from the larger pool would perform worse than those currently being chosen.

If we return to the recommendations of Darling-Hammond and Podolsky (2019, p. 9) for addressing teacher shortages, we conclude that the Noyce program is a promising yet partial approach. It addresses the problem of teacher production directly, and the problem of teacher retention indirectly. The financial penalties directed towards scholarship recipients do seem to be effective in motivating Noyce Scholars to teach marginalized students, but ineffective in motivating them to remain for a long time in teaching, or – if they stay in teaching – to stay with low-income students. We found Noyce scholars are more likely than other teachers from their universities to teach marginalized students, but are less likely to teach for five years, and are more likely to switch to non-high-need schools.

Because the Noyce Scholarships are focused mainly on the preservice teachers themselves, they do not fully leverage the capabilities of universities to support teacher retention in the form of induction support and long-term professional development. From the perspective of educator preparation faculty, Noyce Scholarships are both highly desirable and somewhat unreliable. They come and go on five-year timescales. One cannot count on them to build a long-term program although they provide great help whenever they arrive. The reporting requirements are more onerous than with other NSF grants; the requirement to track scholarship recipients can remain in place more than a decade after the grant ends. This unfunded requirement is in itself a deterrent for faculty considering Noyce applications.

The Noyce program remains only 10% as large as the lower end of the *Gathering Storm* recommendation from 15 years ago. Thus, while it is the largest Federal effort devoted to combatting STEM teacher shortages, it is too small to affect the national problem, which the authors of the *Gathering Storm* assessed as a shortage on the order of 100,000 teachers, and which has not diminished since (Marder, 2021). If as a nation we are serious about addressing that problem, we

can learn from what has been accomplished so far, but need to think about larger, broader, and bolder solutions.

One such solution is to provide sustained support to university programs so faculty and staff are not constantly chasing grants, can focus on the work of preparing teachers, and also can put in place induction support and professional development to improve teacher retention as well as production. Federal block grants given to the states funding STEM teacher preparation and support at universities could provide such stability. The Eisenhower Math and Science State Grant Programs (Government Accounting Office, 1992), which were administered by the US Department of Education provide one example, although they were not centered on teacher preparation. A more recent example, also not centered on teacher preparation, is the Mathematics and Science Partnership grants, which also were from the US Department of Education, and last were funded in 2015 (US Department of Education, 2015).

Another solution is to support postsecondary students at a high enough financial level to change career trajectories, which the current grants rarely do. The NSF's premier support for future researchers, the Graduate Research Fellowship Program (National Science Foundation, 2021), gives a $34,000 stipend to students plus a $12,000 annual allowance to the hosting institution for five years. There is no payback provision, no requirement that recipients do research for 10 years in a high-need jurisdiction or else give the money back to the Treasury. An even more promising model is suggested by the most common form of support delivered to students who pursue STEM research careers through post-baccalaureate PhD programs. They typically receive tuition remission and a salary of around $2000 a month to serve as Research Assistants. There is no payback provision because they are earning their support.

An analogous way to support degree-holders to pursue teaching certificates would be to pay them as Teaching Assistants (TAs). A track of the Noyce program that paid future teachers to teach would form an elegant complement to the customary programs that pay future researchers to do research. Providing TAships to future teachers would make it possible to back away from financial penalties. And the students of these TAs in laboratory and recitation sections would likely benefit from instructors receiving professional preparation to teach.

Despite report after report with the most strident warnings imaginable about teacher shortages, our national remedy supports too few people with too little money and is structured around coercion that betrays a lack of trust. It is unthinkable for most NSF PIs to have to track their former research students and send collection agencies after them if they do not do the right sort of research in the right places after their PhD. For Noyce PIs it is unavoidable. The Noyce program is doing great good, has enhanced the STEM education ecosystem at universities in ways that will never fully be captured, and has supported at very least 12,000 people on their way to becoming high quality, greatly needed STEM teachers. But to address the glaring national inequities that result from the continuing lack of STEM teachers, transformative change is needed for the Robert Noyce Teacher Scholarship Program to realize its promise.

Acknowledgements

References

107th Congress. (2002). *National Science Foundation Authorization Act of 2002*. https://www.congress.gov/bill/107th-congress/house-bill/4664/text?overview=closed

COMPETES Act, (2007) (testimony of 110th Congress). https://www.congress.gov/bill/110th-congress/house-bill/2272/text?overview=closed

America COMPETES Reauthorization, (2010) (testimony of 111lth Congress). https://www.congress.gov/111/plaws/publ358/PLAW-111publ358.pdf

Alemdar, M., Cappelli, C. J., Criswell, B. A., & Rushton, G. T. (2018). Evaluation of a Noyce program: Development of teacher leaders in STEM education. *Evaluation and Program Planning*. https://doi.org/10.1016/j.evalprogplan.2018.06.005

Allen, M. (2003). Eight questions on teacher preparation. What does research say? *Education Commission of the States (ECS)*, 7.

American Association for Employment in Education. (2020). *Educator supply and demand report 2019-20*. https://specialedshortages.org/wp-content/uploads/2021/07/2019-2020-AAEE-Educator-Supply-and-Demand-Report.pdf

Amrein-Beardsley, A. (2009). Value-added tests: Buyer, be aware. *Educational Leadership*, *67*(3), 38–42. https://doi.org/10.1007/s00253-014-5773-9

Backes, B., Goldhaber, D., Cade, W., Sullivan, K., & Dodson, M. (2018). Can UTeach? Assessing the relative effectiveness of STEM teachers. *Economics of Education Review*, *64*, 184–198. https://doi.org/10.1016/j.econedurev.2018.05.002

Bendinelli, A., & Marder, M. (2012). Visualization of longitudinal student data. *Physical Review Special Topics - Physics Education Research*, *8*, 020119/1-15.

Bischoff, P., French, P., & Schaumloffel, J. (2014). Reflective pathways: Analysis of an urban science teaching field experience on Noyce Scholar-Science Education Awardees' decisions to teach science in a high-need New York City school. *School Science and Mathematics*, *114*(1), 40–49.

Bolker, B., Bates, D., Maechler, M., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Grothendieck, G., & Green, P. (2016). *Package lme4*. AAEE.

Bowe, A., Braam, M., Lawrenz, F., & Kirchhoff, A. (2011). Comparison of alternative and traditional teacher certification programs in terms of effectiveness in encouraging STEM pre-service teachers to teach in high need schools. *Journal of the National Association for Alternative Certification*, *6*(1), 26–45.

Brewer, S. (2019, October 23). *UTA's UTeach program shuts down after nine years of operation*.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Houghton Mifflin.

Cappelli, P. H. (2015). Skill gaps, skill shortages, and skill mismatches: Evidence and arguments for the United States. *Industrial and Labor Relations Review*, *68*(2). https://doi.org/10.1177/0019793914564961

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood* (Working Paper No. 17699). National Bureau of Economic Research. https://doi.org/10.1017/S0018246X06005280

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014a). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, *104*(9), 2633–2679. https://doi.org/10.1257/aer.104.9.2633

Chetty, R., Friedman, J., & Rockoff, J. (2014b). Discussion of the American Statistical Association's Statement (2014) on using value-added models for educational assessment. *Statistics and Public Policy*, *1*(1), 111–113. https://doi.org/10.1080/2330443X.2014.955227

Code.org Advocacy Coalition, Computer Science Teaching Association, & Expanding Computing Education Pathways. (2021). *2021 State of computer science education.* https://advocacy.code.org/2021_state_of_cs.pdf

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan*, *93*(6), 8–15. https://doi.org/10.1177/003172171209300603

Darling-Hammond, L., & Podolsky, A. (2019). Breaking the cycle of teacher shortages: What kind of policies can make a difference? *Education Policy Analysis Archives*, *27*(34).

Friedman, J. N., Rockoff, J. E., & Chetty, R. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104*(9), 2593–2632. https://www.aeaweb.org/articles?id=10.1257/aer.104.9.2593

Friedman, T. (2005). *The world is flat: A brief history of the twenty-first century.* Farrar, Straus and Giroux. https://doi.org/10.1021/nn403389h

Government Accounting Office. (1992). *The Eisenhower Math and Science State Grant Program.* https://www.gao.gov/products/hrd-93-25

Hanushek, E. (1971). Teacher characteristics and gains in student achievement: Estimation using micro data. *American Economic Review*, *61*, 280–288.

Hanushek, E. (2011). The economic value of higher teacher quality. *Economics of Education Review*, *30*(3), 466–479. https://doi.org/10.1016/j.econedurev.2010.12.006

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics*. https://doi.org/10.1016/j.jpubeco.2010.11.009

Hill, H. C., Kapitula, L., & Umland, K. (2011). A Validity Argument Approach to Evaluating Teacher Value-Added Scores. *American Educational Research Journal*, *48*(3), 794–831. https://doi.org/10.3102/0002831210387916

Hobson Hargraves, R. (2022). *NSF Noyce Program 20th Anniversary.* https://www.nsf.gov/ehr/Videos/NoyceAnniversary.jsp. Also announced at Fall 2019 Midwestern Noyce conference.

Hoeg, D., & Bencze, L. (2017). Rising against a gathering storm: A biopolitical analysis of citizenship in STEM policy. *Cultural Studies of Science Education*, *12*(4). https://doi.org/10.1007/s11422-017-9838-9

Imbens, G., & Rubin, D. B. (2015). *Causal inference for statistics, social, and biomedical sciences: An introduction.* Cambridge University Press.

Jackson, C. K. (2014). Teacher quality at the high school level: The importance of accounting for tracks. *Journal of Labor Economics*, *32*, 645–684. https://doi.org/10.1086/676017

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, *27*(6), 615–631. https://doi.org/10.1016/j.econedurev.2007.05.005

Kane, T. J., & Staiger, D. O. (2008). *Estimating teacher impacts on student achievement: An experimental evaluation (NBER Working Paper No. 14607).* http://www.nber.org/papers/w14607

Kirchhoff, A., & Lawrenz, F. (2011). The use of grounded theory to investigate the role of teacher education on STEM teachers' career paths in high-need schools. *Journal of Teacher Education.* https://doi.org/10.1177/0022487110397840

Koedel, C., Mihaly, K., & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review*, *47*, 180–195. https://doi.org/10.1016/j.econedurev.2015.01.006

Koedel, C., Parsons, E., Podgursky, M., & Ehlert, M. (2015). Teacher preparation programs and teacher quality: Are there real differences across programs? *Education Finance and Policy*, *10*, 508–534. https://doi.org/10.1162/EDFP_a_00172

Lang, S. (2010). *CALS to close Department of Education over next two years*. https://news.cornell.edu/stories/2010/10/cals-close-education-department-over-next-two-years

Liou, P.-Y., Desjardins, C. D., & Lawrenz, F. (2010). Influence of scholarships on STEM teachers: Cluster analysis and characteristics. *School Science & Mathematics*. https://doi.org/10.1111/j.1949-8594.2010.00016.x

Liou, P.-Y., Kirchhoff, A., & Lawrenz, F. (2010). Perceived effects of scholarships on STEM majors' commitment to teaching in high need schools. *Journal of Science Teacher Education*. https://doi.org/10.1007/s10972-010-9191-2

Liou, P.-Y., & Lawrenz, F. (2011). Optimizing teacher preparation loan forgiveness programs: Variables related to perceived influence. *Science Education*. https://doi.org/10.1002/sce.20409

Liou, P.-Y., Lawrenz, F., Madsen, C., Braam, M., & Medhanie, A. (2009). Using factor analysis and hierarchical linear modeling to analyze STEM majors' perceptions of the commitment to teach in high need schools survey. In *Online Submission*. Online Submission.

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S., & Busick, M. D. (2012). Translating the statistical representation of the effects of education interventions into more readily interpretable forms. *National Center for Special Education Research*. https://doi.org/10.2528/PIER98040300

Marder, M. (2021, October 25). *How bad is the U.S. STEM teacher shortage?* https://uteachinstitute.medium.com/u-s-stem-teacher-shortages-in-2017-2018-1d8314a93ba1

Marder, M., David, B., & Hamrock, C. (2020). Math and science outcomes for students of teachers from standard and alternative pathways in Texas. *Education Policy Analysis Archives*. https://doi.org/10.14507/epaa.28.4863

Mellor, L., Lummus-Robinson, M., Brinson, V., & Dougherty, C. (2008). *Linking teacher preparation programs to student achievement*. [Unpublished report, available on request from corresponding author of this article].

Metzler, J., & Woessmann, L. (2012). The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation. *Journal of Development Economics*. https://doi.org/10.1016/j.jdeveco.2012.06.002

Morrell, P. D., & Salomone, S. (2017). Impact of a Robert Noyce Scholarship on STEM teacher recruitment. *Journal of College Science Teaching*, *47*(2), 16–21.

Mundy, M.-A., & Ratcliff, M. (2021). Attracting Noyce students to become STEM teachers. *Journal of Instructional Pedagogies*, *25*.

National Center for Education Statistics. (2019). *Integrated postsecondary education data system*. https://nces.ed.gov/ipeds/datacenter/InstitutionList.aspx

National Commission on Excellence in Education. (1983). *A nation at risk*. https://eric.ed.gov/?id=ED226006

National Research Council. (2005). *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. http://www.nap.edu/catalog/11463.html

National Science Foundation. (2003). *Robert Noyce Scholarship Program*. https://www.nsf.gov/pubs/2003/nsf03544/nsf03544.txt

National Science Foundation. (2017). *Robert Noyce Teacher Scholarship Program*. https://www.nsf.gov/pubs/2017/nsf17541/nsf17541.htm

National Science Foundation. (2021). *Graduate Research Fellowship Program*. https://www.nsfgrfp.org/

Partelow, L. (2019). *What to make of declining enrollment in teacher preparation programs*. https://www.americanprogress.org/article/make-declining-enrollment-teacher-preparation-programs/

Pearl, J. (2009). *Causality* (2nd ed.). Cambridge University Press.

Pivovarova, M., Amrein-Beardsley, A., & Broatch, J. (2016). Value-added models (VAMs): Caveat emptor. *Statistics and Public Policy*, *3*(1). https://doi.org/10.1080/2330443X.2016.1164641

Rivkin, S., Hanushek, E., & Kain, J. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*, 417–458. https://doi.org/10.1111/j.1468-0262.2005.00584.x

Sanders, W., & Rivers, J. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. University of Tennessee Value-Added Research and Assessment Center.

Scott, T. P., Milam, J. L., Stuessy, C. L., Blount, K. P., & Bentz, A. (2006). Math and Science Scholars (MASS) Program: A model program for the recruitment and retention of preservice mathematics and science teachers. *Journal of Science Teacher Education*. https://doi.org/10.1007/s10972-006-9026-3

Sharma, A. (2016). STEM-ification of education: The zombie reform strikes again. *Journal for Activist Science & Technology Education*, *7*(1).

Sharma, A., & Hudson, C. (2021). Depoliticization of educational reforms: the STEM story. *Cultural Studies of Science Education*. https://doi.org/10.1007/s11422-021-10024-0

Sutcher, L., Darling-Hammond, L., & Carver-Thomas, D. (2019). Understanding teacher shortages: An analysis of teacher supply and demand in the United States. *Education Policy Analysis Archives*, *27*, 1–36. https://doi.org/10.14507/epaa.27.3696

Ticknor, C. S., Gober, D., Howard, T., Shaw, K., & Mathis, L. A. (2017). The influence of the CSU Robert Noyce Teacher Scholarship Program on undergraduates' teaching plans. *Georgia Educational Researcher*, *14*(1), 69–97.

Turner, H. M., Goodman, D., Adachi, E., Brite, J., & Decker, L. E. (2012). *Evaluation of Teach For America in Texas schools*. https://www.academia.edu/21833004/Evaluation_of_Teach_For_America_in_Texas_Schools

US Department of Education. (2015). *Mathematics and science partnerships*. https://www2.ed.gov/programs/mathsci/funding.html

US Department of Education. (2020). *Civil Rights Data Collection (CRDC) for the 2017-2018 school year*. https://ocrdata.ed.gov/resources/downloaddatafile

US Department of Education. (2021). *Title II Higher Education Act data tables*. https://title2.ed.gov/Public/DataTools/Tables.aspx

von Hippel, P. T., & Bellows, L. (2018). How much does teacher quality vary across teacher preparation programs? Reanalyses from six states. In *Economics of Education Review*. https://doi.org/10.1016/j.econedurev.2018.01.005

von Hippel, P. T., Osborne, C., Lincove, J., Bellow, L., & Mills, N. (2016). Teacher quality differences between teacher preparation programs: How big? How reliable? Which programs are different? *Economics of Education Review*, *53*, 31–45. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2506935

Walkington, C., & Marder, M. (2015). Classroom observation and value-added models give complementary information about quality of mathematics teaching. In *Designing teacher evaluation systems*. https://doi.org/10.1002/9781119210856.ch8

White House. (2022). *The Build Back Better Framework*. https://www.whitehouse.gov/build-back-better/

Zahner, W., Chapin, S., Levine, R., He, L., & Afonso, R. (2019). Examining the recruitment, placement, and career trajectories of secondary mathematics teachers prepared for high-need S=schools. *Teachers College Record*, *121*(2).

Zeidler, D. L. (2016). STEM education: A deficit framework for the twenty first century? A sociocultural socioscientific response. *Cultural Studies of Science Education*, *11*(1). https://doi.org/10.1007/s11422-014-9578-z

About the Authors

Michael Marder
The University of Texas at Austin
marder@chaos.utexas.edu
https://orcid.org/0000-0003-0197-4459
Michael Marder is Professor of Physics and Executive Director of UTeach at The University of Texas at Austin. His interest in studying educational data stemmed from over 20 years of work preparing STEM teachers at UT Austin and dozens of other universities.

Catherine Horn
University of Houston
clhorn2@uh.edu
https://orcid.org/0000-0003-4959-5810
Dr. Catherine Horn is Moores Professor and Chair, Department of Educational Leadership and Policy Studies, and Executive Director of the Institute for Educational Policy Research and Evaluation within the College of Education at the University of Houston. She is also the Director for the Center for Research and Advancement of Teacher Education and focuses on the systemic influences of secondary and postsecondary assessment and related policies on the learning trajectories of students especially for students traditionally underserved by the education and social sectors.

Sarah Stephens
University of Wisconsin-Milwaukee
stephese@uwm.edu
https://orcid.org/0000-0003-1729-1812
Dr. Sarah Stephens is an Associate Lecturer of Physics at The University of Wisconsin-Milwaukee, and previously at The University of Texas at Austin. As a graduate student and post-doctoral researcher at UT Austin and the Texas Education Research Center, Sarah worked on a variety of educational research projects that investigated topics including longitudinal outcomes of K-12 assessments, teacher preparation pathways, long-term career and technology education outcomes, and STEM course-taking patterns.

Annelies Rhodes
The University of Texas at Austin
arhodes@utexas.edu
https://orcid.org/0000-0003-2888-0698
Annelies Rhodes is a graduate student in STEM Education at the University of Texas at Austin. Her interests lie in the use of educational data to inform policy. Her current projects include the use of GPA as a criterion for teacher certification and the effects of teachers from various certification pathways on student learning.

education policy analysis archives

About the Editorial Team: https://epaa.asu.edu/ojs/index.php/epaa/about/editorialTeam

Please send errata notes to Audrey Amrein-Beardsley at audrey.beardsley@asu.edu

Join EPAA's Facebook community at https://www.facebook.com/EPAAAAPE and **Twitter feed** @epaa_aape.