THREE ESSAYS ON ECONOMICS OF EDUCATION

A Dissertation

Presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

YANG AN

Dr. Cory Koedel, Dissertation Supervisor

DECEMBER 2020

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

THREE ESSAYS ON ECONOMICS OF EDUCATION

presented by Yang An,

a candidate for the degree of Doctor of Philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

_____

Professor Cory Koedel

_____

Professor David Kaplan

_____

Professor Oded Gurantz

_____

Professor Peter Mueser

ACKNOWLEDGEMENTS

I am very grateful to my advisor, Dr. Cory Koedel, for guiding me throughout my Ph.D. studies. Without his invaluable advice and continuous encouragement, I would not have earned my doctoral degree. I would like to thank my other committee members, Dr. Oded Gurantz, Dr. David Kaplan, and Dr. Peter Mueser, for their valuable suggestions on my dissertation. I would also like to thank Dr. Mark Ehlert and Dr. Eric Parsons for their support during my time as a research assistant for the Economic and Policy Analysis Research Center (EPARC). Last but not least, I would like to express my gratitude to my parents and my wife, Yating, for always supporting me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This dissertation includes three chapters. In Chapter 1, I estimate the effects of gender and race/ethnicity matching between students and high school math and science teachers on students' STEM enrollment and completion in colleges. My sample includes over 100,000 Missouri high school graduates entering a four-year public university in Missouri between 2001 and 2010. Potential bias from non-random teacher-student sorting within and across high schools is mitigated by examining student exposure to teaching staff broadly, and the use of within-high school variation for identification. I find no evidence of matching effects on students' postsecondary STEM outcomes. My findings for gender matching are more precise than for race/ethnicity and rule out modestly positive impacts.

In Chapter 2, I estimate the effects of nursing articulation agreements between the University of Missouri-Columbia (MU) and numerous public community colleges on the likelihood that community college students transfer to the MU four-year nursing degree program. I use difference-in-differences specifications that leverage variation in the availability of nursing articulation agreements with MU across community colleges and over time to identify the effects of these agreements. I supplement these models with triple-difference specifications that additionally leverage variation in transfer rates across initial majors in community colleges. I find no statistical evidence that nursing articulation agreements affect student transfer rates. Supplementary analyses indicate that my null findings are not driven by supply constraints in the MU nursing program and instead reflect the failure of the articulation agreements to spur demand for nursing education among community college students.

In Chapter 3, we examine how teachers from two alternative preparation programs—Teach for America (TFA) and Kansas City Teacher Residency (KCTR)—contribute to the teacher labor market in and around Kansas City, Missouri. We show that TFA and KCTR teachers are more likely than other teachers to work in charter schools, and more broadly, in schools with high concentrations of low-income, low-performing, and underrepresented minority (Black and Hispanic) students. TFA and KCTR teachers are themselves more racially/ethnically diverse than the larger local-area teaching workforce, but only KCTR teachers are more diverse than teachers in the same districts in which they work. In math in grades 4-8, we find sizeable, positive impacts of TFA and KCTR teachers on test-score growth relative to non-program teachers. We also estimate positive impacts on test-score growth in English Language Arts (ELA) for teachers from both programs, but our ELA estimates are smaller in magnitude.

Chapter 1

High School Teacher-student Matching and Postsecondary STEM Outcomes

1.1. Introduction

Labor demand in science, technology, engineering, and mathematics (STEM) fields has been increasing; however, there are not enough native skilled workers in these fields (U.S. Congress Joint Economic Committee, 2012; Atkinson, 2013). The proportion of U.S. students declaring STEM majors is low and attrition rates in STEM fields are high, especially for women and underrepresented minorities (Chen, 2013). For example, in 2011-12, only 11.7 percent of those who declared a STEM major were women (National Science Board, 2016). Despite relatively low attrition rates for women in biological science, the attrition rates are large for women in computer science, engineering, physical science, and mathematics, and even larger for underrepresented minority women (National Girls Collaborative Project, 2015).

Moreover, blacks and Hispanics comprise only 12 percent of STEM workers, but 28 percent of overall employment (Beede et al., 2011). The underrepresentation of these racial/ethnic groups in STEM fields stretches racial wealth gaps. This is because STEM workers earn more than non-STEM workers, regardless of the fields in which they work (Noonan, 2017). However, students from these racial/ethnic groups are less likely to major in STEM fields or to find a STEM job (U.S. Congress Joint Economic Committee, 2012). In addition, as documented in Arcidiacono, Aucejo, and Spenner (2012), more black students have shifted their interests in STEM fields to non-STEM fields. Hence, increasing

the representation of women and minorities in STEM fields is an increasingly important education policy issue. To address this concern, more research is needed to understand the factors that influence women's and minorities' decisions to declare and graduate with STEM majors. One hypothesis is that students are likely to perform better if they are taught by a STEM teacher of the same gender and/or race/ethnicity in high schools. If true, a policy response to improve STEM outcomes for women and underrepresented minorities in college would be to improve the diversity of the STEM teaching workforce along these dimensions.

Many studies examine the effects of teacher-student matching by gender and race/ethnicity on student outcomes, and results are mixed. It has been shown that in secondary schools, female students perform better on tests when assigned to a female teacher (Dee, 2007; Lim and Meer, 2017) and female teachers are less likely to see female students as disruptive or inattentive (Dee, 2005; Dee, 2007). Gender matching effects are also found in postsecondary education. Studies find that female professors make female students more likely to major in the subjects that they teach (Bettinger and Long, 2005) and perform substantially better (Carrell, Page, and West, 2010). In addition, Kofoed and McGovney (2019) use the random assignment of cadets to tactical officers at West Point and show that female cadets assigned to a female officer are more likely to choose their officer's occupation. On the other hand, some researchers show that there is no evidence of gender matching effects (Canes and Rosen, 1995; Ehrenberg, Goldhaber and Brewer, 1995; Price, 2010). A particularly notable study by Carrell, Page, and West (2010) relies on the random assignment of students to professors and finds that having a female professor does not increase a female student's probability of graduating with a STEM major except

for those with advanced math skills. In contrast to the literature on gender matching, research on racial/ethnic matching is more consistent in the sense that students seem to benefit more from same-race/ethnicity instructors or mentors. For example, Ehrenberg and Brewer (1994) show that the high school dropout rate for Hispanic students decreases when the share of Hispanic teachers increases. Moreover, Dee (2004, 2005) finds that assignment to an own-race teacher increases a student's achievement and decreases the likelihood of the student being perceived as disruptive. Also, black teachers have higher expectations for black students compared to non-black teachers (Gershenson, Holt, and Papageorge, 2016). In postsecondary education, black students are more likely to persist after one year when first taught by black faculty (Price, 2010).[1]

The literature focused on understanding teacher-student matching and how it affects outcomes contained within secondary or postsecondary education has been widely developed. However, there is less research examining the effects of high school teacher-student matching on students' postsecondary STEM outcomes. The study most closely relevant to my work is Bottia et al. (2015). They use data on students who graduated from North Carolina public high schools in 2004 and attended an institution in the University of North Carolina system. They find that the proportion of female math and science teachers has positive and statistically significant effects on female students' probabilities of declaring and graduating with STEM majors. They rely on variation in the proportion of female math and science teachers across high schools for identification. However, teacher-student sorting across schools may cause bias in estimates of gender matching effects.

[1] In addition to the finding of gender matching effects, Kofoed and McGovney (2019) find that black cadets assigned to a black officer are more likely to choose their officer's occupation.

This paper contributes to this limited literature by examining the effects of high school teacher-student matching by gender and race/ethnicity on college STEM enrollment and degree attainment. To draw causal inference, I must alleviate endogeneity from non-random teacher-student sorting, as within- and cross-school teacher-student sorting could bias my estimates. To mitigate bias from within-school sorting, I follow on the work of Rose and Betts (2004) and consider student exposure to teaching staff broadly, rather than the precise teachers of each student. Following Darolia et al. (2020), I address potential endogeneity bias due to cross-school sorting using high school fixed effects. Therefore, identification in my study relies on variation in the gender and racial compositions of teachers within high schools over time.

I find no evidence of matching effects on students' postsecondary outcomes. My estimates for gender matching are more informative and precise than for race/ethnicity. A one-standard-deviation increase in the share of female math and science teachers affects female students' STEM enrollment and completion rates by about 0.38 and 0.03 percentage points, respectively. With 95 percent confidence, I can rule out moderately positive effects and the estimates of Bottia et al. (2015).

1.2. Methodology

Non-random sorting will cause bias in estimates of teacher-student matching effects and make it hard to draw causal inference. As noted by Clotfelter, Ladd, and Vigdor (2006), teacher characteristics may correlate with unobserved factors which also affect student outcomes. This endogeneity can be driven by within-school and cross-school teacher-student sorting.

To address within-school sorting, I follow the approach used by Rose and Betts (2004) and consider the effects of exposure to female and underrepresented minority (defined here as blacks and Hispanics) STEM teachers in high schools:

$$Y_{is} = \emptyset_0 + \emptyset_1 X_i + \emptyset_2 Z_s + \emptyset_3 FT_s + \emptyset_4 FT_s FS_i + \emptyset_5 MT_s + \emptyset_6 MT_s MS_i + \epsilon_{is} \qquad (1)$$

In Equation (1), Y_{is} is an indicator variable measuring a STEM outcome for student *i* who graduated from high school *s*. One STEM outcome is whether a student declared a STEM major when entering a university and the other is whether the student graduated from a university with a STEM degree. X_i is a vector of observed student variables, including the student's race/ethnicity, gender, ACT math and English scores, high school percentile ranks, and missing-value indicators to account for missing ACT scores and high school percentile ranks. Z_s is a vector of high school controls, containing total enrollment, the pupil-teacher ratio, the share of female students, the share of underrepresented minority (again as blacks and Hispanics) students, the share of students who are free or reduced-price lunch eligible, the share of experienced (at least 10-year experience in all public school systems) teachers, the share of teachers with a master's or doctoral degree, and average total teacher salaries. FS_i and MS_i are indicator variables equal to one if student *i* is female and an underrepresented minority, respectively. FT_s and MT_s are the share of female math and science teachers and the share of underrepresented minority math and science teachers in high school *s*, respectively. ϵ_{is} is an error term.

Within a high school, FT_s and MT_s measure student exposure to female and underrepresented minority STEM teachers, respectively. Using these shares, rather than measures based on the precise teachers of each student, mitigates the concern of bias from within-school teacher-student sorting. Next, I need to address cross-school sorting.

Variation across high schools in these shares may be correlated with unobserved factors which also influence students' STEM outcomes.

To deal with cross-school sorting, I employ a model following Darolia et al. (2020):

$$Y_{ist} = \alpha_0 + \alpha_1 X_{it} + \alpha_2 Z_{st} + \alpha_3 FT_{st} + \alpha_4 FT_{st} FS_{it} + \alpha_5 MT_{st} + \alpha_6 MT_{st} MS_{it} + h_s + \pi_t + \epsilon_{ist} \quad (2)$$

Rich administrative data allow me to observe multiple cohorts of high school graduates over time. Thus, I can implement a panel regression based on Equation (2). Time index *t* represents the year of high school graduation for each cohort. High school and year fixed effects are denoted by h_s and π_t, respectively. Potential bias from cross-school sorting is diminished by including high school fixed effects. Identification relies on variation in FT_{st} and MT_{st} within high schools over time.

1.3. Data

Student-level administrative records are provided by the Missouri Department of Higher Education (DHE). I build a 10-year data panel tracking first-time, full-time, degree-seeking freshmen who graduated from a Missouri public high school and attended a Missouri four-year public university. As shown in Appendix Table A.1, my analytical sample includes 109,930 students entering the university system between 2001 and 2010.

I track each student's graduation status within six years of entry. Intended and completed majors are identified by CIP (Classification of Instructional Programs) codes developed by the U.S. Department of Education's National Center for Education Statistics (NCES). I follow the STEM designated degree program list provided by the Department of Homeland Security to determine whether a major is in a STEM field.

As shown in Appendix Table A.2, Missouri has 13 public four-year universities. In the sample, 25 percent of students enrolled at the University of Missouri-Columbia, which accounted for 30 percent of total STEM entrants and 32 percent of total STEM graduates. Although the Missouri University of Science and Technology had only 5 percent of sample enrollees, it accounted for 21 percent of STEM entrants and 22 percent of STEM graduates.

The DHE data include student-level variables that I use as controls in the model, such as race/ethnicity, gender, ACT math and English scores, and high school percentile ranks. Sample summary statistics are shown in Table 1.1. Female students make up 56 percent of the sample. Eighty-three percent of the sample is white, 9 percent is black, 2 percent is Hispanic, 2 percent is Asian/Pacific Islander, and 4 percent is some other race/ethnicity.[2]

Table 1.2 shows that 22 percent of the sample entrants into the four-year public university system initially chose a STEM major, and 13 percent graduated with a STEM degree. Among entrants with an initial STEM major, 44 percent completed a degree within a STEM field. Entrants with initial non-STEM majors are unlikely to graduate with a STEM major (4 percent). Men are more likely than women to declare and complete a STEM major. For example, among entrants, 32 percent of men initially enrolled in a STEM program and 19 percent completed a STEM degree. In comparison, only 14 percent of women initially enrolled in a STEM program and 8 percent completed a STEM degree. Among graduates, 34 percent of men graduated with a STEM degree, while this number is just 13 percent for women. As shown in the last two columns of Table 1.2, men are more likely than women to attain a STEM degree no matter what initial majors they chose.

[2] Fifty-one percent of Missouri's population is female, 81 percent is white, 12 percent is black, 3 percent is Hispanic, 2 percent Asian/Pacific Islander, and 3 percent is some other race/ethnicity.

Asians/Pacific Islanders are the most likely to choose and graduate with a STEM major, while blacks are the least likely.

The DHE data are merged by high school codes with teacher-level administrative records provided by the Missouri Department of Elementary and Secondary Education (DESE). For each student, I construct measures of the teacher gender and race/ethnicity shares in math and science over the last three years of high schools. I choose three years because some high schools only include grades 10-12. As shown in Table 1.1, the student-weighted mean and standard deviation of the share of female math and science teachers are 0.56 and 0.15, respectively. The share of underrepresented minority math and science teachers has a mean of 0.03 and a standard deviation of 0.08.

The shares of female and underrepresented minority STEM teachers are each measured in two ways. For each high school, I first calculate class-enrollment-weighted total teaching minutes of female teachers and underrepresented minority teachers in math and science. I then create a primary measure of the share of female or underrepresented minority STEM teachers by taking the ratio of each of these calculated teaching minutes to total class-enrollment-weighted teaching minutes of all teachers in math and science. The second measure is based on each teacher's primary teaching subject, measured as the subject with the most teaching minutes. The shares of female and underrepresented minority STEM teachers are defined as the proportions of female STEM teachers and underrepresented minority STEM teachers.

High school controls, such as total enrollment, the pupil-teacher ratio, the share of female students, the share of underrepresented minority students, and the share of students who are free or reduced-price lunch eligible, are from the Common Core of Data (CCD)

provided by the NCES. I calculate a three-year average for each of these measures and merge these variables to the DHE data by each student's high school code and year of high school graduation. The analytical sample has 109,930 students who graduated from 507 Missouri public high schools between 2001 and 2010 and attended a Missouri four-year public university.

My analysis is conditional on students' university attendance. If female and underrepresented minority STEM teachers affect university enrollment, it would be difficult to interpret the matching estimates. This is because students on the margin of attending colleges may be less likely to declare and complete a STEM major, given that students in STEM fields are high performers in their high schools on average.[3] If no such effects are found, I can focus on college students' changes in initial majors and degree attainment. Conditional on enrollment, I test for an effect of the female and underrepresented minority STEM teacher shares on enrollment in the Missouri four-year public university system as follows. For each high school in my sample, I consider the number of students who matriculated to a Missouri four-year public university in each corresponding cohort as the dependent variable. The share of female STEM teachers, the share of underrepresented minority STEM teachers, and the other high school controls are the independent variables. High school and year fixed effects are also included. As shown in Appendix Table A.3, I find no statistically significant effects of female or underrepresented minority STEM teachers on students' college attendance. Although this is not a conclusive test, it reduces concerns about bias from selection into my sample.

1.4. Results

[3] The average high school percentile rank (77.1 percent) among STEM entrants is higher than the sample average (71.9 percent).

1.4.1. Primary Findings

Linear probability models are estimated under different specifications, and regression results using the primary measures of the female and underrepresented minority STEM teacher shares are reported in Table 1.3. Columns (3) and (6) present the results using the full specification as shown in Equation (2). The dependent variable in Columns (1)-(3) measures whether a student declared a STEM major when entering a university, and in Columns (4)-(6) it is whether a student graduated with a STEM major.

The second row of Table 1.3 shows estimates of gender matching effects. The coefficients are small in magnitude, not statistically significant, and inconsistent in sign across specifications. Given that the standard deviation of the share of female math and science teachers is 0.15 in my sample, Columns (3) and (6) show that a one-standard-deviation increase in the share of female math and science teachers to which a female student is exposed has effects of about 0.38 and 0.03 percentage points on STEM enrollment and completion, respectively. The sample means of STEM enrollment and attainment for women are 0.14 and 0.08, respectively. These effects account for about 2.7 and 0.38 percent of the sample means of these outcomes. Moreover, with 95 percent confidence, I can rule out positive effects larger than 1.8 and 6.6 percent of the sample means of STEM enrollment and attainment for women, respectively.

The last row of Table 1.3 presents estimates of race/ethnicity matching effects. As shown in Columns (3) and (6), there is no evidence that racial/ethnic matching between students and high school STEM teachers affects students' postsecondary STEM outcomes, and the point estimates are negative. There is less identifying variation in the share of underrepresented minority STEM teachers, however, given that it has a standard deviation

of just 0.08 in the sample. The lack of statistical power resulting in larger standard errors makes the estimates less precise and informative compared to those of gender matching.

1.4.2. Heterogeneity

The models in Table 1.3 examine the effects of gender and race/ethnicity matching between students and high school math and science teachers. Next, I explore the possibility of differential effects of female or underrepresented minority math teachers versus female or underrepresented minority science teachers. Math courses are further categorized into high school (regular) and college (advanced) levels based on course titles, and the possibility of heterogeneous effects along this dimension is also considered. The definitions of the shares of female and underrepresented minority STEM teachers in each case are analogous to my primary measures defined above.

As shown in Table 1.4, there is no evidence of gender and race/ethnicity matching effects on STEM enrollment and degree attainment in advanced math. However, although I do find statistically significant and negative gender matching effects on STEM enrollment in regular math, I am cautious to draw strong inference from these results because the point estimates are small in magnitude. The point estimates of matching effects in science are statistically insignificant, with only one exception, which is the coefficient of race/ethnicity matching in Column 8 for the degree attainment model. Given the imprecise estimates of race/ethnicity matching, taken as a whole, there is no indication of matching effects in science.

1.5. Robustness and Extension

1.5.1. Robustness

In order to ensure that my results using primary measures are robust, I estimate the models specified in Table 1.3 using alternative measures of the shares of female and underrepresented minority math and science teachers. The alternative measures are based on the subject in which a teacher has taught the most minutes. For example, a female teacher who has taught more minutes of math than in any other subject is considered a female math teacher. For brevity, I only report estimates from the fully specified models. As shown in Appendix Table A.4, the estimates are close to those using the primary measures. There is no indication that exposure to female and underrepresented minority STEM teachers in high schools affects female and underrepresented minority students' respective postsecondary STEM outcomes.

Given that the identification relies on variation in the gender and racial compositions of teachers within high schools over time, endogenous changes to the shares of female and underrepresented minority math and science teachers within high schools over time could cause bias in estimates of teacher-student matching effects. For example, high schools aiming to improve STEM interests or training for female or underrepresented minority students may hire more female or underrepresented minority STEM teachers. If true, estimates of matching effects would be biased upward. Since endogenous changes are less likely to happen in a shorter period, I split the data panel in half and test the possibility of this endogeneity by estimating the fully specified models in each period separately. If the estimates are not substantially different from those using the full data panel reported in Table 1.3, this reduces concerns about endogenous changes within schools in the full data panel. Per Appendix Table A.5, the matching estimates for the STEM enrollment and completion models using the partial samples are similar to my findings in Table 1.3 and

mostly statistically insignificant. Overall, I conclude that endogenous changes of the shares of female and underrepresented minority STEM teachers within high schools over time are unlikely to drive my findings.

1.5.2. Extension

Female students are underrepresented regarding initial enrollment and degree completion in natural sciences, which include engineering, mathematics, computer sciences, and physical sciences. Specifically, as shown in Table 1.5, 20 percent of entrants and 21 percent of graduates are female. However, there is no female underrepresentation in biology, agriculture, or health, given that 61 percent of entrants and 56 percent of graduates are female. Despite the underrepresentation of women in natural sciences, the distributions of underrepresented minority students are more even in both areas regarding initial enrollment and degree completion. For example, six percent of graduates are underrepresented minorities in each area.

Although my primary findings indicate no evidence of matching effects on students' STEM outcomes, it is still worth examining the possibility of differential effects of matching between high school STEM teachers and students majoring in biology, agriculture, or health versus other natural sciences. Given the underrepresentation of female students in natural sciences, a policy response to improve their STEM outcomes would be to increase the share of female high school STEM teachers if there is evidence of gender matching effects in natural sciences. Thus, I analyze the effects of high school teacher-student matching by gender and race/ethnicity on STEM enrollment and degree attainment in biology, agriculture, or health versus other natural sciences. As shown in Table 1.6, all matching estimates are statistically insignificant. Therefore, there is no

evidence of gender or race/ethnicity matching effects on students' STEM enrollment and completion in both fields.

1.6. Conclusion

Role model effects predict that students are likely to perform better when assigned to a teacher of the same gender and/or race/ethnicity. The effects of teacher-student matching by gender and race/ethnicity on student outcomes in both secondary and postsecondary education have been studied broadly. However, the body of literature examining the effects of high school teacher-student matching on students' college outcomes is limited. I use administrative student records from Missouri, including 10 cohorts of entering college students to estimate the effects of gender and race/ethnicity matching between students and high school math and science teachers on students' STEM enrollment and completion in colleges.

Non-random high school teacher-student sorting will potentially cause bias in estimates of teacher-student matching effects. I use student exposure to teaching staff broadly at the school-cohort level, rather than measure exposure based on the precise teachers of each student, as well as high school fixed effects. This mitigates concerns about bias from within- and cross-school sorting. The identification relies on variation in the gender and racial compositions of teachers within high schools over time. I find no evidence of matching effects along either dimension on students' postsecondary STEM outcomes. My findings for gender matching are more informative and precise than for race/ethnicity. Given that there are not many underrepresented minority STEM teachers in Missouri, there is less identifying variation in the share of underrepresented minority STEM teachers. The lack of statistical power makes my estimates for race/ethnicity

matching imprecise and less informative. Moreover, my insignificant estimates for gender matching imply small effects even taken at face value and I can rule out moderately positive impacts. Specifically, a one-standard-deviation increase in the share of female math and science teachers to which a female student is exposed has effects of about 0.38 and 0.03 percentage points on STEM enrollment and completion, respectively. With 95 percent confidence, I can rule out effects larger than 1.8 and 6.6 percent of the sample means of these outcomes.

Overall, my findings suggest that a high school STEM teacher of the same gender and/or race/ethnicity does not affect student postsecondary STEM interests and success. As suggested by Darolia et al. (2020), interventions on the intensive margin, such as changes in the norm of high school STEM instruction, may be an option to affect postsecondary STEM outcomes. Therefore, to increase the representation and success of women and underrepresented minorities in STEM fields, more research is needed on examining the effectiveness of the interventions.

Table 1.1. Sample Summary Statistics.

		Mean	SD		
A. Students					
Male		0.44			
Female		0.56			
White		0.83			
Black		0.09			
Hispanic		0.02			
Asian/Pacific Islander		0.02			
Other Race/Ethnicity		0.04			
High School Percentile Rank		71.9	21.0		
High School Percentile Rank Missing Indicator		0.04	0.20		
ACT Math		22.7	4.7		
ACT Math Missing Indicator		0.01	0.12		
ACT English		23.4	5.2		
ACT English Missing Indicator		0.01	0.12		
Number of Students	109,930				
B. High Schools		School-year Weighted		Student Weighted	
		Mean	SD	Mean	SD
Female Math and Science Teacher Share		0.57	0.22	0.56	0.15
Minority Math and Science Teacher Share		0.02	0.10	0.03	0.08
Female Student Share		0.49	0.03	0.49	0.02
Minority Student Share		0.10	0.21	0.14	0.20
Free or Reduce-price Lunch Share		0.33	0.16	0.22	0.15
Experienced Teacher Share		0.53	0.11	0.55	0.09
The Share of Teachers with Advanced Degrees		0.41	0.15	0.52	0.15
Average Teacher Salaries		36,477	6,999	42,099	7,101
Enrollment		575	545	1,198	659
Pupil-Teacher Ratio		14	3	16	3
Number of High Schools	507				

Table 1.2. STEM Enrollment and Completion.

	Initial STEM Majors/Entrants	STEM Degrees/Entrants	STEM Degrees/Graduates	STEM Degrees/Entrants with Initial STEM Majors	STEM Degrees/Entrants with Initial Non-STEM Majors
All Students	0.22	0.13	0.22	0.44	0.04
Male	0.32	0.19	0.34	0.47	0.06
Female	0.14	0.08	0.13	0.40	0.03
White	0.22	0.13	0.22	0.46	0.04
Black	0.18	0.06	0.16	0.24	0.02
Hispanic	0.23	0.12	0.22	0.38	0.03
Asian/Pacific Islander	0.31	0.22	0.35	0.51	0.09
Other Race/Ethnicity	0.23	0.13	0.25	0.42	0.05

Table 1.3. STEM Enrollment and Completion Models, Matching Effects.

	STEM Enrollment			STEM Completion		
	(1)	(2)	(3)	(4)	(5)	(6)
FT	0.024 (0.019)	0.020 (0.019)	0.007 (0.022)	-0.001 (0.016)	-0.001 (0.015)	-0.017 (0.019)
FT X FS	-0.025 (0.020)	-0.025 (0.020)	-0.025 (0.021)	0.001 (0.017)	0.0004 (0.017)	0.002 (0.017)
MT	0.100 (0.051)**	0.145 (0.050)***	0.130 (0.070)*	0.045 (0.036)	0.075 (0.031)**	0.087 (0.045)*
MT X MS	-0.072 (0.051)	-0.138 (0.051)***	-0.064 (0.067)	-0.107 (0.040)***	-0.083 (0.035)**	-0.024 (0.048)
Individual Controls & Year FE	X	X	X	X	X	X
HS Controls		X	X		X	X
HS FE			X			X
N	109,930	109,930	109,930	109,930	109,930	109,930

Notes: FT and MT denote the share of female math & science teachers and the share of underrepresented minority math & science teachers, respectively. FS and MS are indicator variables equal to one if a student is female and an underrepresented minority, respectively. Standard errors clustered by high schools are shown in the parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.4. STEM Enrollment and Completion Models, Subject and Math-level Heterogeneity.

	STEM Enrollment				STEM Completion			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
FT_adv	0.009 (0.015)	0.0003 (0.016)		0.001 (0.016)	-0.019 (0.012)	-0.023 (0.014)*		-0.022 (0.014)
FT_adv X FS	-0.002 (0.013)	0.016 (0.015)		0.016 (0.015)	0.009 (0.012)	0.021 (0.014)		0.021 (0.014)
MT_adv	0.033 (0.044)	-0.009 (0.044)		-0.007 (0.044)	0.018 (0.023)	0.009 (0.025)		0.007 (0.023)
MT_adv X MS	-0.025 (0.042)	0.006 (0.044)		0.006 (0.044)	0.011 (0.030)	-0.007 (0.037)		-0.0004 (0.032)
FT_reg		0.018 (0.014)		0.017 (0.014)		0.008 (0.013)		0.008 (0.013)
FT_reg X FS		-0.033 (0.015)**		-0.032 (0.015)**		-0.022 (0.014)		-0.023 (0.014)
MT_reg		0.083 (0.043)*		0.076 (0.043)*		0.033 (0.019)*		0.019 (0.020)
MT_reg X MS		-0.040 (0.041)		-0.035 (0.041)		0.021 (0.035)		0.043 (0.033)
FT_sci			-0.003 (0.016)	-0.004 (0.016)			-0.002 (0.014)	-0.002 (0.014)
FT_sci X FS			-0.012 (0.016)	-0.010 (0.016)			0.004 (0.013)	0.005 (0.013)
MT_sci			0.076 (0.069)	0.047 (0.071)			0.077 (0.050)	0.085 (0.053)
MT_sci X MS			-0.050 (0.076)	-0.022 (0.081)			-0.077 (0.048)	-0.111 (0.052)**
Individual Controls & Year FE	X	X	X	X	X	X	X	X
HS Controls	X	X	X	X	X	X	X	X
HS FE	X	X	X	X	X	X	X	X
N	109,930	109,930	109,930	109,930	109,930	109,930	109,930	109,930

Notes: FT_adv and MT_adv denote the shares of female and underrepresented minority advanced math teachers. FT_reg and MT_reg denote the shares of female and underrepresented minority regular math teachers. FT_sci and MT_sci denote the shares of female and underrepresented minority science teachers. FS and MS are indicator variables equal to one if a student is female and an underrepresented minority, respectively. Standard errors clustered by high schools are shown in the parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.5. Gender and Race/Ethnicity and STEM Majors.

	Biology/Agriculture/Health		Natural Sciences	
	STEM Enrollment	STEM Completion	STEM Enrollment	STEM Completion
Male	39%	44%	80%	79%
Female	61%	56%	20%	21%
White	83%	86%	84%	87%
Black	7%	4%	8%	4%
Hispanic	2%	2%	2%	2%
Asian/Pacific Islander	4%	4%	3%	3%
Other Race/Ethnicity	4%	4%	5%	5%
Number of Students	8,750	5,251	14,895	8,289

Table 1.6. STEM Enrollment and Completion Models, Matching Effects, by Majors.

	STEM Enrollment		STEM Completion	
	Biology/Agriculture/Health	Natural Sciences	Biology/Agriculture/Health	Natural Sciences
FT	0.023 (0.013)*	-0.017 (0.020)	0.001 (0.009)	-0.017 (0.015)
FT X FS	-0.011 (0.013)	-0.017 (0.020)	-0.009 (0.008)	0.007 (0.015)
MT	0.048 (0.043)	0.080 (0.062)	-0.007 (0.027)	0.106 (0.034)***
MT X MS	-0.022 (0.035)	-0.038 (0.056)	0.006 (0.023)	-0.035 (0.040)
Individual Controls & Year FE	X	X	X	X
HS Controls	X	X	X	X
HS FE	X	X	X	X
N	109,930	109,930	109,930	109,930

Notes: FT and MT denote the share of female math & science teachers and the share of underrepresented minority math & science teachers, respectively. FS and MS are indicator variables equal to one if a student is female and an underrepresented minority, respectively. Standard errors clustered by high schools are shown in the parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Chapter 2

Nursing Articulation Agreements and Transfers to a Baccalaureate Nursing Program

2.1. Introduction

The American Association of Colleges of Nursing (AACN, 2019) expects the U.S. to experience a persistent shortage of Registered Nurses (RNs). Demand for RNs is increasing with the aging population and corresponding increase in the need for health care. According to Torpey (2018), projected total job openings for RNs are over two million between 2016 and 2026. However, the production of nurses from postsecondary nursing programs across the country is not keeping pace. Labor shortages for nurses with four-year college degrees are particularly stark and professional organizations have argued that the U.S. must increase the educational credentials of the nursing workforce. For example, in 2010, at which point the share of all nurses with baccalaureate-level credentials stood at 0.5, the Institute of Medicine (IOM, 2011) proposed to increase this share to 0.8 by 2020.

Hospitals and postsecondary nursing programs have made some efforts to address the nursing shortage. For example, hospitals typically offer hiring bonuses, work to improve job satisfaction, and use competitive pay schemes to attract and retain nurses. Postsecondary nursing programs have formed partnerships with local health care providers to increase the number of nurses. A specific example is a partnership established in 2013 between the School of Nursing at the University of Minnesota and the Minneapolis VA Health Care System, which aimed to produce 100 additional nurses with a Bachelor of Science in nursing (BSN) over five years (U.S. Department of Veterans Affairs, 2013).

More broadly, many two- and four-year colleges have reached nursing articulation agreements between associate degree in nursing (ADN) and BSN programs to assist community college students in transferring from an associate degree program to a bachelor's degree program, with the goal of ultimately earning a BSN. The objective of these agreements is to expand the scope of the pipeline into BSN programs. However, while intuitively appealing, there is little empirical evidence on how articulation agreements between two- and four-year colleges affect the nursing pipeline.

In this study, I examine the effects of ADN-to-BSN nursing articulation agreements between the University of Missouri-Columbia (MU) and multiple public community colleges in Missouri. I estimate the effects of these agreements on the likelihood that community college students transfer to MU, and to the MU four-year nursing degree program specifically. Using administrative student records provided by the Missouri Department of Higher Education (DHE), I build a nine-year data panel tracking first-time students entering one of the 14 public community colleges in Missouri between 2006 and 2014. To estimate the causal effects of these agreements, I use difference-in-differences specifications where identification relies on variation in the availability of nursing articulation agreements with MU across community colleges and over time. I also supplement the difference-in-differences models with a triple-difference design that further leverages variation in transfer rates across initial majors in community colleges.

I find no statistical evidence that nursing articulation agreements increase student transfer rates. Moreover, my insignificant point estimates are small in magnitude. For example, the point estimate from my preferred specification for the effect of having an articulation agreement in place implies an increase in the likelihood that students transfer

to the MU nursing program by just 0.01 percentage points. Even taking this insignificant estimate at face value indicates a very small effect on the nursing pipeline. It suggests that the number of students transferring to the MU nursing program across all community colleges that entered into articulation agreements with MU increased by only about 7 students in total between 2010 and 2014. Moreover, although my estimates are somewhat imprecise, I can still rule out a positive effect larger than 50 students in total transferring to the MU nursing program during this period.

2.2. Background

2.2.1. Nursing Shortages

The American Nurses Association reports that the U.S. needs 1.1 million new RNs to fill newly created positions and to replace retiring nurses by 2022.[4] Moreover, Torpey (2018) shows that job openings for RNs are projected to increase by just over 200,000 each year, on average, from 2016 to 2026, corresponding to a cumulative increase of over two million during this period. There are many factors that contribute to the persistently strong demand for nurses. First, as the baby boomer generation continues to reach retirement age, the demand for caregivers, such as nurses, is anticipated to rise. Second, the growing use of preventative and community-based care may also increase the demand for nurses. Third, incumbent nurses are aging and a large number of them are expected to retire in coming years. According to the fact sheet of the AACN (2019), over half of RNs are at least 50 years old and one million RNs will retire between 2017 and 2030. The retirement of experienced RNs may affect the quality of patient care and induce healthcare providers to demand more RNs.

[4] Retrieved from https://www.nursingworld.org/practice-policy/workforce/

U.S. postsecondary nursing programs are struggling to produce enough nurses to meet the demand. In addition to the need for a generally larger pipeline of undergraduate students interested in nursing, existing four-year nursing degree programs are facing challenges in their expansion efforts. According to the report of the AACN (2019), more than 75,000 qualified applicants were rejected by nursing programs across the country in 2018. Two main factors could lead to constrained supply. First, nursing faculty are also aging and retiring, and nursing graduate programs are not producing enough graduate-degree-prepared nurses to replace them as faculty. This makes it difficult to build large and robust nursing degree programs, which in turn limits enrollment in nursing programs. Second, nursing programs must restrict the number of students admitted because they do not have enough clinical sites and classrooms. Put differently, nursing education is expensive and program budgets limit expansion efforts. These constraints on supply could reduce the ability of ADN-to-BSN nursing articulation agreements to increase the supply of nurses with a BSN. That is, if nursing programs cannot accept the students who are already qualified and interested in enrolling, increasing the number of qualified and interested students via articulation agreements is unlikely to increase the nursing supply.

2.2.2. Bachelor of Science in Nursing versus Associate Degree in Nursing

A person can pursue a career in nursing through different routes. The two main degree qualifications for entering the profession are an associate degree in nursing (ADN) and a Bachelor of Science in nursing (BSN). Individuals must pass a national licensure exam, the NCLEX-RN, to be licensed as a registered nurse. Although both degrees can satisfy the degree qualification at entry, BSN-level nurses are considered to be more qualified than ADN-level nurses. This is because BSN degree programs are designed to

provide students with more complete training, including not only the basics of nursing and clinical skills but also communication skills and leadership development. Several descriptive studies show that highly trained BSN-level nurses are able to provide better patient care, are associated with lower mortality and failure-to-rescue rates, and give more proficient diagnoses (AACN, 2019).

Moreover, nurses with a BSN have a better job outlook than those with an ADN. First, BSN-level nurses usually have more jobs to choose from. This is because a BSN is necessary for some advanced positions in nursing, such as nursing specialties and leadership positions, which require nurses to have broader duties. In addition, following the IOM's recommendation, a BSN may become the new entry-level degree required by many health care employers by 2020. Thus, nurses with a BSN will have more job opportunities. Second, BSN-level nurses working in high-level positions make more money than ADN-level nurses, even though the starting salaries for new BSN- and ADN-level nurses are not significantly different.

2.2.3. Articulation Agreements and Transfers

Community colleges play an important role in providing students with affordable, flexible, and easy access to higher education (Rouse, 1995). Community college students have an opportunity to transfer to a four-year university and earn a bachelor's degree; however, transfer rates and bachelor's degree completion rates are low. According to the National Student Clearinghouse's Transfer Tracking report (2017), of first-time community college students in 2010, 31.5 percent transferred to a four-year college and 13 percent attained a bachelor's degree within six years of entry, respectively. The low rates may be attributed to many factors. One factor is that community college students are much

less academically qualified on average. Another is the complex transfer process in which students take time to navigate transfer requirements and course equivalency of different institutions, which could lead to the low transfer rates. A third and related issue is that some courses needed to earn an associate degree from community colleges may not be accepted by four-year colleges. This makes the transfer process even more complicated and burdensome.

Articulation agreements between two- and four-year institutions have been developed to simplify the transfer process and make it more transparent for students. The National Association for College Admission Counseling defines articulation agreements as "Formal arrangements between two or more colleges and universities that specify how courses, a general education plan, and/or major requirements transfer from one institution of higher education to another."[5] Many states have developed statewide articulation agreements, some of which mandate that the completion of a core set of courses can transfer among all public institutions of higher education within a state and satisfy the general education requirements of receiving institutions. Moreover, independent articulation agreements have been established at the academic program level among institutions. Both statewide agreements and program-level agreements between two-year and four-year colleges assist community college students in transferring to four-year colleges and completing specific degrees.

Much of the extant research examines the effects of statewide articulation agreements on student outcomes and most studies do not find statistically significant effects (Anderson, Sun, and Alfonso, 2006; Gross and Goldhaber, 2009; Roksa and Keith,

[5] Retrieved from https://www.nacacnet.org/knowledge-center/transfer/the-transfer-process-defined/

2008). For example, a descriptive analysis by Anderson, Sun, and Alfonso (2006) uses the Beginning Postsecondary Student Longitudinal Study (BPS89) and shows that statewide articulation agreements do not increase the likelihood that students transfer from two- to four-year colleges. Similarly, Gross and Goldhaber (2009) find no effects of articulation agreements on student transfers and bachelor's degree attainment, but they do find small and positive effects on the transfer rates of Hispanic students. One exception is a study by Boatman and Soliz (2018), who find suggestive evidence of positive effects on transfer rates. Specifically, they use Ohio's longitudinal student-level data to examine the effects of a statewide articulation agreement that requires each public institution to build a set of general education courses known as a transfer module, which can be transferred among Ohio's public institutions. The study finds that completing the transfer module increases the likelihood that students transfer from two- to four-year colleges and attain associate degrees.

The literature on statewide articulation transfer policies is thin, but there is even less research on sub-state agreements. A recent study by Baker (2016) examines the effects of one such agreement known as the associate degrees for transfer (ADT) program, which facilitates transfers from California community colleges to the California State University (CSU) system. Baker identifies the effects of the ADT program using a triple-difference model and finds that attainment of associate degrees rises in departments offering the ADT, but the policy does not increase the number of students transferring from community colleges to the CSU system, although there is suggestive evidence of positive impacts in later years.

The current analysis contributes to the limited literature on articulation agreements by investigating a narrower articulation program focused on nursing specifically. The focus on nursing is of interest due to the aforementioned strong labor-market demand for individuals with this training. Whereas many of the previously studied articulation agreements can be described as attempting to increase access to four-year colleges broadly, the nursing articulation agreements I study focus on increasing access to a specific type of training in a field characterized by strong labor demand and high salaries. The individual articulation agreements are between MU, the state flagship university in Missouri, and multiple community colleges. I use difference-in-differences specifications that leverage variation in the availability of nursing articulation agreements with MU across community colleges and over time to identify the effects of these agreements. Moreover, I supplement these models with a triple-difference design that additionally leverages variation in transfer rates across initial majors in community colleges.

2.3. Methodology

My main model is a difference-in-differences regression with multiple periods as well as multiple treatment and comparison groups:

$$Y_{ist} = \alpha + \beta Policy_{st} + X_{ist}\gamma + \delta_s + \lambda_t + \epsilon_{ist} \quad (3)$$

In Equation (3), Y_{ist} is an indicator variable measuring a transfer outcome for student *i* who initially enrolled in community college *s* in year *t*. I consider two transfer outcomes: whether a student transferred to MU, and to the MU nursing program, both within three years of entry into a community college. X_{ist} is a vector of observed student variables including the student's race/ethnicity, gender, ACT math and English scores, high school percentile ranks, and missing-value indicators to account for missing ACT scores and high

school percentile ranks. Male and white students are the omitted gender and racial/ethnic categories. $Policy_{st}$ is the treatment variable, set equal to one if community college *s* had a nursing agreement with MU in year *t* and zero otherwise. δ_s and λ_t are community college and year fixed effects, respectively, and ϵ_{ist} is the error term. Error bands are estimated using the wild-cluster bootstrap with clustering at the community college level. I use the wild-cluster bootstrap because typical clustered errors estimated using few clusters can cause over-rejection of the null hypothesis (Cameron, Gelbach, and Miller, 2008). I also consider Goodman-Bacon's (2018) point that the estimate from a difference-in-differences setting with variation in treatment timing is a weighted average of all possible canonical difference-in-differences estimates and units treated in the middle of the data panel have higher weights. In my application, the treatment timing for community colleges is concentrated in the middle of my study period and this reduces his concern about bias in difference-in-differences estimates from time-varying heterogeneity.

The key identifying assumption of the difference-in-differences method is parallel trends. Specifically, the trend for the control group is taken as an accurate counterfactual for the treatment group if the treatment had not occurred. While this assumption is not directly testable, a common piece of evidence consistent with parallel trends is that trends in the outcome variables are the same for both treated and control groups prior to the treatment. To test this in my setting with staggered articulation agreement adoptions, I estimate an event-time model with leads and lags of the treatment:

$$\begin{aligned} Y_{ist} = \alpha + \beta_{-4}Policy_{st}^{-4} + \beta_{-3}Policy_{st}^{-3} + \beta_{-2}Policy_{st}^{-2} + \beta_0 Policy_{st}^{0} \\ + \beta_1 Policy_{st}^{1} + \beta_2 Policy_{st}^{2} + X_{ist}\gamma + \delta_s + \lambda_t + \epsilon_{ist} \end{aligned} \tag{4}$$

In Equation (4), $Policy_{st}^{0}$ is an indicator variable set to one for the year in which a nursing articulation agreement was signed between MU and community college s. $Policy_{st}^{1}$ is an indicator variable set to one for one year after the agreement and $Policy_{st}^{2}$ is an indicator variable set to one for two or more years after the agreement. The leads of the treatment are defined similarly. For example, $Policy_{st}^{-4}$ is an indicator variable for four or more years before the agreement, $Policy_{st}^{-3}$ is an indicator variable for three years before the agreement, etc. I omit $Policy_{st}^{-1}$, which is an indicator variable for one year before the agreement, to serve as a baseline. The coefficients of the event-time indicator variables in Equation (4) measure the treatment effects in the given year relative to the base year. For example, β_{-2} is the difference-in-differences parameter measuring the "treatment effect" two years before the agreement, relative to the base year. If the parallel trends assumption is satisfied, then β_{-4} , β_{-3} , and β_{-2} should be statistically insignificant (i.e., indistinguishable from each other and the omitted baseline). Although recent studies have argued that the approach of using the event-time model above to test parallel trends assumption is imperfect (Freyaldenhoven, Hansen, and Shapiro, 2019; Roth, 2018), it may not imply that detecting no pre-trends using this commonly used method is meaningless (Abraham and Sun, 2018; Goodman-Bacon, 2018). In addition, my estimated upper bound of the 95 percent confidence interval even implies small effects given my somewhat imprecise point estimate. This somehow reduces Roth's (2018) concern about bias resulting from parallel trends tests conditional on detecting no pre-trends in outcome variables.

To supplement the difference-in-differences research design, I use triple-difference specifications that further leverage variation in transfer rates across initial majors in

community colleges. I identify students who initially enrolled in a program in liberal arts and science, general studies, or humanities (hereafter: general studies) in community colleges as the treated group in these models. It may seem counterintuitive at first that I do not identify students who declared nursing majors at entry into community colleges as the treated group for the triple-difference models. However, empirically, students who major in general studies are much more likely to transfer to a four-year institution (Baum and Holzer, 2017) and these students are more likely to be affected by changes in transfer policies. Moreover, in the bottom panel of Table 2.2, I show that during the period (2006-2009) prior to any nursing articulation agreements being signed in Missouri, over 90 percent of MU nursing transfers initially declared general studies majors in community colleges. Furthermore, later (in Tables 2.8 and 2.9 below) I show that nursing articulation agreements do not induce students to choose nursing as their initial community college majors. All of this evidence is consistent with my choice of general studies as the treated majors.

Following Baker (2016), the triple-difference regression model is written as:

$$\begin{aligned} Y_{iskt} = \alpha + \beta Policy_{st} * Major_k + X_{iskt}\gamma + \lambda_t \\ + \pi_k + \delta_s + \varphi_{kt} + \theta_{ks} + \sigma_{st} + \epsilon_{iskt} \end{aligned} \tag{5}$$

In Equation (5), Y_{iskt} is a transfer outcome for student *i* in community college *s* with initial major k in year t.[6] $Policy_{st}$ is the treatment variable, which equals one if community college *s* had a nursing agreement with MU in year *t* and zero otherwise. $Major_k$ is an indicator variable set to one if major k is general studies and zero otherwise. φ_{kt}, θ_{ks},

[6] As shown in Table 2.2, the majority of MU transfers initially declared general studies as their majors in community colleges. Given this empirical evidence, I categorize initial community college majors into general studies and non-general studies.

and σ_{st} are major-by-year, major-by-community college, and community college-by-year fixed effects, respectively. λ_t, π_k, and δ_s are year, major, and community college fixed effects, respectively. This model controls for potential confounders including community college-specific time-varying changes, time-varying major-specific effects, and community college-specific major-specific effects (Baker, 2016). The triple-difference parameter is β, which captures the effects of nursing agreements on student transfer outcomes in treated majors, treated colleges, and treated years.

2.4. Data

My analysis is based on student-level administrative records provided by the Missouri Department of Higher Education (DHE). The data include student-level information, such as gender, race/ethnicity, ACT math and English scores, and high school percentile ranks. I build a nine-year data panel tracking first-time students entering one of the 14 public community colleges in Missouri between 2006 and 2014. The DHE data allow me to access each student's term registration information and track the student's transfer status within three years of entry. With the information about the school the student attended and her declared major(s), I define two transfer outcomes: whether the student transferred to MU at all and whether the student transferred to the MU baccalaureate nursing program. Majors are identified by CIP (Classification of Instructional Programs) codes developed by the U.S. Department of Education's National Center for Education Statistics (NCES).

Figure 2.1 shows trends in the share of students transferring to MU and the share of students transferring to the MU nursing program for the community college cohorts in my sample from 2006 to 2014. The share of MU transfers trended downward, decreasing

by about 32 percent from 2006 to 2014 overall. Similarly, the share of MU nursing transfers decreased by 50 percent, although this is hard to see in the figure because the base is so much smaller. Note that trends in the number of total transfers to MU and the number of total transfers to the MU nursing program moved similarly as the share of MU transfers and the share of MU nursing transfers, respectively.

Descriptive statistics for the data are reported in Table 2.1. My analytical sample includes 210,417 first-time students entering a community college between 2006 and 2014. Within three years of entry, approximately 1.6 percent of entrants transferred to MU and 0.1 percent transferred to the MU nursing program. Female students make up about 54 percent of the sample. About 71 percent of the sample is white, 14 percent is black, 3 percent is Hispanic, 2 percent is Asian/Pacific Islander, and 11 percent is some other race. Compared to the racial composition of the entire sample, black and Hispanic students are underrepresented among students who transferred to MU. For example, only about 3 percent of MU transfers are black students. In addition, the average ACT math and English scores and the average high school percentile rank are higher for MU transfers.

I retrieved the nursing articulation agreements between MU and the public community colleges from the MU office of admissions website. The purpose of these agreements is to make transfers from an associate degree program to the MU baccalaureate nursing program easier and provide students who previously attained an ADN with a pathway to earn a BSN at the MU Sinclair School of Nursing. These agreements state conditions of transfer, describe transfer of credit, and list program plans, which provide detailed course requirements in community colleges and at MU. The academic years in which the agreements became effective and the signatures of chief executive officers of

both institutions are included. There are 14 public community colleges in Missouri. Two colleges signed a nursing agreement in the academic year 2010-11, six in 2011-12, and two in 2012-13.[7] If a community college signed a nursing articulation agreement with MU in the academic year 2010-11, I treat the 2010 cohort of entering community college students as the first treated cohort.

As noted above, I identify the treated majors in the triple-difference specifications by examining the pre-treatment period distribution of initial community college majors for students who transferred to MU. The top panel of Table 2.2 shows that 1,692 students transferred to MU and 86 percent of these students initially declared general studies as their majors in community colleges. The second panel of Table 2.2 isolates the 538 students who transferred to an MU program with an articulation agreement, regardless of fields. Similarly, about 85 percent of these students initially majored in general studies in community colleges. In addition, the bottom panel of Table 2.2 presents that about 91 percent of students who transferred to the MU nursing program chose general studies as their initial community college majors.

2.5. Results

2.5.1. Primary Findings

I use linear probability models to estimate the difference-in-differences specification shown in Equation (3). Regression results are reported in Table 2.3. The dependent variable in Column (1) is whether a community college student transferred to

[7] Jefferson College, North Central Missouri College, St. Louis Community College, and State Technical College of Missouri did not enter into a nursing articulation agreement with MU by 2014 (the end of my data panel).

MU, and in Column (2) it is whether a student transferred to the MU nursing program. Both regressions include community college and year fixed effects and individual controls.

Table 2.3 reports point estimates, bootstrapped 95 percent confidence intervals, and bootstrapped p-values. The point estimates are positive, small in magnitude, and statistically insignificant. The insignificant estimates in Columns (1) and (2) of Table 2.3, taken at face value, imply that a nursing articulation agreement increases the likelihood that students transfer to MU and the MU nursing program by only about 0.07 and 0.01 percentage points, respectively. These percentages correspond to increases in the number of transfers of about 49 and 7 students in total, respectively, over the five year period between 2010 and 2014.[8] Thus, the estimates are not only statistically insignificant, but they imply small effects. Additionally, Table 2.3 shows that the upper bound of the 95 percent confidence interval for the coefficient on transferring to the MU nursing program is 0.07 percentage points. Hence, although the estimate is somewhat imprecise, I can still rule out a positive effect larger than 50 students transferring to the MU nursing program during this five-year period. This effect is still small in terms of effectively expanding the nursing supply even though it is comparable to the total of MU nursing transfers among pre-treatment period cohorts, as noted in the bottom panel of Table 2.2.

Next, I estimate the event-time model shown in Equation (4) to test the parallel trends assumption. As noted above, the indicator variable for one year before the agreement is omitted and serves as a baseline. If the coefficients on $Policy_{st}^{-4}$, $Policy_{st}^{-3}$, and $Policy_{st}^{-2}$ are jointly equal to zero, this would be consistent with the parallel trends

[8] Increases in the number of students transferring to MU and the MU nursing program are calculated based on the formula: $\left[\sum_{j=1}^{J} y_j * enroll_j\right] * \hat{\beta}$, where J is the total number of treated colleges, y_j and $enroll_j$ are the number of treated years and the average annual enrollment in treated years for college j, respectively, and $\hat{\beta}$ is the corresponding point estimate in Table 2.3.

assumption being upheld. Figures 2.2 and 2.3 present the coefficients of these event-time indicator variables and corresponding 95 percent confidence intervals from the models of the two different transfer outcomes. As shown in the figures, the coefficients on all leads of the treatment are close to zero in magnitude and statistically insignificant.

2.5.2. Heterogeneity

Full-time students who register for 12 or more credits in their first semester in community colleges are more likely to transfer to four-year colleges. In the sample, the majority of students who transferred to public four-year colleges in Missouri were full-time enrollees using the 12-credit definition, while they first enrolled in community colleges. Specifically, about 75 percent of students who ultimately transferred to a four-year college in Missouri were full-time enrollees in community colleges and about 73 percent of MU transfers were full-time students in community colleges. Correspondingly, full-time students might be affected by changes in transfer policies. Next, I restrict the full sample to only full-time students and examine the possible effects of nursing agreements on these students. The sample restriction is used to focus on students interested in transferring and exclude a case where for example, an elderly person takes only one community college course.

Regression results are shown in Table 2.4. There is some statistical evidence that a nursing agreement increases the likelihood that full-time students transfer to MU and the MU nursing program. The coefficients for both outcomes are statistically significant at the 10 percent level, but imply small effects. Specifically, a nursing agreement increases the likelihood that full-time students transfer to the MU nursing program by just about 0.03 percentage points. This change corresponds to an increase in the number of full-time

community college students transferring to the MU nursing program, across all community colleges that entered articulation agreements with MU, of about 15 students in total between 2010 and 2014. The emergence of the significant estimates may be due to the fact that full-time students are more likely to be affected by nursing agreements. However, I am cautious to draw strong inference from these results because the point estimates are only weakly significant and small in magnitude. Also, as explained in the next section, they are not robust to the triple-difference specifications.

2.5.3. Triple-difference Specifications

Many community college students declare general studies as their initial majors (Baum and Holzer, 2017). These fields are designed for students expecting to transfer to four-year institutions (Jaschik, 2015). Correspondingly, Table 2.2 shows that the majority of students who transferred to MU initially chose these majors in community colleges. Given this empirical evidence, community college students in these fields might be more responsive to changes in transfer policies, such as articulation agreements. This is true even for nursing. Despite the fact that the articulation agreements I study formally link ADN and BSN programs, empirically it is clear that like other community college transfer students, transfer students to the MU nursing program primarily chose general studies as their beginning community college majors. Therefore, I regard general studies as the treated majors in the triple-difference specifications, which supplement the difference-in-difference specifications and further leverage variation in transfer rates across initial majors in community colleges for identification.

I first use the full sample to estimate the model specified in Equation (5), which provides full controls for major-by-year, major-by-community college, and community

college-by-year fixed effects. The dependent variables are the same as those in Table 2.3 and regression results are reported in Table 2.5. I find no evidence that nursing agreements increase student transfer rates in the triple-difference specifications. The estimated effects on transfers are small, close to zero, statistically insignificant, and similar to those in Table 2.3. When I re-estimate Equation (5) using only full-time students, as noted above and shown in Table 2.6, the marginally significant effects found for full-time students using the difference-in-differences specifications are not robust to the triple-difference specifications.

2.6. Extensions

2.6.1. Transfers to Other Public Four-year Institutions

There are 13 public four-year institutions in Missouri. Thus far, the focus has been on examining the effects of nursing articulation agreements between MU and community colleges on the likelihood that community college students transfer to the MU nursing program. However, there is an omitted-variable bias concern. Specifically, it is possible that other four-year colleges made nursing articulation agreements with community colleges around the same time as MU. Unfortunately, comprehensive data on program-level articulation agreements involving other four-year institutions are unavailable to investigate this hypothesis directly. However, if there were concurrent treatments with other four-year colleges in the state, some students may have transferred to the other programs instead of MU. This would cause negative bias in the estimates above.

To explore this issue further, I first examine trends in transfers to other public four-year colleges in Missouri descriptively, and then test indirectly for the possibility of bias from potentially unobserved treatments. Figure 2.4 shows trends in the share of students

transferring to non-MU public four-year institutions and the share of students transferring to non-MU public four-year degree programs in nursing for community college cohorts during my data panel. Trends in the share of MU transfers and the share of MU nursing transfers are also shown in the figure for comparison. Panel (a) is for all transfers and panel (b) is for transfers to nursing programs. Generally, the curves in each panel moved similarly. However, panel (a) shows that in 2013, there was an uptick in total transfers statewide, but not at MU. Additionally, as presented in panel (b), there was slight divergence in the trends in nursing transfers between 2010 and 2011.[9] Moreover, due to the smaller magnitude of the share of MU nursing transfers, it is not obvious from panel (b) that the changes in the share of MU nursing transfers between 2006 and 2008 were similar to the share of non-MU nursing transfers.

Next, I test indirectly for the possibility of omitted-variable bias by examining the effects of MU nursing agreements on the likelihood that community college students transfer to nursing programs at *other* public four-year institutions in Missouri. I estimate the model specified in Equation (3); however, the dependent variables are transfers to a non-MU four-year institution and transfers to a non-MU four-year degree program in nursing, respectively. Table 2.7 shows that the coefficients on the MU nursing agreement treatment are negative and statistically insignificant in both models. Thus, there is no evidence that my weak findings reflect increases in transfers from community colleges to other four-year public universities in Missouri.

2.6.2. Effects on Declaring Nursing Majors

[9] It is unclear what is driving the divergence in the trends in total transfers and nursing transfers. However, in results omitted for brevity, I confirm that my primary findings are not qualitatively different if I exclude the 2013 and 2014 cohorts, and the 2010 and 2011 cohorts, respectively, from my sample (results omitted for brevity).

Based on the descriptive output in Table 2.2, I define community college students who initially enrolled in a program in general studies, rather than in nursing, as the group treated by nursing agreements in the triple-difference specifications. This makes sense given that very few transfers to the MU nursing program initially enrolled in a community college nursing program based on pre-treatment data. However, it is possible that the implementation of a nursing articulation agreement could change the initial majors for entering community college students. That is, it might induce more students to declare nursing as their initial majors in community colleges. In turn, this would weaken the case for using general studies as the "treated majors" in the triple-difference specifications.

To explore this possibility empirically, I test for an effect of nursing agreements on whether students choose nursing as their initial majors in community colleges. I re-estimate Equation (3) using the full sample and only full-time students, respectively. As shown in Tables 2.8 and 2.9, the coefficients are negative and statistically insignificant. There is no evidence that nursing agreements increase students' likelihood to initially declare nursing majors in community colleges.

2.6.3. Effects on ADN Attainment

The nursing articulation agreements between MU and the community colleges formally govern transfers from an associate degree program to the MU baccalaureate nursing program and provide students who previously obtained an ADN with a gateway to earn a BSN at MU. Although the agreements do not affect whether students initially enroll in community college nursing programs per Tables 2.8 and 2.9, they could affect transfers to ADN programs in community colleges. To test this possibility, I estimate models where the outcome variable is the intermediary outcome of an associate degree in nursing. A

positive effect would imply that these agreements encourage students to earn an associate degree in nursing due to the transfer incentive. The implication of such a result, combined with my null findings for transfers to the MU nursing program, would be that program constraints limit the ability of articulation agreements to expand the nursing supply. As discussed above in Section 2.1, nursing programs nationally are turning away many interested applicants as they struggle with capacity issues.

I examine the relationship between nursing agreements and attainment of an associate degree in nursing by estimating Equation (3) using the full sample. The dependent variable is whether a student attained an associate degree in nursing within three years of entry into a community college. As presented in Table 2.10, the point estimate is positive but statistically insignificant. Thus, there is no evidence of such relationship. I also test whether the parallel trends assumption holds for this outcome variable and as shown in Figure 2.5, the coefficients on all leads of the treatment are close to zero in magnitude and statistically insignificant. Table 2.11 shows similar results using only full-time students. There is no statistical evidence that nursing agreements affect attainment of associate degrees in nursing among this sample either.

Taken as a whole, I find no evidence that MU nursing agreements influence the likelihood that students earn an associate degree in nursing in community colleges. This indicates that supply constraints in the MU nursing program do not drive the null findings for transfer outcomes. Instead, these results are consistent with a lack of demand for nursing education among community college students spurred by the articulation agreements.

2.7. Conclusion

The American Association of Colleges of Nursing expects the U.S. to experience excess demand for registered nurses. Moreover, labor shortages for baccalaureate-prepared nurses are of particular policy relevance as the current proportion of nurses with a Bachelor of Science degree is far below the recommendation of the Institute of Medicine. One policy solution that may help to address this problem is program-level nursing articulation agreements between two- and four-year colleges. These agreements are intuitively appealing options for expanding the nursing pipeline. However, little is known about their effects.

I study nursing articulation agreements between MU and multiple public community colleges in Missouri. I estimate the effect of these agreements on the likelihood that community college students transfer to MU, and more importantly, to the MU baccalaureate nursing program. I apply administrative data provided by the DHE to construct my data panel, which includes 210,417 first-time students entering public community colleges in Missouri between 2006 and 2014. I use difference-in-differences specifications that leverage variation in the availability of nursing articulation agreements with MU across community colleges and over time for identification. I also supplement the difference-in-differences models with a triple differences design that further leverages variation in transfer rates across initial majors in community colleges.

I find no statistical evidence that the nursing agreements at MU increase transfers to the MU nursing degree program. Additionally, my insignificant point estimates imply small effects even when taken at face value. For example, in my preferred specification, I estimate that nursing articulation agreements, on average, increase the likelihood that community college students transfer to the MU nursing program by just 0.01 percentage

points. As a result, continuing to take the insignificant point estimate at face value, the number of students transferring to the MU nursing program across all community colleges that entered articulation agreements with MU is estimated to have increased by only about 7 students in total between 2010 and 2014. Moreover, although my estimates are somewhat imprecise, I can rule out a positive effect larger than 50 students in total transferring to the MU nursing program during this period.

In an extension of my main analysis, I also show that nursing articulation agreements do not increase the likelihood that community college students attain an associate degree in nursing. This finding is important because it helps to distinguish the mechanism for my null results regarding student transfers. In particular, given that many nursing programs nationally are oversubscribed, one hypothesis for why the nursing articulation agreements I study have no effects is that the MU nursing program is at or beyond its capacity. One way this condition could manifest in my analysis, although not the only way, is if the articulation agreements expanded the production of associate degrees in nursing with no subsequent change in transfer rates.[10] However, the results indicate that the production of associate degrees in nursing is not affected. This implies that the nursing articulation agreements in Missouri have failed to stimulate the demand for nursing education at community colleges.

[10] My findings do not entirely rule out capacity constraints at MU as the explanation. For example, it could be that the MU nursing program is capacity constrained and community college students know this and are forward looking. In this case they may not expect to be admitted even if they adhere to the articulation agreement conditions. My empirical analysis is not sufficient to rule out this possibility, although recent evidence suggests that the quality of information students possess about many aspects of the education system is limited (Deming, Goldin, and Katz, 2013; Hoxby and Turner, 2013; Baker, Bettinger, Jacob, and Marinescu, 2018).

Figure 2.1. Trends in MU Transfers and MU Nursing Transfers for Community College Cohorts.

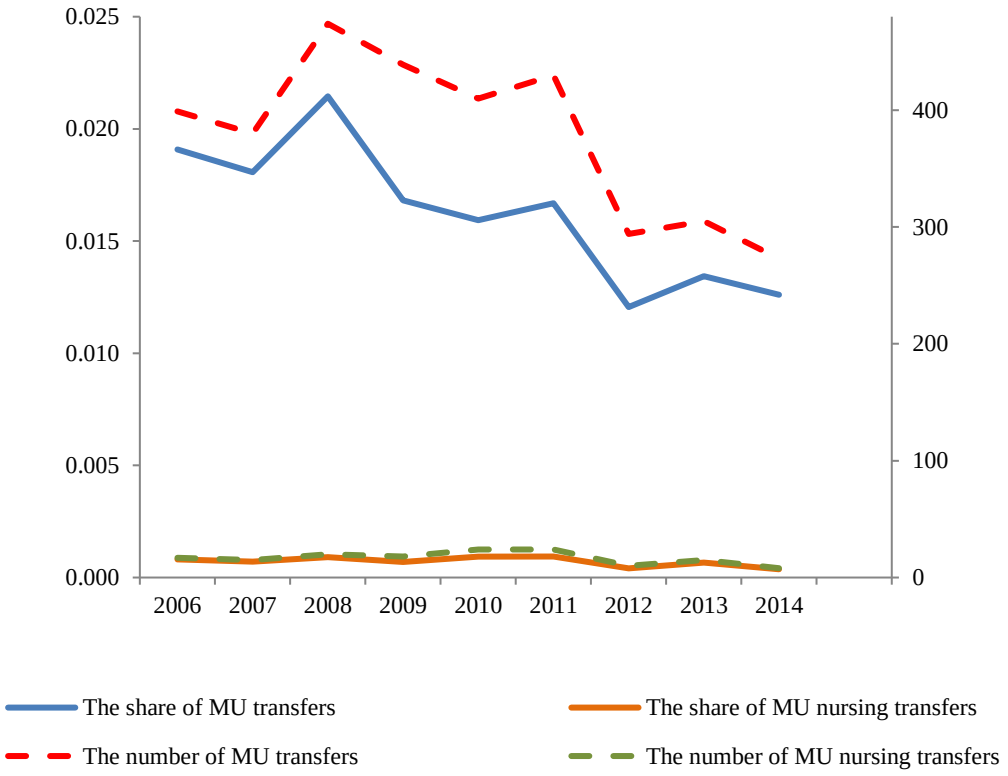
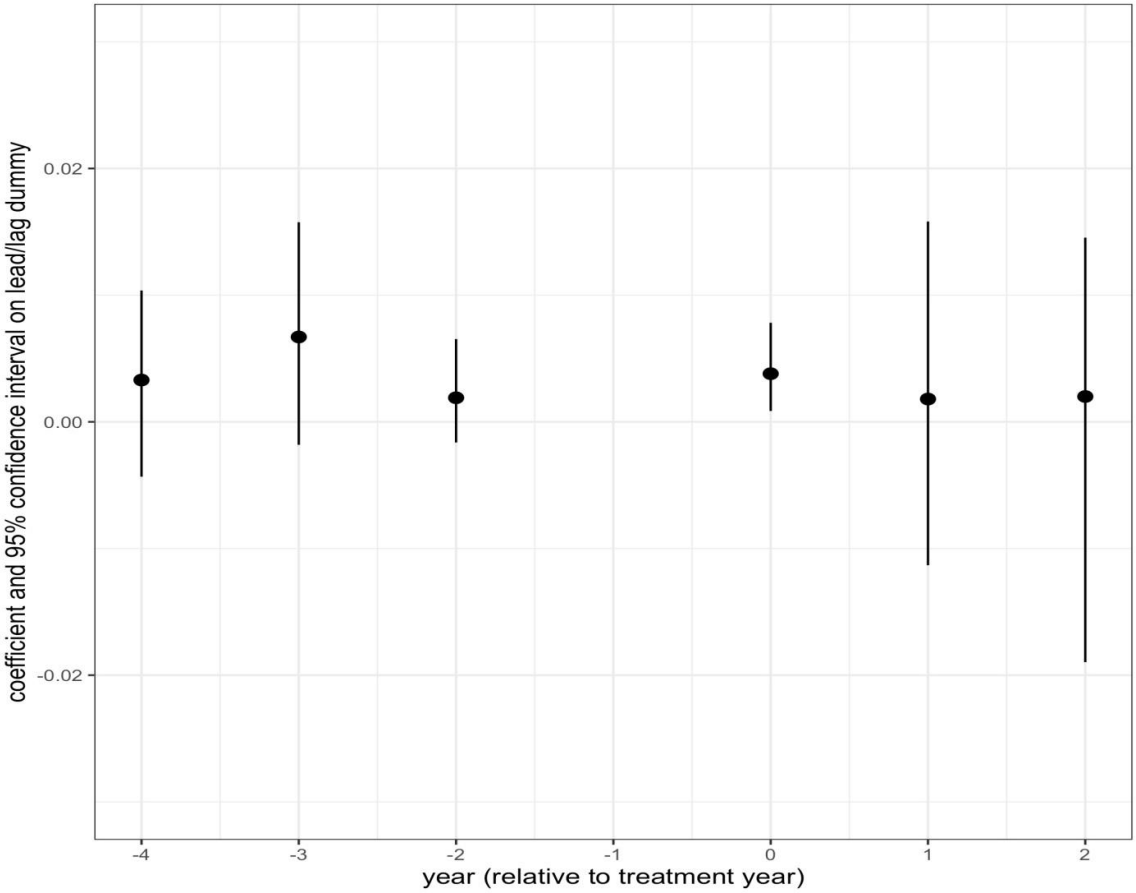
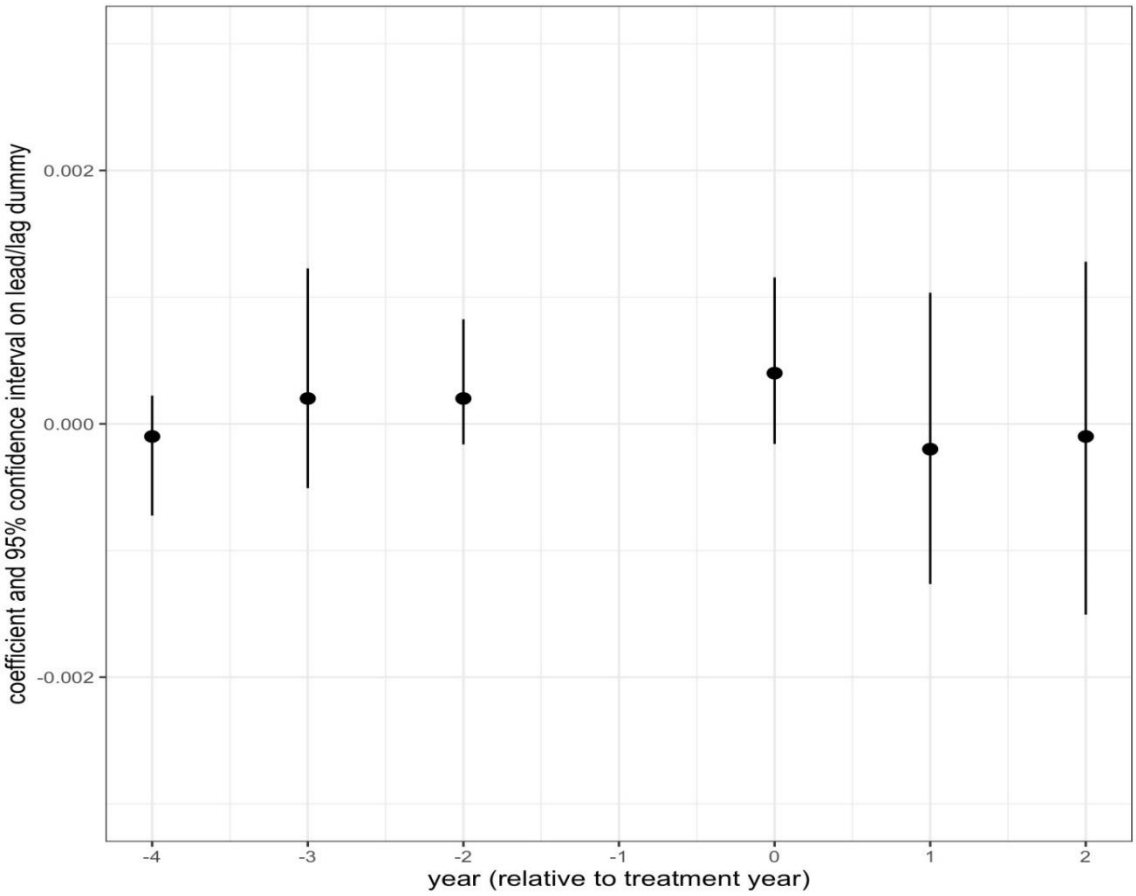


Figure 2.2. Parallel Trends Test for Transferring to MU.



Note: Black dots and vertical bars represent coefficients and 95 percent confidence intervals, respectively.

Figure 2.3. Parallel Trends Test for Transferring to the MU Nursing Program.



Note: Black dots and vertical bars represent coefficients and 95 percent confidence intervals, respectively.

Figure 2.4. Trends in the Share of Non-MU Transfers and the Share of Non-MU Nursing Transfers for Community College Cohorts.

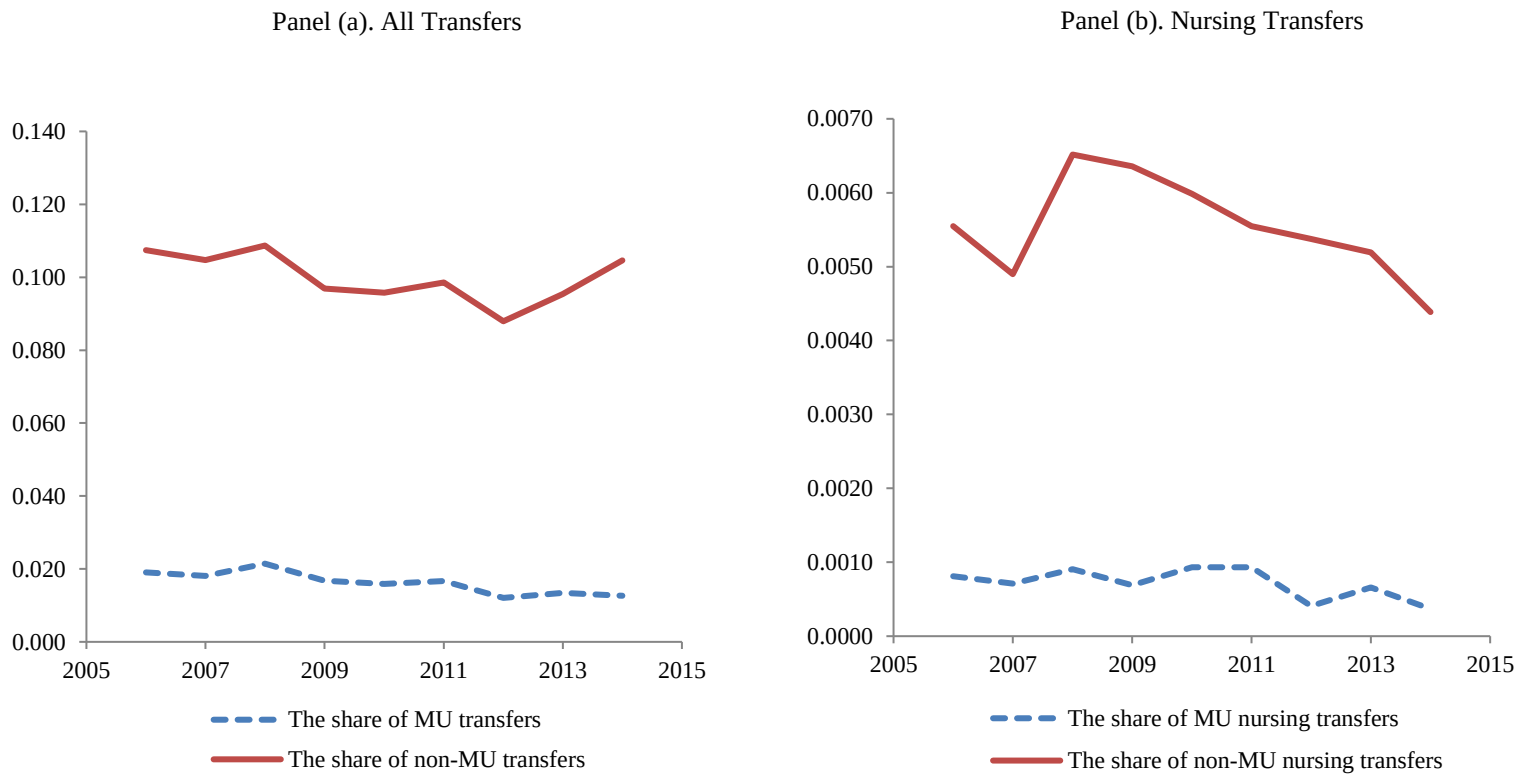
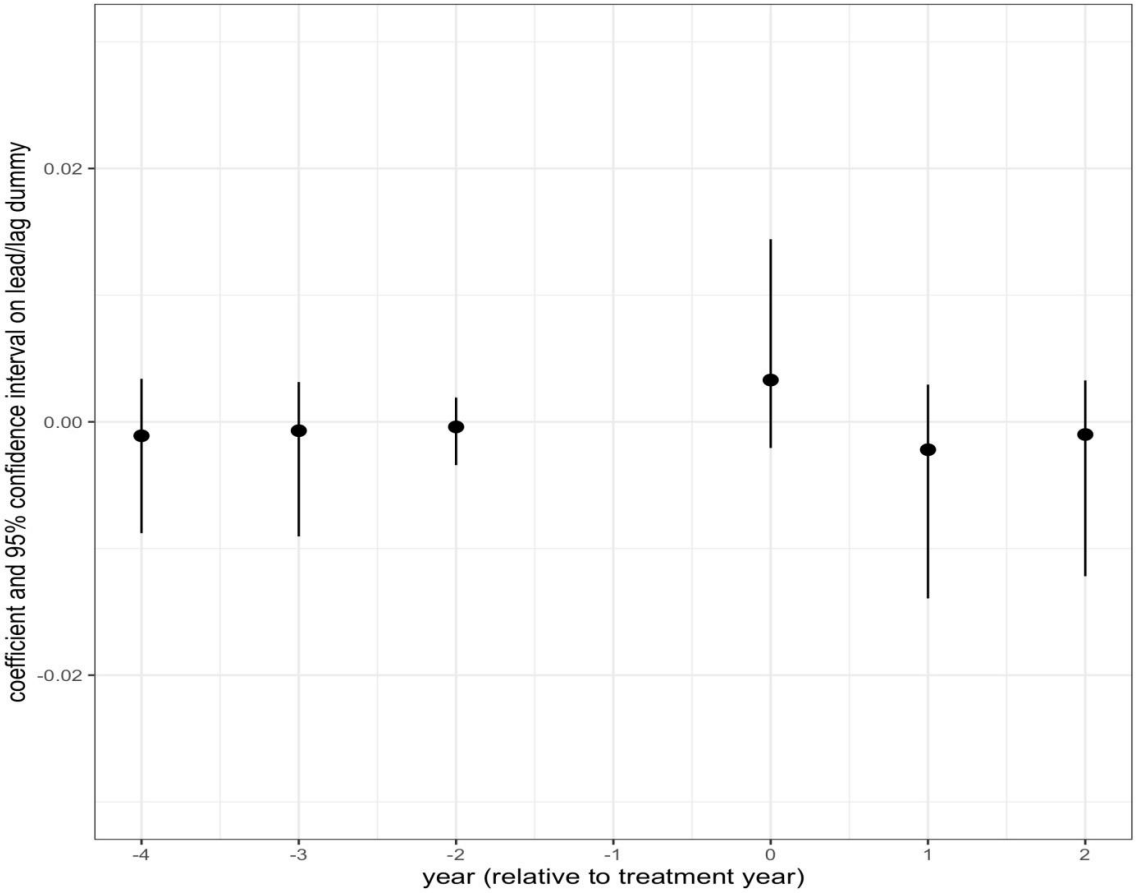


Figure 2.5. Parallel Trends Test for an Associate Degree in Nursing.



Note: Black dots and vertical bars represent coefficients and 95 percent confidence intervals, respectively.

Table 2.1. Descriptive Statistics.

Student Outcomes	Mean (stdev)	
Transfer to MU	0.016 (0.126)	
Transfer to the MU Nursing Program	0.001 (0.027)	
	All Students	MU Transfers
Student Characteristics	Mean (stdev)	Mean (stdev)
White	0.709 (0.454)	0.831 (0.374)
Black	0.138 (0.344)	0.027 (0.161)
Hispanic	0.027 (0.162)	0.017 (0.129)
Asian/Pacific Islander	0.015 (0.121)	0.019 (0.137)
Other Race	0.112 (0.315)	0.106 (0.308)
Female	0.538 (0.499)	0.463 (0.499)
ACT English	19.2 (3.0)	20.7 (3.6)
ACT English Missing Indicator	0.615 (0.487)	0.414 (0.493)
ACT Math	19.1 (2.4)	20.4 (3.2)
ACT Math Missing Indicator	0.616 (0.486)	0.414 (0.493)
High School Percentile Rank	49.1 (16.5)	55.7 (17.8)
High School Percentile Rank Missing Indicator	0.558 (0.497)	0.498 (0.500)
Number of Students	210,417	3,404

Table 2.2. The Pre-treatment Period (2006-2009) Distribution of Initial Community College Majors.

	Number of Students	% of Total
Students Transferring to Any Major at MU		
General Studies	1,458	86
Other Majors	234	14
Total	1,692	100
Students Transferring to Majors with an Articulation Agreement		
General Studies	458	85
Other Majors	80	15
Total	538	100
Students Transferring to the MU Nursing Program		
General Studies	64	91
Other Majors	6	9
Total	70	100

Table 2.3. Effects of Nursing Articulation Agreements on Transferring to MU and to the MU Nursing Program.

	(1)	(2)
	Transfer to MU	Transfer to the MU Nursing Program
Policy	0.0007	0.0001
	[-0.0083, 0.0094]	[-0.0004, 0.0007]
	p-value: 0.864	p-value: 0.705
Individual Controls	X	X
College FE	X	X
Year FE	X	X
N	210,417	210,417

Table 2.4. Effects of Nursing Articulation Agreements on Transferring to MU and to the MU Nursing Program, Only Full-time Students.

	(1)	(2)
	Transfer to MU	Transfer to the MU Nursing Program
Policy	0.0043*	0.0003*
	[-0.0008, 0.0096]	[-0.00002, 0.0009]
	p-value: 0.086	p-value: 0.070
Individual Controls	X	X
College FE	X	X
Year FE	X	X
N	142,820	142,820

Table 2.5. Effects of Nursing Articulation Agreements on Transferring to MU and to the MU Nursing Program, Triple-difference Specification.

	(1)	(2)
	Transfer to MU	Transfer to the MU Nursing Program
Policy	0.0009	-0.0002
	[-0.0081, 0.0100]	[-0.0009, 0.0006]
	p-value: 0.806	p-value: 0.459
Individual Controls	X	X
College FE X Year FE	X	X
Treated Major Indicator X Year FE	X	X
Treated Major Indicator X College FE	X	X
N	210,417	210,417

Table 2.6. Effects of Nursing Articulation Agreements on Transferring to MU and to the MU Nursing Program, Only Full-time Students and Triple-difference Specification.

	(1)	(2)
	Transfer to MU	Transfer to the MU Nursing Program
Policy	0.0030	-0.0003
	[-0.0043, 0.0109]	[-0.0013, 0.0005]
	p-value: 0.449	p-value: 0.535
Individual Controls	X	X
College FE X Year FE	X	X
Treated Major Indicator X Year FE	X	X
Treated Major Indicator X College FE	X	X
N	142,820	142,820

Table 2.7. Effects of MU Nursing Articulation Agreements on Transferring to Other MO Four-year Public Institutions and to Other Nursing Programs.

	(1)	(2)
	Transfer to Other Institutions	Transfer to Other Nursing Programs
Policy	-0.0057	-0.0013
	[-0.0405, 0.0257]	[-0.0032, 0.0009]
	p-value: 0.784	p-value: 0.170
Individual Controls	X	X
College FE	X	X
Year FE	X	X
N	210,417	210,417

Table 2.8. Effects of Nursing Articulation Agreements on Declaring Nursing Majors in Community Colleges.

	Nursing Majors in Community Colleges
Policy	-0.030
	[-0.1045, 0.0293]
	p-value: 0.437
Individual Controls	X
College FE	X
Year FE	X
N	210,417

Table 2.9. Effects of Nursing Articulation Agreements on Declaring Nursing Majors in Community Colleges, Only Full-time Students.

	Nursing Majors in Community Colleges
Policy	-0.028
	[-0.0885, 0.0284]
	p-value: 0.451
Individual Controls	X
College FE	X
Year FE	X
N	142,820

Table 2.10. Effects of Nursing Articulation Agreements on Earning an Associate Degree in Nursing.

	Associate Degree in Nursing
Policy	0.0012
	[-0.0028, 0.0069]
	p-value: 0.465
Individual Controls	X
College FE	X
Year FE	X
N	210,417

Table 2.11. Effects of Nursing Articulation Agreements on Earning an Associate Degree in Nursing, Only Full-time Students.

	Associate Degree in Nursing
Policy	0.0008
	[-0.0035, 0.0049]
	p-value: 0.627
Individual Controls	X
College FE	X
Year FE	X
N	142,820

Chapter 3

How Do Teachers from Alternative Pathways Contribute to the Teaching Workforce in Urban Areas? Evidence from Kansas City

3.1. Introduction

It is well-documented that urban school districts have difficulty recruiting and retaining high-quality teachers (Boyd et al., 2005; Boyd et al., 2006; Lankford, Loeb, and Wyckoff, 2002; Papay et al., 2017). Moreover, recent evidence suggests that accountability policies and improved measures of teaching effectiveness, which have increased the demand for and ability to identify effective teachers, respectively, have exacerbated staffing challenges for high-need schools (Bates, 2020; Cullen, Koedel, and Parsons, forthcoming). Although policy efforts in some states designed to combat these challenges have had some success, teacher recruitment in high-need, urban areas is an ongoing challenge (Glazerman et al., 2013; Springer, Swain, and Rodriguez, 2016; Swain, Rodriguez, and Springer, 2019).

Alternative teacher preparation programs (ATPPs) can be a source of labor supply in localized labor markets that face supply-side challenges. Indeed, many ATPPs explicitly build this idea into their mission statements. A well-known example is the national Teach for America (TFA) program, which we study here. In addition, regionally-based programs with similar goals include New York City Teaching Fellows, the Mississippi Teaching Corps, and Kansas City Teacher Residency (which we also study), among others. Compared to traditional university-based teacher preparation programs, which remain the predominant pipeline into the teaching profession nationally, ATPPs typically provide an

accelerated pathway into the classroom. A rationale is that rigid licensing requirements create barriers to entry that keep some qualified teachers out of the classroom (Sass, 2015). By reducing these barriers, ATPPs can increase the appeal and accessibility of the profession for a broader population of potential teachers.

ATPPs also offer pathways to teaching permanency (i.e., paths toward full licensure that would be required for a full career in teaching), although the structure of the pathways differs across programs. There are mixed views about whether ATPPs induce churn in the teaching profession, most notably with respect to TFA, but empirically the evidence suggests the high turnover rate of TFA teachers is not meaningfully different from the rate of other young teachers working in the same challenging environments (Donaldson and Johnson, 2011; Papay et al., 2017). Teacher residency programs, which are an increasingly common form of ATPP, typically include explicit supports to help promote teacher retention and some of these programs produce teachers with much higher retention rates than traditionally-trained teachers (e.g., see Papay et al., 2012).

In this paper, we examine how two ATPPs—TFA and Kansas City Teacher Residency (KCTR)—contribute to the local teacher labor market in and around Kansas City, Missouri. The city school district, Kansas City Public Schools (KCPS), is a high-poverty urban district with low achievement. Surrounding districts are more advantaged than KCPS, albeit marginally in some cases. There is also a large and vibrant charter sector in Kansas City, which is an interesting dimension along which to consider the role of ATPPs in serving the market.

We begin with a descriptive analysis of the TFA and KCTR teacher placements. We show that teachers from both programs are placed disproportionately in charter schools,

and more broadly, in schools with larger shares of low-income, low-performing, and underrepresented minority (Black and Hispanic) students. We also examine the diversity of the teachers themselves, motivated by the large minority enrollment share in Kansas City area schools and a rapidly evolving body of research pointing to the importance of demographic representation in the teaching workforce (e.g., Dee, 2005; Egalite and Kisida, 2017; Egalite, Kisida, and Winters, 2015; Holt and Gershenson, 2019; Lindsay and Hart, 2017; Papageorge, Gershenson, and Kang, 2020) . Relative to the larger local area, we find that both TFA and KCTR teachers are more racial-ethnically diverse than other teachers. However, only KCTR teachers are more racial-ethnically diverse than other teachers working in the same districts. TFA, KCTR, and the larger teaching workforce in the Kansas City area are all female-dominated—that said, TFA and KCTR are modestly diversity improving along the dimension of gender.

Next we examine the efficacy of TFA and KCTR teachers as estimated by value added to student achievement in math and English Language Arts (ELA) in grades 4-8. First, for TFA, we estimate that TFA teachers raise student test scores by 0.11 and 0.03 student standard deviations in math and ELA, respectively, compared to non-program teachers on average. These estimates contribute to a large literature on the efficacy of TFA teachers, but to the best of our knowledge are the first estimates from Kansas City. Our findings are consistent with previous evidence that TFA teachers are much more effective than other teachers in similar circumstances in terms of raising math achievement; and either similar to, or marginally more effective than, other teachers in terms of raising ELA achievement.[11]

[11] This description of the empirical literature on TFA value-added is broadly accurate, although several studies that have been conducted in New York City find null TFA results. For example, Decker, Mayer, and

We are not aware of any previous efficacy evidence for KCTR teachers, for whom our efficacy findings are similar to what we find for TFA. Specifically, we find that KCTR teachers increase student achievement by 0.15 and 0.05 student standard deviations in math and ELA, respectively, compared to non-program teachers on average. Despite strong interest in the teacher residency model among teacher educators (Guha, Hyler, and Darling-Hammond, 2016), our results for KCTR contribute to a very thin literature on the efficacy of teachers from residency programs in terms of their ability to improve student achievement. We are aware of just two previous points of empirical evidence. First, Papay et al. (2012) evaluate the Boston Teacher Residency and find negative impacts on student achievement in mathematics, although they find evidence of a positive performance trajectory among these teachers. The other efficacy evidence is from the Memphis Teacher Residency, which is evaluated as part of Tennessee's Report Card on the Effectiveness of Teacher Training Programs (Tennessee Higher Education Commission, 2014). The report presents mixed results for the Memphis Teacher Residency program, although overall the evidence is more positive than negative.

Taken on the whole, our analysis provides an area-level overview of how the TFA and KCTR programs contribute to the teacher labor market in Kansas City, Missouri. We show that these programs are being used to fill teaching needs in generally disadvantaged

---

Glazerman (2004) use a within-school randomized research design to study the effects on student achievement in math and ELA of TFA teachers and estimate that TFA teachers raise student achievement by about 0.15 student standard deviations in math relative to control teachers in their same schools. Backes et al. (2019) use value-added models and data from Miami-Dade County and find that TFA teachers outperform other teachers by about 0.10 student standard deviations in math. Xu, Hannaway, and Taylor (2011) study TFA effects on achievement in high school and find that TFA teachers increase math test scores by about 0.13 student standard deviations. Two studies using data from New York City find smaller-to-null TFA effects in math (Kane, Rockoff, and Staiger, 2008; Boyd et al., 2006). In terms of the effects in ELA, TFA value-added is smaller (Backes et al., 2019; Decker, Mayer, and Glazerman, 2004; Kane, Rockoff, and Staiger, 2008), although in high school, Xu, Hannaway, and Taylor (2011) find that TFA teachers have similar effects on math and English tests.

districts and schools, including charter schools. And at least as measured by achievement impacts, teachers from these programs are more effective than their non-program peers. In a final, supplementary analysis we examine teacher retention among teachers who enter the labor market via these programs compared to non-program teachers. Consistent with the findings from Papay et al. (2012) on the Boston Teacher Residency, we find that early-career retention for KCTR teachers is far above that of other teachers in the same districts. TFA teachers have higher retention after 1 and 2 years of service, but by year 5 are less likely to remain as teachers in the Kansas City area than other local-area teachers.

3.2. Brief Program Descriptions

3.2.1. TFA

Teach For America (TFA) recruits high-performing college students who commit to teach for two years in a low-income community where TFA has partnered with local school districts. Pre-placement TFA summer training varies by region but typically includes a 5-7 week accelerated training program, which includes teaching practice and coaching, and a 1-2 week regional induction and orientation program. TFA partners with local certification programs to help corps members pursue full teacher certification during their 2-year commitment period. Donaldson and Johnson (2011) find that the majority of TFA teachers continue to teach beyond the 2-year commitment, although the TFA exit rate increases significantly from the second to third year.

3.2.2. KCTR

KCTR is an urban teacher residency program operating in the Kansas City area. Residents are college graduates who train with a mentor, receive coaching, and enroll in a Master's program through the University of Missouri-Kansas City. KCTR participants earn

credit toward their master's degrees and teach four days a week for a full academic year in their mentor's classroom during the program. At the end of the residency year, residents become certified teachers in Missouri and agree to teach in a high-need school in Kansas City for three additional years. During the first post-residency year, program participants complete their Master's degrees, and they continue to receive instructional coaching throughout the three-year post-residency commitment.

3.3. Data

We received comprehensive lists of TFA and KCTR participants placed in Missouri schools from the programs themselves. The data include the year and school of each participant's initial placement after the training. Our TFA data cover seven cohorts who received training between fall-2011 and fall-2017 (inclusive). KCTR is a newer program and the first post-residency cohort was not placed until fall-2017; from KCTR we received program placements for the three cohorts that began their teaching placements in fall-2017, fall-2018, and fall-2019.[12] Hereafter, we refer to each school year by the spring year; e.g., 2017-18 as 2018.

We matched the listed participants to their employment records in administrative data provided by the Missouri Department of Elementary and Secondary Education (DESE). The DESE data provide additional information about the participants themselves, their placements, and their students. We were able to match all of the teachers on the program lists in the DESE data.

Table 3.1 shows the counts of program participants by the year of the first post-program placement, again noting that school years are denoted by the spring year. For TFA,

[12] To be more precise, we do not treat the during-residence year as a teaching placement. The first KCTR cohort finished the residency year in spring-2017 and was placed in teaching positions in fall-2017.

we use data for teachers who entered the program between 2012 and 2018 for our evaluation. A small number of TFA teachers entered the workforce with a lag, which is why Table 3.1 shows non-zero TFA placements in 2019 and 2020. KCTR's initial cohort went through residency during the 2017 school year and our analysis is based on program participants whose first post-residency years were in 2018, 2019 and 2020.

Table 3.1 also shows the numbers of program participants whose first placements were in teaching positions, who are the focus of our analysis. As expected, the vast majority of program participants were placed in teaching positions. Exceptions include a small number of individuals whose initial placements were not in standard teaching roles. For our descriptive analysis we analyze all teaching placements. For our teacher efficacy analysis based on value-added to student achievement, we use teachers of math and ELA in grades 4-8 (for which sample details are provided below). The value-added sample includes "self-contained" elementary teachers and subject-specific teachers in higher grades.

3.4. Methodology

3.4.1. Descriptive Labor Market Analysis

We begin by describing the composition of teachers and their initial teaching placements for each program compared to other public school teachers in the Kansas City area. As a first step, to define the "local labor market area" or "Kansas City area," we retrieved the address of the central office for each local education agency (LEA) operating over the span of our data from 2012-2020 (including six LEA's that were open in at least one of these years, but closed by 2020). Note that LEA's include both traditional school districts and charter school operators, where the LEA is defined at the level of the operator

for charter networks with more than one school in the area. For ease of presentation, we use the terms "LEA" and "district" interchangeably in the text.

We define the local labor market area as including all districts with a Kansas City, Missouri address. There are 30 such districts, including charter authorizers. We also include two additional districts with addresses in nearby Independence and Raytown (which are each about 8 miles from central Kansas City). In total, we define the area to include 32 districts, which combine to represent the region of effect for the programs we evaluate.[13]

Table 3.2 lists the 32 school districts, ordered from highest to lowest by the percentage of local-area non-program teachers employed, shown in the last column of the table. Noting that the vast majority of local-area teachers are non-program teachers, the ordering is essentially by district size. For each focal program, we report the percent of teachers in our sample from that program placed in each district. For ease of presentation, the data are aggregated for programs over relevant years in the 2012-2020 range (per Table 3.1).

The primary takeaway from Table 3.2 for KCTR and TFA is that they disproportionately place teachers in the central city school district, Kansas City Public Schools (KCPS). Over 60 percent of TFA teachers are placed in KCPS, whereas no other LEA has a double-digit share of TFA teachers. KCTR's representation in KCPS is also large—it accounts for about 27.6 percent of KCTR placements—but smaller than for TFA. Other districts with double-digit shares of KCTR teachers include Hickman Mills and the

[13] We made one exception in our geographic definition of the Kansas City area, which is to exclude Park Hill school district. While Park Hill has a Kansas City address, it is about 13 miles away from central Kansas City and is a highly advantaged school district. Park Hill did not receive any TFA or KCTR teachers during the period we study.

network of Crossroads Charter Schools. North Kansas City is the largest school district in the region (based on total enrollment and workforce size), but employs relatively few program teachers, all from KCTR. The North Kansas City student population is much wealthier than the neighboring KCPS population and has a lower share of underrepresented minority (URM; i.e. Black and Hispanic) enrollment.

We compare the composition of teachers and their placements from each program to teachers in the larger Kansas City area in terms of (a) the sector (charter or not), level (elementary, middle/junior high, or high school), and subject of the placement, (b) the characteristics of students in the school, and (c) teachers' own race/ethnicities and genders. Each program is compared to the local area using two different benchmarks. First, we use a simple teacher-weighted average from all 32 districts listed in Table 3.2 over the years 2012-2020 as a common benchmark for both programs. Second, we construct program-specific benchmarks calculated as district-by-year weighted averages that are unique to each program, where the district-by-year weights are the program-specific teacher shares of initial placements.

Formally, the district-by-year-weighted benchmark value of characteristic X for program j, which sends teachers to Kansas City area districts k in years t, can be written as:

$$\bar{X}_j = \sum_{n=kt}^{N_{kt}} w_{jkt} X_{kt} \tag{6}$$

In Equation (6), the weighting variable w_{jkt} is the fraction of all teachers produced by program j who are placed in district k in year t, and X_{kt} is the value of characteristic X for district k in year t. N_{kt} is the total number of district-by-year cells in which a teacher from

any of the two focal programs is placed. In district-years when no teacher from program *j* is placed, $w_{jkt} = 0$. For each program *j*, $\sum_{kt}^{N_{kt}} w_{jkt} = 1$.

The first benchmark, to the simple average over all teachers in the Kansas City area, compares teachers from each program to the region as a whole. The second benchmark, using the district-by-year weights, compares teachers from each program to other teachers in the same districts and years in which teachers from that program are placed. Both are useful for understanding the ways in which the programs influence the regional labor market.

3.4.2. Efficacy Analysis

We estimate the effects of teachers from each program on student achievement in grades 4-8 in math and ELA, on average, compared to non-program teachers during the period 2012-2019 using the following value-added model, structured based on Koedel, Mihaly and Rockoff (2015):

$$\begin{aligned} Y_{igmpqt} = \beta_0 + Y_{imt-1}\beta_1 + X_{it}\beta_2 + \bar{Y}_{mpt-1}\beta_3 + \bar{X}_{pt}\beta_4 \\ + T_{it}\beta_5 + P_{iqt}\beta_6 + \gamma_g + \delta_t + \varepsilon_{igmpqt} \end{aligned} \tag{7}$$

In Equation (7), Y_{igmpqt} is a standardized test score (standardized by grade-subject-year) for student *i* in grade *g* and subject *m*, who attended school *p* and was taught by teacher *q* in year *t*.[14] Y_{imt-1} is a 4-element vector of lagged test-score information. The first element is the same-subject lagged score, which we require of all students for inclusion in each subject-specific model (i.e., math or ELA). The second element is the lagged off-subject score—in our models of math achievement we include the lagged ELA score, and in the

[14] Some students take the algebra-I end-of-course test in the eighth grade instead of the standard grade-level test. We include these students in the analysis and their scores on the algebra-I test are separately standardized.

ELA model we include the lagged math score. To facilitate the inclusion of students who are missing just the off-subject lagged score (but still have the required same-subject score), we impute the missing score to the mean and include an indicator variable that we set equal to one if the score is missing. Finally, we add an interaction between the missing indicator variable and the lagged same-subject score, which improves estimation efficiency by allowing the model to rely more heavily on same-subject lagged performance to predict current performance for students who are missing the off-subject lagged score. The vector $\bar{Y}_{mpt-1}$ includes school-average values of the lagged test-score variables (lagged math achievement, lagged ELA achievement, and the fraction missing the off-subject test).

The vector X_{it} contains student characteristics. We include indicators for racial/ethnic and gender designations, free and reduced-price lunch (FRL) status, individualized education program (IEP) status, English language learner (ELL) status, and mobility status (i.e., an indicator for whether the student changed schools mid-year during year *t*).[15] We also include school percentages of these variables in the vector $\bar{X}_{pt}$.[16] The vector T_{it} controls for teacher experience. In our preferred specification we bin teachers into experience groups as in Clotfelter, Ladd, and Vigdor (2007): (1) 0 years prior experience (omitted category), (2) 1-2 years, (3) 3-5 years, (4) 6-12 years, (5) 13-20 years, (6) 21-27 years, and (7) 28+ years. The inclusion of the experience bins ensures teacher comparisons are restricted to occur within these experience bands. We also estimate a version of the model where we omit teacher experience entirely.[17] We elaborate below on

[15] The racial-ethnic categories we include are American Indian, Asian/Pacific Islander, Black, Hispanic, and multi-race (White is the omitted group).
[16] For parsimony we condense the racial-ethnic school percentage variable to capture just the percentage of non-White, non-Asian/Pacific Islander students; this has no substantive effect on our results.
[17] In results omitted for brevity we also confirm that all of our main findings are qualitatively upheld if we control for experience using linear and quadratic terms in place of the bins.

the insights afforded by the comparison of models with and without experience controls. γ_g and δ_t are grade and year fixed effects, respectively, and ε_{igmpqt} is the error term, which we cluster at the teacher level following Koedel et al. (2015). The vector P_{iqt} includes the treatment variables of interest: two separate indicator variables for whether student i's teacher in year t is from one of the focal programs. The omitted comparison group consists of non-program teachers in the Kansas City area.[18]

We can only estimate value-added for the subsample of teachers in grades 4-8 in math and ELA. In math, we observe 146 and 20 TFA and KCTR teachers, respectively. In ELA, the teacher sample sizes are 147 and 24 (some of these are overlapping—i.e., self-contained teachers in elementary schools). Due to the clustering structure of the models, the teacher sample sizes are the key determinants of statistical power. Our large TFA sample allows for fairly precise inference regarding program-level value-added. Our standard errors for the KCTR estimates are larger (about 60-100 percent larger depending on the outcome); they are still informative, but future research on KCTR (and other teacher residency models) would benefit from analyses at greater scale. A general challenge is that the scale of teacher residency programs is often modest, especially when one accounts for the fact that not all teachers are placed in tested grades and subjects. For example, in Papay

[18] The comparison group includes teachers from a subset of the 32 LEAs listed in Table 3.2. This is because a few LEAs do not cover grades 4-8 during the period 2012-2019 (e.g., a K-3 charter school). A small data issue also arises for teachers that start in one of the 32 focal districts, but subsequently move to a different district. For our main models, we include all teacher-year observations in any of the 32 focal school districts, and drop all observations outside of these districts (e.g., when a teacher moves out of the area but remains in Missouri). That said, how we handle data outside of the 32 focal districts is inconsequential to our results. For example, we have confirmed that the value-added results are qualitatively insensitive to including teacher-years from outside districts when teachers move. We have also confirmed our results are similar if we pull in more control teachers from the districts that teachers move to from our original sample. The robustness of our results to modifying the sample-inclusion criteria is consistent with the model's ability to control for student and school circumstances to isolate teacher effects on student learning (Koedel, Mihaly, and Rockoff, 2015).

et al. (2012)—the only other published study we are aware of focused on a teacher residency program that estimates value-added—their sample of residency teachers is similarly modest in size (N≈50).

3.5. Results

3.5.1. Descriptive Analysis

Figures 3.1-3.4 document the compositions of program teachers along the dimensions of school type and level of placements, characteristics of students at placement schools, and the demographics of teachers themselves. The figures are structured so that there is one graph for each program in each figure. For a given characteristic, the blue bars show average values for teachers in the focal program. The orange bars show average values for non-program teachers in the local area—i.e., the simple averages across all non-program teachers in the districts listed in Table 3.2. The grey bars show district-by-year weighted average values for non-program teachers as calculated by Equation (6). Note that the non-program group excludes teachers from both focal programs to facilitate its consistency across comparisons. In the appendix, we provide data tables with all of the information presented in the figures (Appendix Tables B.1-B.9). In addition, the appendix tables show comparisons restricted to only novice teachers (0-2 years) and provide some additional details that we omit from the figures for ease of exposition.

We illustrate the substantive difference between the orange and grey bars in the figures using TFA as an example. Returning to Table 3.2, note that KCPS employs 19.14 percent of all non-program local-area teachers, and thus the orange bars in our comparisons involving TFA (implicitly) give a 19.14 percent weight to KCPS when setting the comparison group. However, Table 3.2 also shows that TFA places a disproportionate

fraction of teachers in KCPS—specifically, 60.58 percent of TFA teachers are initially placed in KCPS. The grey bars re-weight KCPS so that it has a 60.58 percent weight in the TFA-specific comparison group. In other words, the orange bars compare TFA teachers to the local-area average on the whole, whereas the grey bars compare TFA teachers to other teachers in the districts (and years) that match TFA's own placement profile.

Beginning with Figure 3.1, we document teacher placements in terms of the schooling level and whether the placement is in a charter or non-charter school. We use DESE's rules to categorize each school as either an elementary school, middle/junior high school, or high school, as follows: Elementary schools are defined as schools with any combination that includes grades K-8, middle schools are those with any combination that includes grades 4-8 and is at least partly departmentalized, junior high schools have any combination that includes departmentalized grades 7-9, and high schools typically include grades 9-12 but may include grades 7-12.

For TFA, Figure 3.1 indicates that about 48 percent of teachers in our sample were placed in elementary schools. This value is below the simple average of the local area, which is about 56 percent, and also below the TFA-specific weighted average comparison group in the same districts and years, which is about 55 percent. Thus, from the first set of bars we conclude that TFA teachers are less likely than other teachers in the local area, and other teachers in the same districts and years in which TFA placements occur, to teach in elementary schools. The graph shows that the underrepresentation of TFA teachers in elementary grades is made up in high schools, where TFA teachers are disproportionately likely to be placed. In contrast, for KCTR teachers, the figure shows that they are more

likely to be placed in elementary schools but less likely to be placed in high schools relative to the larger local-area labor market.

The final set of bars in each graph in Figure 3.1 shows the percentages of teachers across all schooling levels placed in charter schools. Both TFA and KCTR teachers are much more likely to teach in charter schools than the average non-program teacher, which highlights the charter sector's disproportionate reliance on these programs for staffing. As indicated by the blue bars, the charter percentages for TFA and KCTR are about 38 and 49.5 percent, respectively. In comparison, as indicated by the orange bars, just 15.5 percent of non-program teachers in the area teach in charter schools. For the charter school comparison in particular, the weighted average comparisons given by the grey bars are not especially informative because almost all charter school operators are coded as their own districts in Missouri. Because the weighted-average comparison group forces weights proportional to each program's own district placements, it is by construction that the percentage of teachers in charter schools for each program virtually matches its program-specific weighted-average value indicated by the grey bar.

Next, in Figure 3.2 we document average student characteristics at teachers' placement schools. The structure of the figure is the same as Figure 3.1. We compare program teachers' placement schools using four school-level student characteristics: (1) the underrepresented minority (URM) enrollment percentage, which we calculate as the percentage of Black and Hispanic students (note that given the demographics of the local area, the URM percentage primarily captures the percentage of students who are Black), (2) the free and reduced-price lunch (FRL) eligible enrollment percentage, (3) the

percentage of students on an individualized education program (IEP), (4) the percentage of students who are English Language Learners (ELL).[19]

In terms of demographics, Figure 3.2 shows pronounced differences in the URM percentages between program and non-program teachers in the local area for TFA and KCTR—program teachers are much more likely to work in schools with higher URM student populations than non-program teachers. This can be seen by the gap between the blue and orange bars in each graph corresponding to the URM percentages. Note that for both programs, the URM-percentage gap disappears when the comparison shifts to the program-specific weighted averages, represented by the grey bars. This is informative about the mechanism by which program teachers are disproportionately working with high URM populations. Specifically, it means that the sorting is all occurring at the district level. Put another way, conditional on the district and year of the placement, teachers from the focal programs are working in schools with similar URM percentages as other, non-program teachers. But because the districts in which they are placed are high-URM districts, their exposure to URM students is higher than the local-area average of all teachers.

With respect to student FRL and ELL percentages, a qualitatively similar pattern plays out for each program, with modest variability in the magnitude of exposure gaps between program and non-program teachers. There is no indication of differences in schools' IEP percentages for program and non-program teachers.

[19] The FRL percentage is measured imperfectly because of the community eligibility provision, or CEP (Koedel and Parsons, 2020). The CEP-induced measurement error in the FRL percentage is likely to understate differences between program and non-program teachers along this dimension, but directionally the comparisons are still informative.

Figure 3.3 provides related evidence using average standardized test scores at teachers' placement schools. We make two notes about these comparisons. First, average test scores in all teachers' schools, even in the larger sample of non-program teachers, are large and negative in both subjects. This is because we standardize scores using the state distribution. The implication is that on average, students in the Kansas City area (as we have defined it) perform below the state average. Second, the test score results in Figure 3.3 are descriptive only. They may embody some program effects to the extent that program teachers impact test scores, about which we provide some evidence below. However, noting that non-schooling factors explain the majority of the variance in student test score levels (Parsons, Koedel, and Tan, 2019), and program teachers represent just a small fraction of the local-area workforce, our primary use of school-average test scores here, like in the previous figures, is to provide information about placement context.

Figure 3.3 shows that both TFA and KCTR teachers are placed in schools with substantially lower test score levels than other local-area teachers, as indicated by the large gaps between the blue and orange bars for these programs in Figure 3.3. Like the comparisons using the other student characteristics in Figure 3.2, the gaps shrink when we use the program-specific weighted averages, although they do not completely close as was the case for the previous comparisons. The fact that they mostly close points to district placements, and not school placements within districts, as the primary mechanism that drives the sorting of TFA and KCTR teachers into schools with lower achievement. But the fact that the gaps do not close all the way indicates that there is some additional within-district sorting of TFA and KCTR teachers that leads them to teach at lower-achieving schools compared to other teachers in the same districts and years. In the appendix

(Appendix Tables B.5 and B.6), we show that the small gaps that remain are partly explained by teacher experience. Specifically, noting that TFA and KCTR teachers are themselves inexperienced, if we restrict the weighted comparison group for each program to include only inexperienced non-program teachers, the achievement gap between program and non-program teacher placements declines further.[20]

We additionally note one other finding from the achievement comparisons not shown in Figure 3.3, but available in the appendix: the sorting patterns of program teachers in tested grades and subjects largely mirrors the sorting patterns of all program teachers discussed thus far. This is important for informing our analysis of value-added. If program teachers in tested grades and subjects are sorted to schools differently from other program teachers, it could have implications for inference from the value-added models and how the results relate to the descriptive comparisons. However, Appendix Tables B.4-B.6 show that the sorting patterns are similar, which aids in the interpretation of the value-added results.

Finally, Figure 3.4 documents the racial-ethnic and gender compositions of program teachers themselves relative to the local area using the same structure as the previous figures. We divide teachers into the following racial/ethnic groups: Asian/Pacific Islander, Black, Hispanic, White, and Other. The Other category is suppressed for ease of presentation, but results are reported in the appendix (Appendix Tables B.7-B.9). Compared to local area teacher demographics overall, as represented by the orange bars in the graphs, both programs are at least modestly diversifying, with generally larger

[20] This is consistent with well-documented evidence that inexperienced teachers, on average, are more likely to work in disadvantaged schools (Clotfelter, Ladd, and Vigdor, 2006; Goldhaber, Quince, and Theobald, 2018).

percentages of Asian/Pacific Islander, Black, and Hispanic teachers, and smaller percentages of White teachers, than the local-area average. KCTR is the most diverse program, particularly with respect to the percentage of Black teachers (37 percent, which is about 2.5 times higher than the local-area average of 13.55 percent). When we compare program teachers to the program-specific, district-and-year weighted average comparison groups (grey bars), KCTR teachers remain more diverse racial-ethnically than their non-program teaching peers; however, TFA teachers are a less diverse group.

We also examine gender diversity. Like the national teaching workforce, the workforce in the Kansas City area is female-dominated, as is each focal program. That said, the both TFA and KCTR are modestly diversity-improving along the dimension of gender.

3.5.2. Efficacy Analysis

Figure 3.5 shows the main value-added results for teachers in grades 4-8. We estimate four different models for each subject—all variants of Equation (7)—to recover estimates of the average value-added of teachers from each program relative to non-program teachers. A solid bar in the figure indicates the estimate is statistically distinguishable from the average value-added of non-program teachers at the 10 percent level or better and a clear bar indicates the estimate cannot be distinguished from the value-added of non-program teachers. The results underlying the figure are also available in tabular form in Appendix Tables B.10 (math) and B.11 (ELA).

The four different value-added specifications are labeled as models 1-4 in the figure. Model 4 is the full specification shown in Equation (7) and models 1-3 are sparser variants that build up to the full model. First, model 1 is a base specification that only controls for individual lagged achievement (Y_{imt-1}) and grade and year fixed effects.

Model 2 adds the individual student characteristics in the vector X_{it}. Model 3 further adds the school-level control vectors $\bar{Y}_{mpt-1}$ and $\bar{X}_{pt}$ to account for schooling context. The last component of the full model is the vector of teacher experience bins, denoted by T_{it}, which is added to model 4.

Before describing the results, we first note that our value-added estimates reflect the combined effects of (a) any selection into the programs and (b) any incremental improvement in teaching caused by the programs conditional on who enrolls. A program can have high value-added through either or both channels. For example, if a program recruits individuals who are predisposed to be strong teachers (i.e., positive selection) but does nothing via training to improve their performance, it will have high value-added; similarly, a program that recruits average teachers but offers exemplary training will also have high value-added. While our inability to disentangle the "selection" and "training" effect mechanisms is a limitation for some research questions, the combined effect is likely of first-order policy interest for districts looking to hire effective teachers.[21]

We first focus on the results from model 4, which is our preferred specification because it controls for student and school circumstances and compares teachers with similar experience levels. In math, the findings indicate that TFA and KCTR teachers outperform non-program teachers in the Kansas City area. Their value-added estimates are 0.11 and 0.15 student standard deviations higher on average, respectively. To give these estimates some context, the best research on teacher quality indicates that a one-standard-deviation move in the distribution of teacher quality as measured by math value-added—

[21] That is, districts will care more about whether effective teachers come out of a particular pipeline than why one pipeline produces stronger teachers than another. Disentangling the mechanisms is of greater interest from the perspective of informing teacher training organizations, for which knowing more about what aspects of training lead to greater improvements in efficacy for the teachers who participate is important.

e.g., a move from about the 50th to 85th percentile in the distribution of teacher quality—corresponds to a move in the student test distribution of about 0.10-0.15 standard deviations. Thus, our estimates of 0.11 and 0.15 imply that TFA and KCTR teachers are about 0.70-1.50 teacher standard deviations more effective than comparable non-program teachers in our data, on average. These are very large effects.

Comparing our findings to previous studies on the value-added of TFA teachers in math, our positive estimate for TFA is qualitatively consistent with the literature outside of the New York City (e.g., Boyd et al., 2006; Kane, Rockoff, and Staiger, 2008), and inclusive of the New York City estimates, falls somewhere in the middle of the range of estimates in previous research.[22] We are not aware of any comparable prior literature on KCTR specifically. However, a similar evaluation of the efficacy of teachers from the Boston Teacher Residency (BTR) by Papay et al. (2012) finds negative achievement effects in math.

Looking at the estimates for TFA and KCTR across models in Figure 3.5 is also instructive. In the sparse model—model 1—there are no statistically detectable differences between the program and non-program teachers. However, once we control for student characteristics in model 2, the differences emerge and persist as the specifications become richer. This finding is previewed by the descriptive analysis above, which shows that TFA and KCTR teachers are more likely to be placed in schools with more disadvantaged and lower-achieving students. Model 1 does not account for these placement differences except to the extent that they are captured by students' own lagged test scores. The more robust

[22] Again, these studies include Decker, Mayer, and Glazerman (2004), Backes et al. (2019), and Xu, Hannaway, and Taylor (2011).

accounting for teaching context in models 2-4 reveals important performance gaps between TFA and KCTR teachers compared to other Kansas City area teachers.

Another aspect of the cross-model estimates that merits attention is the difference between models 3 and 4. These models differ only by whether we control for teacher experience. On the one hand, the experience-conditional comparisons in model 4 are useful for gauging the efficacy of TFA teachers relative to their similarly-experienced non-program peers. However, it is also desirable to compare TFA teachers to all program teachers without conditioning on experience because part of the TFA treatment, arguably, is increased student exposure to relatively inexperienced teachers. Model 3 does not separately control for experience, so implicitly it compares TFA teachers to all non-program teachers, who are much more experienced on average. The estimates from both models are informative about the TFA treatment effect. We have less information on the longevity of KCTR teachers and no *ex ante* reason to believe they will have shorter careers, on average, than non-program teachers, so for KCTR teachers the comparisons in model 4 are preferred.

As a practical matter, our estimates change little going from models 3 to 4 for both TFA and KCTR. Upon further investigation, the reason for the modest change in the estimates is that the experience-efficacy gradient among non-program local area teachers is modest, and effectively flat over a large range of the experience distribution after the first year (results suppressed for brevity). While this result is not entirely out of line with what has been found elsewhere in the literature (Clotfelter, Ladd, and Vigdor, 2006;

Wiswall, 2013), the gradient in Kansas City is especially flat.[23] Thus, whether we condition on teacher experience in our comparisons involving TFA and KCTR teachers has little bearing on the findings.

Next, we turn to the ELA estimates in the second panel of Figure 3.5. The results from model 4 suggest small positive effects of TFA and KCTR teachers on ELA growth, compared to similarly-experienced non-program teachers, on the order of about 0.03-0.05 student standard deviations. Our smaller findings in ELA are consistent with the broad empirical regularity documented in research that teacher effects in math are larger than in English Language Arts (e.g., Lefgren and Sims, 2012; Goldhaber, Cowan, and Walsh, 2013). For TFA specifically, these findings also align with findings from related studies (e.g., Backes et al., 2019; Decker, Mayer, and Glazerman, 2004; Kane, Rockoff, and Staiger, 2008).[24] The ELA effects are not statistically preserved in model 3, where we do not condition on teacher experience, although the absolute impact of going from model 3 to model 4 is similar in ELA and math. The difference is that the estimates from model 4 for ELA are already small, so the modest reduction in the coefficients going to model 3 pushes them below the significance threshold.

3.6. Extension: Teacher Retention

We briefly extend our analysis to assess whether program teachers are more or less likely to remain in the Kansas City area compared to non-program teachers. For KCTR we can perform the retention analysis for the 2018 placement cohort only—the 2019 and 2020

[23] Wiswall (2013) shows that one explanation for the generally flat, or only slightly upward sloping experience-performance gradient, is negative selection into who stays in teaching. That said, for our purposes the distinction of mechanisms is not critical.

[24] Decker, Mayer, and Glazerman (2004) and Kane, Rockoff, and Staiger (2008) find no statistical evidence of a TFA effect in ELA, although their standard errors also cannot rule out modest positive impacts, and Backes et al. (2019) estimate a statistically significant TFA effect on ELA of 0.02 student standard deviations (which is very close to our estimate).

cohorts are too new to credibly study retention. We look to see if the 2018 cohort of KCTR teachers remain in the workforce in 2019 and 2020 (N=31, noting that one 2018 placement was in a non-teaching position). Because we have more cohort data for TFA, we expand the retention analysis to look forward up to five years for TFA teachers whose initial placements were between 2012-2016 (inclusive; N=340).

Retention rates in the Kansas City area for both programs are reported in Figure 3.6 compared to retention rates for non-program first-year teachers in the same years. We define the Kansas City area more broadly for examining retention than in the previous analysis. Specifically, we treat a teacher as retained in the Kansas City area if she is observed teaching in a public school in the Missouri portion of the formal metropolitan statistical area (MSA) as defined by the U.S. Census. We also report retention rates in the larger Missouri workforce in the appendix (which are slightly higher but similar; see Appendix Tables B.12 and B.13).

In Figure 3.6, retention after one year indicates that the teacher was observed working in the year following the initial placement (i.e., in year 2). Retention rates after years 2, 3, and 4 are similarly defined and cumulative (e.g., a value of 50 after year-3 would indicate that 50 percent of the entering teachers were still teaching in the area into the 4th year). As in previous figures, we compare KCTR and TFA teachers to the simple average of teachers in the districts listed in Table 3.2, and the program-specific weighted averages based on the districts and years in which KCTR and TFA teachers were placed. We restrict the comparison groups to new teachers only for this analysis.

Figure 3.6 shows that 1- and 2-year retention rates for KCTR are above the local-area average and the district-by-year weighted average based on KCTR placements. The

retention gaps over the first two years for KCTR teachers are quite large—after one year, KCTR teachers are more than 20 percentage points more likely to remain in teaching in the area relative to non-program new teachers in the same districts and years. After two years the gap shrinks but remains sizeable, at about 14 percentage points. The higher retention rates of KCTR teachers over the span of time we can evaluate is consistent with the feature of the program that residents agree to teach in a high-need school in the Kansas City area for at least three years.

For TFA, more than 99 percent of TFA teachers in our sample return after the first year, which is consistent with the 2-year program commitment. However, there is a stark drop off going into year 3, with only 57 percent of TFA teachers remaining beyond the second year. Retention after the 4th year—i.e., the percent of TFA teachers who are still teaching in year-5—is just 32.35 percent. These retention rates for TFA teachers in our sample are similar to rates calculated using national TFA data (Donaldson and Johnson, 2011).

The seemingly low retention rate of TFA teachers in the metro area—32.35 percent—is perhaps disappointing, but less so given the larger context that non-program teachers in these same districts also have a low 5-year retention rate in the area (40.84 percent). As noted by Donaldson and Johnson (2011), this points to the issue of high turnover being less about the TFA program per se, and more about the difficult teaching environments faced by TFA teachers. While non-program teachers are retained at a higher rate than TFA teachers in the Kansas City area by the fifth year, the retention gap is not so large as to be a first-order difference between TFA and other teachers.

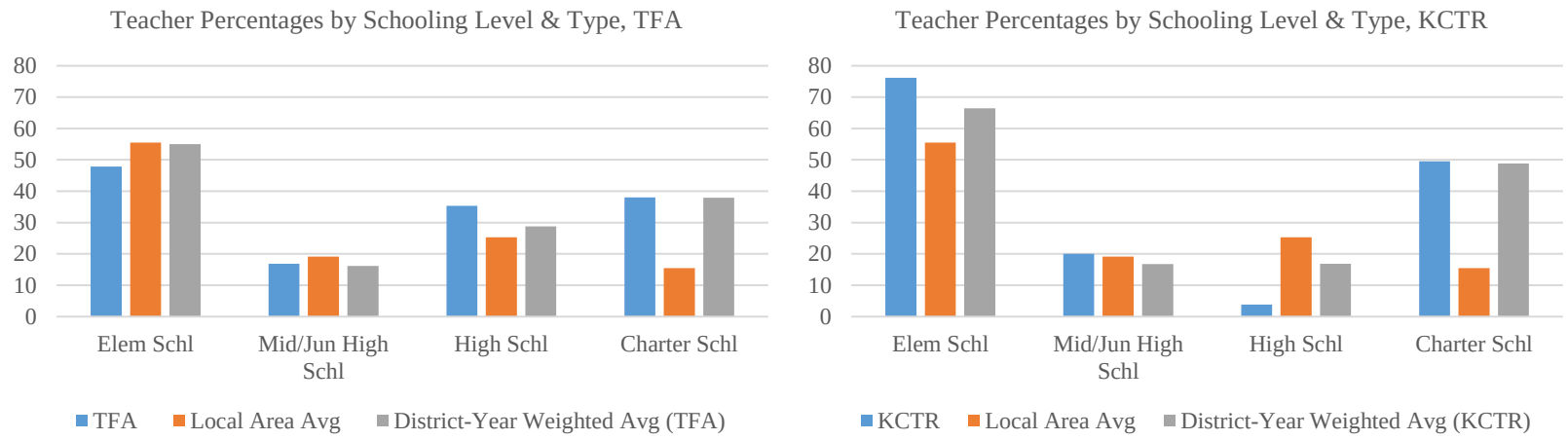
3.7. Conclusion

We evaluate two alternative teacher preparation programs—Teach for America (TFA) and Kansas City Teacher Residency (KCTR)—to assess how they contribute to the teacher labor market in Kansas City, Missouri. Descriptively, we document program placements in terms of school types and levels, characteristics of students taught, and the racial/ethnic and gender diversity of the teachers themselves. Although there is some heterogeneity across the two programs, common themes are that these programs disproportionately place teachers in charter schools, and more broadly, in schools serving disadvantaged students. Teachers from both programs are also more racial-ethnically diverse than the larger local-area teaching population, although only KCTR teachers are more diverse than other teachers in the same districts in which they place. Notably, KCTR seems to be particularly effective as a pathway for Black teachers to enter the profession.

We also estimate the value-added of program teachers relative to non-program teachers. We find that students in grades 4-8 whose teachers are from TFA and KCTR have much higher achievement growth in math than similar students, in similar schools, who are taught by non-program teachers. We also find evidence of small, positive impacts of teachers from these programs on ELA achievement in grades 4-8.

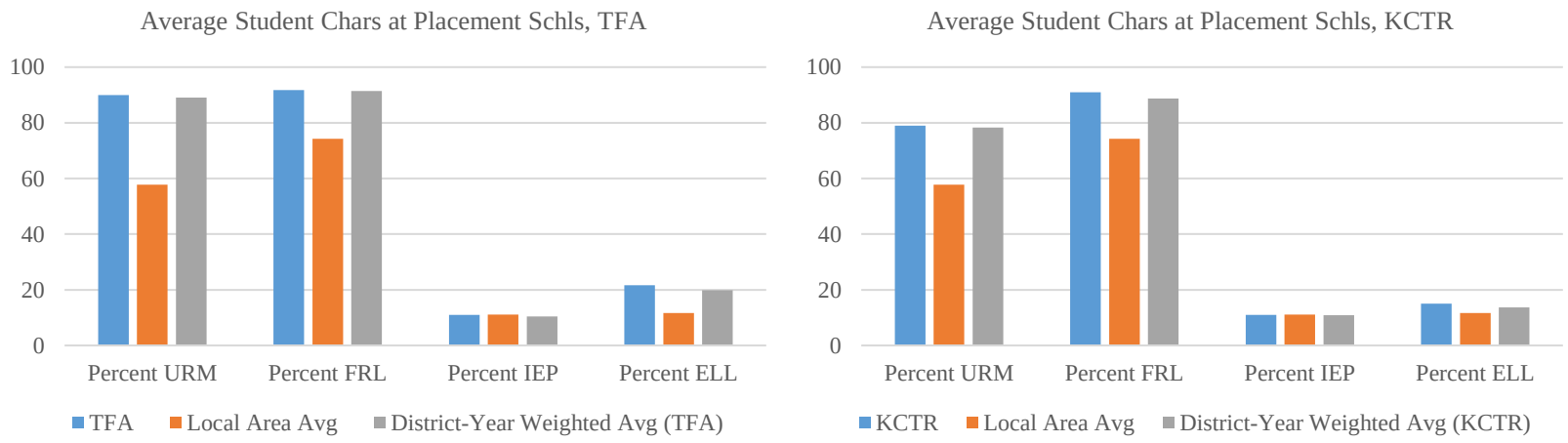
Finally, we analyze teacher retention for TFA and KCTR teachers. We find that KCTR teachers are much more likely to be retained in Kansas City MSA than comparable non-program teachers over the first 3 years post-training. TFA teachers are more likely be retained after the first year, but their retention rates drop off thereafter. Retention rates after 4 years for TFA teachers are below comparable rates for non-program teachers in the same districts, but not markedly (32 versus 41 percent).

Figure 3.1. Teacher Placements by School Type (i.e., Charter or Not) and Level (Elementary, Middle/Junior High, or High School).



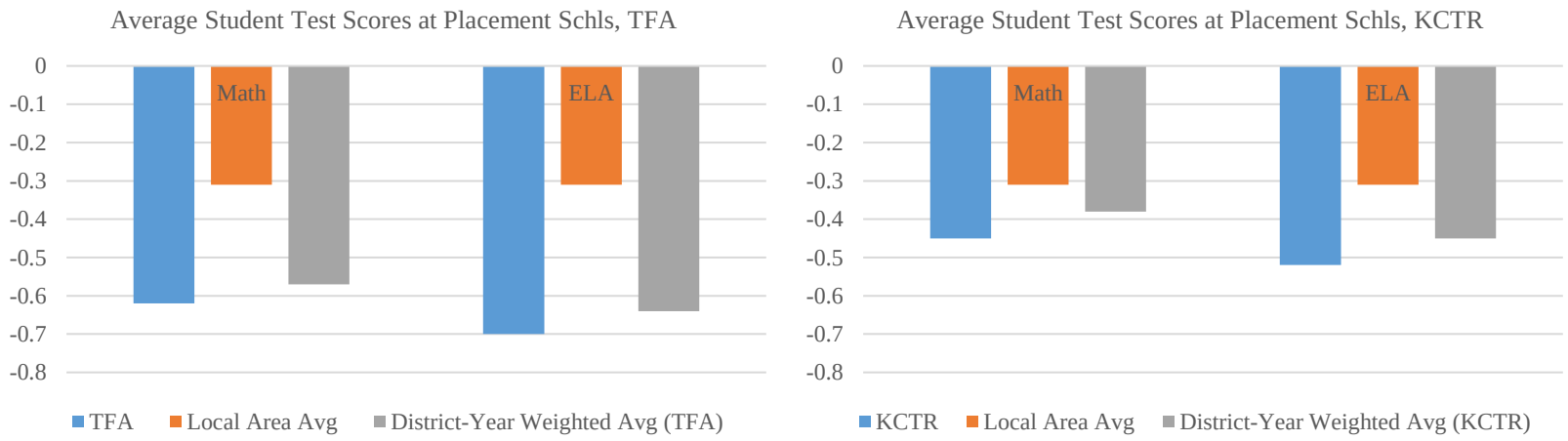
Notes: The local-area averages (orange bars) are for all teachers in the comparison districts shown in Table 3.2—they are not program specific and thus do not change in the graphs. The district-and-year weighted averages (grey bars) are weighted based on each program's own placement patterns to compare program teachers to teachers working in the same districts and years of the placements.

Figure 3.2. Teacher Placements by the Characteristics of Students Attending the Initial Placement School.



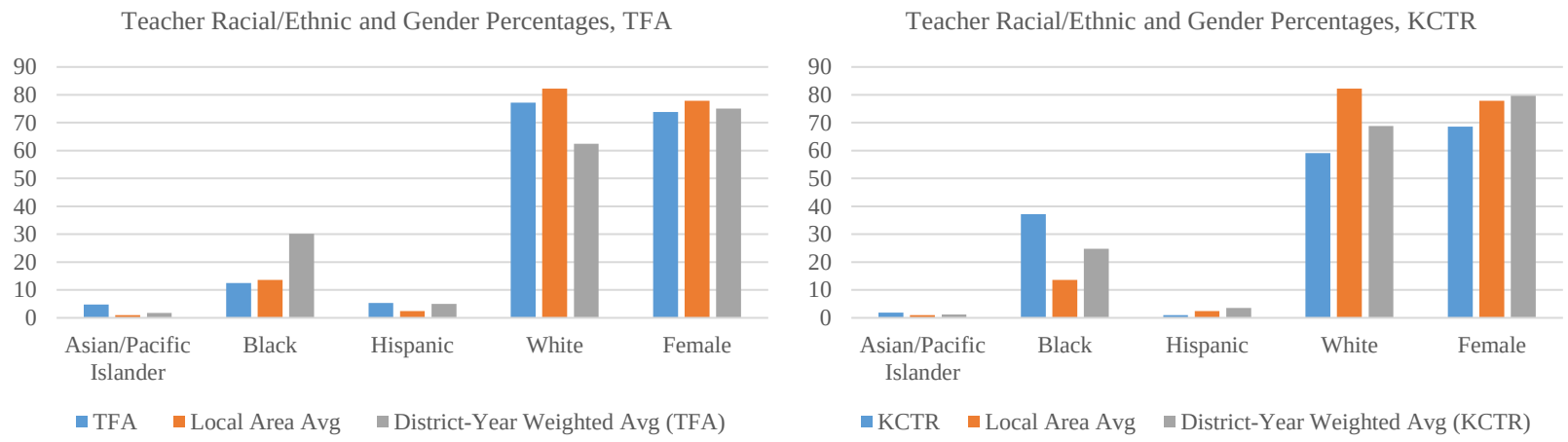
Notes: The local-area averages (orange bars) are for all teachers in the comparison districts shown in Table 3.2—they are not program specific and thus do not change in the graphs. The district-and-year weighted averages (grey bars) are weighted based on each program's own placement patterns to compare program teachers to teachers working in the same districts and years of the placements. URM=underrepresented minority (Black and Hispanic); FRL=free or reduced-price lunch eligible; IEP=individualized education program; ELL=English language learner.

Figure 3.3. Teacher Placements by the Standardized Achievement Level of Students Attending the Initial Placement School in Math and English Language Arts (ELA).



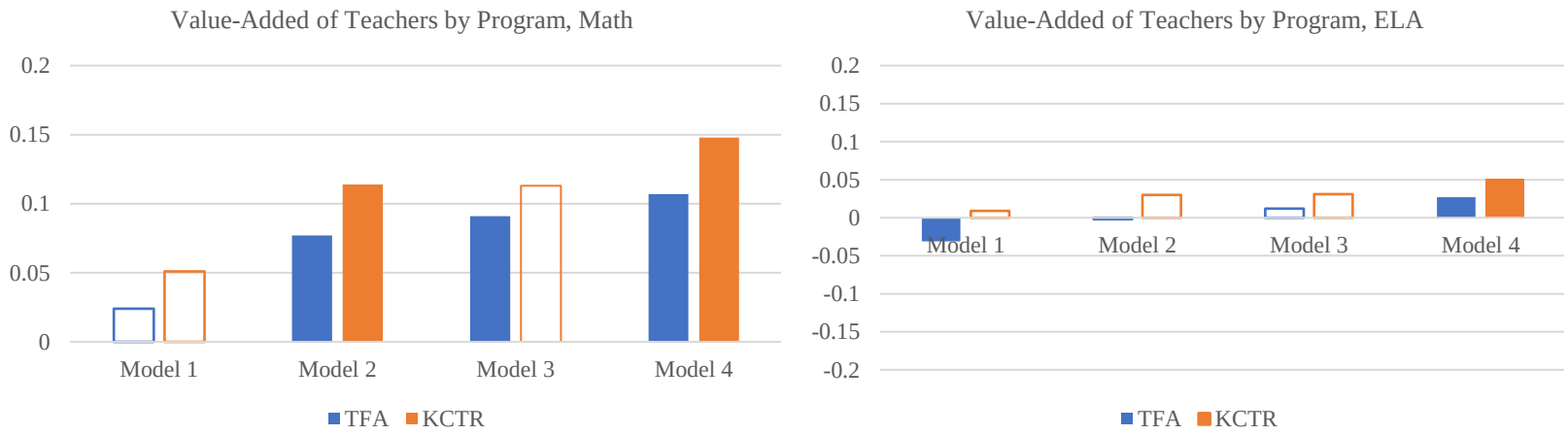
Notes: The local-area averages (orange bars) are for all teachers in the comparison districts shown in Table 3.2—they are not program specific and thus do not change in the graphs. The district-and-year weighted averages (grey bars) are weighted based on each program's own placement patterns to compare program teachers to teachers working in the same districts and years of the placements.

Figure 3.4. Teachers' Racial/Ethnic and Gender Designation Percentages.



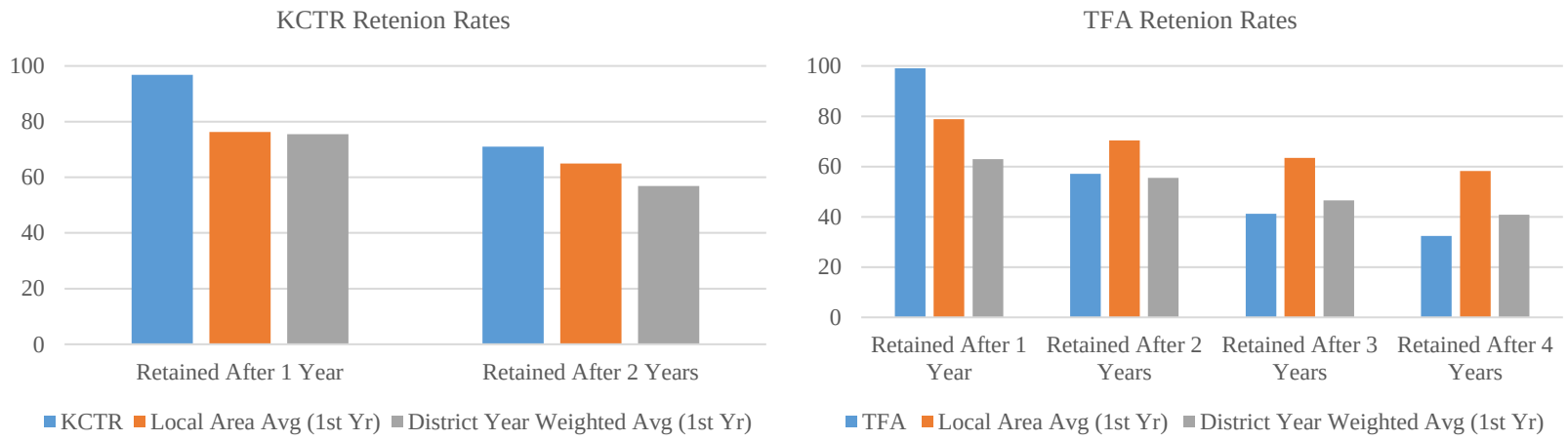
Notes: The local-area averages (orange bars) are for all teachers in the comparison districts shown in Table 3.2—they are not program specific and thus do not change in the graphs. The district-and-year weighted averages (grey bars) are weighted based on each program's own placement patterns to compare program teachers to teachers working in the same districts and years of the placements.

Figure 3.5. Value-added to Achievement in Math and English Language Arts (ELA) in Grades 4-8 for Program Teachers Compared to Non-program Teachers Using Different Value-added Specifications as Described in the Text.



Notes: Moving from model 1 to model 4 increases the comprehensiveness of the value-added model. Model 1 is a base specification that only controls for individual lagged achievement (Y_{imt-1}) and grade and year fixed effects. Model 2 adds the individual student characteristics in the vector X_{it}. Model 3 further adds the school-level control vectors $\bar{Y}_{pmt-1}$ and $\bar{X}_{pt}$. The last component of the full model is the vector of teacher experience bins, denoted by T_{it}, added to model 4, which is the full , preferred specification. Solid bars indicate differences that are statistically significant at the 10 percent level or higher; clear bars indicate statistically insignificant differences.

Figure 3.6. Retention Rates for KCTR and TFA Teachers Relative to Novice Non-program Teachers in the Kansas City Area, as Defined by the Missouri Portion of the Metropolitan Statistical Area.



Notes: The local-area averages (orange bars) are for all first-year teachers in the full set of comparison districts shown in Table 3.2, for the same years as program teachers (the definition of this comparison group differs slightly from the definition of the analogous comparison groups above). The district-and-year weighted averages (grey bars) are weighted based on each program's own placement patterns and thus compare program teachers to other first-year teachers working in the same districts and years of the placements. The retention rates reported in the figure are cumulative (e.g., the "Retained After 3 Years" percentage for TFA reports the number of originally-placed TFA teachers who are still working in the Kansas City MSA in the 4th year after the initial placement). Retention rates for KCTR are reported for the 2018 placement cohort only; retention rates for TFA are reported for the 2012-2016 placement cohorts.

Table 3.1. Teacher Counts by Program and the First Post-program Placement Year.

	TFA[a]	KCTR
2012	127	
2013	69	
2014	64	
2015	48	
2016	40	
2017	33	
2018	46	32
2019	1	28
2020	2	48
Total	430	108
Total (excluding non-teaching placements)	416[b]	105[b]
Number unmatched	0	0

Notes:

[a] TFA provided placement data for cohorts through 2018. The handful of post-2018 TFA placements are teachers who completed their TFA training in an earlier year but delayed entry into the workforce.

[b] Non-teaching positions include central office positions, individuals listed as working in special centers, and teacher coaches, among other positions.

Table 3.2. Teacher Placement Percentages across Districts in the Kansas City Area, Combined across All Years.

	TFA	KCTR	Non-Program Teachers
North Kansas City 74	0	8.57	25.19
Kansas City 33	60.58	27.62	19.14
Independence 30	0	0	16.8
Raytown C-2	1.2	0	10.89
Hickman Mills C-1	0.24	15.24	8.6
Center 58	0	0	4.01
Frontier Schools	0	0	2.16
Hogan Preparatory Academy	6.49	0	1.39
Academie Lafayette	0	0	1.3
University Academy	4.57	3.81	1.21
Guadalupe Centers Schools	6.25	0	1.14
Kc International Academy	0.24	0.95	1.09
Brookside Charter School	1.92	7.62	0.81
Lee A. Tolbert Com. Academy	0.96	3.81	0.77
Allen Village	0	0	0.67
Crossroads Charter Schools	0	10.48	0.67
Ewing Marion Kauffman School	6.25	5.71	0.57
B. Banneker Academy	0	0	0.44
Gordon Parks Elem.	0	0.95	0.4
Scuola Vita Nuova	0	1.9	0.4
Pathway Academy	1.2	0	0.4
Kipp: Endeavor Academy	2.64	7.62	0.3
Delasalle Charter School	1.92	0	0.29
Genesis School Inc.	2.88	2.86	0.24
Derrick Thomas Academy	1.68	0	0.22
Hope Leadership Academy	0	0	0.2
Renaissance Acad Math and Sci	0.48	0	0.18
Academy For Integrated Arts	0	0.95	0.18
Citizens Of The World Charter	0.24	0	0.17
Hope Academy	0	0	0.13
Urban Com. Leadership Academy	0.24	0	0.04
Kansas City Girls Prep Academy	0	1.9	0.01
Sum	100	100	100

Notes: Columns sum to 100 percent.

Appendix A

Supplementary Tables for Chapter 1

Table A.1. Construction of the Analytical Sample.

First-time, full-time, degree-seeking entering freshmen at one of the 13 public four-year universities in Missouri between 2001 and 2010	178,114	
	Records Lost	Remaining Sample
Attended a non-Missouri public high school	-59,291	118,823
Different years of high school graduation and college matriculation	-7,686	111,137
Missing SSN	-62	111,075
Duplicates drop	-14	111,061
Missing DESE high school data	-583	110,478
No reported math or science teachers	-179	110,299
Missing high school enrollment data	-332	109,967
Missing free or reduced-price lunch data	-2	109,965
Unknown student gender	-35	109,930

Table A.2. Public Four-year Universities in Missouri.

	Enrollment Share	Graduation Rate	Share of Total STEM Entrants	Share of Total STEM Graduates
University of Central Missouri	0.09	57%	0.07	0.07
Harris-Stowe State University	0.01	12%	0.00	0.00
Lincoln University	0.02	26%	0.01	0.01
Missouri Southern State University	0.04	39%	0.03	0.02
Missouri Western State University	0.06	34%	0.04	0.03
Truman State University	0.07	78%	0.09	0.08
Northwest Missouri State University	0.07	57%	0.04	0.05
Southeast Missouri State University	0.09	53%	0.06	0.05
Missouri State University	0.17	61%	0.08	0.09
University of Missouri-Columbia	0.25	72%	0.30	0.32
Missouri University of Science and Technology	0.05	67%	0.21	0.22
University of Missouri-Kansas City	0.05	49%	0.04	0.04
University of Missouri-St Louis	0.03	47%	0.02	0.02

Table A.3. The Effects of Female and Underrepresented Minority STEM Teachers on College Attendance.

	Number of Matriculated Students
FT	-0.531 (0.737)
MT	-2.462 (6.046)
HS Controls & Year FE	X
HS FE	X
Dependent Variable Mean (SD)	23 (32)
N (High School by Year)	4,714

Notes: FT and MT denote the share of female math & science teachers and the share of underrepresented minority math & science teachers, respectively. The dependent variable is the number of students matriculated to a Missouri four-year public university in each corresponding cohort. Standard errors clustered by high schools are shown in the parenthesis. *** p<0.01, ** p<0.05, * p<0.10.

Table A.4. STEM Enrollment and Completion Models, Matching Effects, Alternative Measures.

	STEM Enrollment	STEM Completion
FT	0.012 (0.023)	-0.021 (0.019)
FT X FS	-0.029 (0.021)	0.0004 (0.018)
MT	0.120 (0.071)*	0.083 (0.047)*
MT X MS	-0.065 (0.066)	-0.022 (0.047)
Individual Controls & Year FE	X	X
HS Controls	X	X
HS FE	X	X
N	109,930	109,930

Notes: FT and MT denote the share of female math & science teachers and the share of underrepresented minority math & science teachers, respectively. FS and MS are indicator variables equal to one if a student is female and an underrepresented minority, respectively. Standard errors clustered by high schools are shown in the parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Table A.5. STEM Enrollment and Completion Models, Various Time Periods.

	All Years	2001-2005	2006-2010
	(1)	(2)	(3)
STEM Enrollment			
FT	0.007 (0.022)	-0.015 (0.040)	0.020 (0.038)
FT X FS	-0.025 (0.021)	-0.004 (0.027)	-0.049 (0.028)*
MT	0.130 (0.070)*	0.074 (0.115)	0.142 (0.130)
MT X MS	-0.064 (0.067)	0.056 (0.073)	-0.175 (0.104)*
STEM Completion			
FT	-0.017 (0.019)	-0.029 (0.029)	0.013 (0.029)
FT X FS	0.002 (0.017)	-0.001 (0.023)	0.003 (0.023)
MT	0.087 (0.045)*	0.084 (0.078)	0.041 (0.089)
MT X MS	-0.024 (0.048)	-0.047 (0.047)	0.010 (0.072)
Individual Controls & Year FE	X	X	X
HS Controls	X	X	X
HS FE	X	X	X
N	109,930	52,486	57,444

Notes: FT and MT denote the share of female math & science teachers and the share of underrepresented minority math & science teachers, respectively. FS and MS are indicator variables equal to one if a student is female and an underrepresented minority, respectively. Standard errors clustered by high schools are shown in the parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Appendix B

Supplementary Tables for Chapter 3

Table B.1. Teacher Placement Percentages by Grade Span and Subject for Initial Post-program Placements, Combined across All Years, Compared to the Simple Region Average.

	TFA	KCTR	All Non-Program Teachers	All Non-Program Teachers (Novice Only)
Elementary Total	47.84	76.19	55.53	58.67
Tested grades and subjects (4-8)	12.02	28.57	10.05	11.97
PK-3	24.28	36.19	21.94	26.06
Language Specialist	4.33	0.95	3.8	2.21
Special Education	4.33	0	6.15	5.97
Other	2.88	10.48	13.58	12.46
Middle School/Junior High School Total	16.83	20	19.19	19.18
Tested Grades and Subjects (4-8)	8.89	12.38	6.82	7.54
Language Specialist	1.68	0.95	0.47	0.52
Special Education	1.2	0	2.15	1.5
Science	2.88	3.81	2.03	2.14
Social Studies	0.48	0.95	1.77	1.66
Other	1.68	1.9	5.94	5.82
High School Total	35.34	3.81	25.28	22.16
Tested Grades and Subjects (7-8)	7.93	0	0.56	0.55
English Language Arts	6.25	0.95	4.05	4.15
Math	4.33	1.9	2.78	2.66
Science	7.21	0.95	2.61	3.1
Social Studies	1.92	0	2.77	2.08
Special Education	4.81	0	2.93	1.97
Other	2.88	0	9.58	7.64
Sum (Totals)	100	100	100	100
Total Charter Percent	37.98	49.52	15.47	23.60

Notes: Schooling levels are defined as described in the text. The "other" category at each level contains a number of sparsely populated positions including physical education, health, music, and other specialty subjects and non-traditional assignments. The "total charter percent" row combines charter placements across all schooling levels and subjects. Novice-only teachers in the last column are teachers with 0-2 years of experience.

Table B.2. Detailed Analog to Table B.1 for TFA, with Program-specific Weighted-average Comparison.

		TFA Weighted	
	TFA	All Non-Program Teachers	All Non-Program Teachers (Novice Only)
Elementary Total	47.84	55.03	54.62
Tested grades and subjects (4-8)	12.02	9.77	9.51
PK-3	24.28	23.55	27.64
Language Specialist	4.33	3.60	1.39
Special Education	4.33	6.52	5.79
Other	2.88	11.59	10.28
Middle School/Junior High School Total	16.83	16.14	17.77
Tested Grades and Subjects (4-8)	8.89	5.75	7.17
Language Specialist	1.68	1.45	1.11
Special Education	1.2	1.49	0.86
Science	2.88	1.69	1.93
Social Studies	0.48	0.96	1.74
Other	1.68	4.80	4.95
High School Total	35.34	28.83	27.62
Tested Grades and Subjects (7-8)	7.93	2.34	4.18
English Language Arts	6.25	4.12	4.09
Math	4.33	2.64	2.48
Science	7.21	2.86	3.71
Social Studies	1.92	3.60	2.26
Special Education	4.81	2.85	1.23
Other	2.88	10.42	9.67
Sum (Totals)	100	100	100
Total Charter Percent	37.98	37.89	37.91

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience. Program-specific weighted averages are calculated as shown by Equation (6). The weights are the district-by-year initial placement shares of program teachers and the characteristics are weighted across all non-program teachers (i.e., all teachers who are not from one of the two focal programs).

Table B.3. Detailed Analog to Table B.1 for KCTR, with Program-specific Weighted-average Comparison.

		KCTR Weighted	
	KCTR	All Non-Program Teachers	All Non-Program Teachers (Novice Only)
Elementary Total	76.19	66.42	67.01
Tested grades and subjects (4-8)	28.57	11.52	13.23
PK-3	36.19	23.70	26.41
Language Specialist	0.95	5.15	2.83
Special Education	0	7.55	7.91
Other	10.48	18.49	16.63
Middle School/Junior High School Total	20.00	16.74	15.62
Tested Grades and Subjects (4-8)	12.38	5.86	6.01
Language Specialist	0.95	1.38	1.30
Special Education	0	2.28	0.92
Science	3.81	1.24	1.72
Social Studies	0.95	1.40	1.17
Other	1.90	4.58	4.50
High School Total	3.81	16.84	17.37
Tested Grades and Subjects (7-8)	0	0.41	0.33
English Language Arts	0.95	2.83	3.56
Math	1.90	1.82	1.97
Science	0.95	1.89	3.62
Social Studies	0	1.79	1.41
Special Education	0	1.96	1.08
Other	0	6.15	5.40
Sum (Totals)	100	100	100
Total Charter Percent	49.52	48.85	47.45

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience. Program-specific weighted averages are calculated as shown by Equation (6). The weights are the district-by-year initial placement shares of program teachers and the characteristics are weighted across all non-program teachers (i.e., all teachers who are not from one of the two focal programs).

Table B.4. Average Characteristics of Students in the Schools and Years of Teachers' First Placements, Compared to Simple Average School Characteristics (Teacher Weighted) in the Region.

	TFA	KCTR	All Non-Program Teachers	All Non-Program Teachers (Novice Only)
Percent Asian/Pacific Islander	2.40	3.31	2.91	2.71
Percent Black	64.21	62.00	39.90	45.51
Percent Hispanic	25.71	16.95	17.90	19.86
Percent White	6.28	12.46	33.83	27.09
Percent Other	1.40	5.27	5.46	4.84
Percent FRL	91.78	90.97	74.31	80.80
Percent IEP	10.94	10.97	11.05	11.08
Percent ELL	21.67	15.08	11.68	14.06
Average Math achievement (standardized)	-0.62	-0.45	-0.31	-0.41
Average ELA achievement (standardized)	-0.70	-0.52	-0.31	-0.40
Among teachers in tested grades and subjects only (4-8):				
Average Math achievement (standardized)	-0.61	-0.49	-0.32	-0.43
Average ELA achievement (standardized)	-0.69	-0.54	-0.31	-0.41

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience.

Table B.5. Detailed Analog to Table B.4 for TFA, with Program-specific Weighted-average Comparison.

		TFA Weighted	
	TFA	All Non-Program Teachers	All Non-Program Teachers (Novice Only)
Percent Asian/Pacific Islander	2.40	2.46	2.11
Percent Black	64.21	65.55	69.55
Percent Hispanic	25.71	23.55	20.85
Percent White	6.28	6.95	5.99
Percent Other	1.40	1.49	1.50
Percent FRL	91.78	91.38	91.40
Percent IEP	10.94	10.45	10.51
Percent ELL	21.67	19.87	17.35
Average Math achievement (standardized)	-0.62	-0.57	-0.60
Average ELA achievement (standardized)	-0.70	-0.64	-0.66
Among teachers in tested grades and subjects only (4-8):			
Average Math achievement (standardized)	-0.61	-0.55	-0.67
Average ELA achievement (standardized)	-0.69	-0.62	-0.72

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience. Program-specific weighted averages are calculated as shown by Equation (6). The weights are the district-by-year initial placement shares of program teachers and the characteristics are weighted across all non-program teachers (i.e., all teachers who are not from one of the two focal programs).

Table B.6. Detailed Analog to Table B.4 for KCTR, with Program-specific Weighted-average Comparison.

		KCTR Weighted	
	KCTR	All Non-Program Teachers	All Non-Program Teachers (Novice Only)
Percent Asian/Pacific Islander	3.31	2.71	2.66
Percent Black	62.00	61.35	62.31
Percent Hispanic	16.95	16.92	16.66
Percent White	12.46	14.04	13.87
Percent Other	5.27	4.98	4.49
Percent FRL	90.97	88.77	89.43
Percent IEP	10.97	10.84	11.14
Percent ELL	15.08	13.69	13.15
Average Math achievement (standardized)	-0.45	-0.38	-0.41
Average ELA achievement (standardized)	-0.52	-0.45	-0.48
Among teachers in tested grades and subjects only (4-8):			
Average Math achievement (standardized)	-0.49	-0.38	-0.40
Average ELA achievement (standardized)	-0.54	-0.45	-0.48

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience. Program-specific weighted averages are calculated as shown by Equation (6). The weights are the district-by-year initial placement shares of program teachers and the characteristics are weighted across all non-program teachers (i.e., all teachers who are not from one of the two focal programs).

Table B.7. Program Teachers' Race-ethnicity and Gender Percentages Compared to Simple Average (Teacher Weighted) in the Region.

	TFA	KCTR	All Non-Program Teachers	All Non-Program Teachers (Novice Only)
Percent Asian/Pacific Islander	4.81	1.90	1.00	1.24
Percent Black	12.50	37.14	13.55	13.83
Percent Hispanic	5.29	0.95	2.40	3.18
Percent White	77.16	59.05	82.26	80.78
Percent Other	0.24	0.95	0.79	0.97
Percent female	73.80	68.57	77.83	77.68

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience.

Table B.8. Detailed Analog to Table B.7 for TFA, with Program-specific Weighted-average Comparison.

		TFA Weighted	
	TFA	All Non-Program Teachers	All Non-Program Teachers (Novice Only)
Percent Asian/Pacific Islander	4.81	1.76	2.70
Percent Black	12.50	30.20	29.42
Percent Hispanic	5.29	5.00	5.79
Percent White	77.16	62.48	61.09
Percent Other	0.24	0.56	1.00
Percent female	73.80	75.08	74.06

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience. Program-specific weighted averages are calculated as shown by Equation (6). The weights are the district-by-year initial placement shares of program teachers and the characteristics are weighted across all non-program teachers (i.e., all teachers who are not from one of the two focal programs).

Table B.9. Detailed Analog to Table B.7 for KCTR, with Program-specific Weighted-average Comparison.

		KCTR Weighted	
	KCTR	All Non-Program Teachers	All Non-Program Teachers (Novice Only)
Percent Asian/Pacific Islander	1.90	1.25	1.77
Percent Black	37.14	24.77	25.27
Percent Hispanic	0.95	3.55	4.54
Percent White	59.05	68.84	67.84
Percent Other	0.95	1.59	0.57
Percent female	68.57	79.61	75.17

Notes: Novice-only teachers in the last column are teachers with 0-2 years of experience. Program-specific weighted averages are calculated as shown by Equation (6). The weights are the district-by-year initial placement shares of program teachers and the characteristics are weighted across all non-program teachers (i.e., all teachers who are not from one of the two focal programs).

Table B.10. Value-added to Student Achievement by Program, Grades 4-8, Math.

	Model 1	Model 2	Model 3	Model 4
TFA	0.024 (0.036)	0.077 (0.037)**	0.091 (0.038)**	0.107 (0.039)***
KCTR	0.051 (0.069)	0.114 (0.068)*	0.113 (0.069)	0.148 (0.066)**
Lagged test scores, grade & year fixed effects	X	X	X	X
Student-level controls		X	X	X
School-level controls			X	X
Teacher experience controls (bins)				X
R-squared	0.580	0.589	0.593	0.594
N (student-year observations)	185284	185284	185284	185284
N (TFA teachers)	146	146	146	146
N (KCTR teachers)	20	20	20	20
N (non-program teachers)	1953	1953	1953	1953

Note: Models control for teacher experience using indicators for the following experience bins as reported in the main text: (1) 0 years prior experience (omitted), (2) 1-2 years, (3) 3-5 years, (4) 6-12 years, (5) 13-20 years, (6) 21-27 years, and (7) 28+ years. Standard errors clustered by teacher are reported in parentheses. The teacher counts reported at the bottom of the table indicate the number of unique teachers (i.e., clusters). *** p<0.01, ** p<0.05, * p<0.10.

Table B.11. Value-added to Student Achievement by Program, Grades 4-8, ELA.

	Model 1	Model 2	Model 3	Model 4
TFA	-0.031 (0.014)**	-0.002 (0.014)	0.012 (0.013)	0.027 (0.013)**
KCTR	0.009 (0.044)	0.030 (0.036)	0.031 (0.026)	0.051 (0.028)*
Lagged test scores, grade & year fixed effects	X	X	X	X
Student-level controls		X	X	X
School-level controls			X	X
Teacher experience controls (bins)				X
R-squared	0.665	0.673	0.676	0.676
N (student-year observations)	186614	186614	186614	186614
N (TFA teachers)	147	147	147	147
N (KCTR teachers)	24	24	24	24
N (non-program teachers)	2178	2178	2178	2178

Note: Models control for teacher experience using indicators for the following experience bins as reported in the main text: (1) 0 years prior experience (omitted), (2) 1-2 years, (3) 3-5 years, (4) 6-12 years, (5) 13-20 years, (6) 21-27 years, and (7) 28+ years. Standard errors clustered by teacher are reported in parentheses. The teacher counts reported at the bottom of the table indicate the number of unique teachers (i.e., clusters). *** p<0.01, ** p<0.05, * p<0.10.

Table B.12. KCTR Teacher Retention Rates Compared to Other Teachers in the Region.

	KCTR	All Non-Program Teachers (First-year Teachers Only)	All Non-Program Teachers (First-year Teachers Only, District-year Weighted Average)
Kansas City area 1-year retention rate	96.77	76.24	75.48
Kansas City area 2-year retention rate	70.97	64.95	56.82
Missouri 1-year retention rate	96.77	78.02	76.44
Missouri 2-year retention rate	74.19	67.92	58.79

Table B.13. TFA Teacher Retention Rates Compared to Other Teachers in the Region.

	TFA	All Non-Program Teachers (First-year Teachers Only)	All Non-Program Teachers (First-year Teachers Only, District-year Weighted Average)
KC area 1-year retention rate	99.12	78.86	62.95
KC area 2-year retention rate	57.06	70.35	55.51
KC area 3-year retention rate	41.18	63.37	46.53
KC area 4-year retention rate	32.35	58.22	40.84
MO 1-year retention rate	99.12	80.94	65.04
MO 2-year retention rate	58.82	73.27	58.08
MO 3-year retention rate	43.24	66.78	49.49
MO 4-year retention rate	33.82	61.88	43.99

BIBLIOGRAPHY

Abraham, S., and Sun, L. (2018). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Available at SSRN: https://ssrn.com/abstract=3158747

American Association of Colleges of Nursing (AACN). (2019). Fact sheet: Nursing shortage. Retrieved at https://www.aacnnursing.org/News-Information/Fact-Sheets/Nursing-Shortage

American Nurses Association. Nursing workforce. Retrieved at https://www.nursingworld.org/practice-policy/workforce/

Anderson, G., Sun, J. C., and Alfonso, M. (2006). Effectiveness of statewide articulation agreements on the probability of transfer: A preliminary policy analysis. *The Review of Higher Education*, 29(3), 261-291.

Arcidiacono, P., Aucejo, E. M., and Spenner, K. (2012). What happens after enrollment? An analysis of the time path of racial differences in GPA and major choice. *IZA Journal of Labor Economics*, 1(1), 5.

Atkinson, R. (2013). A short and long-term solution to America's STEM crisis. *The Hill*, 11.

Backes, B., Hansen, M., Xu, Z., and Brady, V. (2019). Examining spillover effects from Teach for America corps members in Miami-Dade County Public Schools. *Journal of Teacher Education*, 70(5), 453-471.

Baker, R. (2016). The effects of structured transfer pathways in community colleges. *Educational Evaluation and Policy Analysis*, 38(4), 626-646.

Baker, R., Bettinger, E., Jacob, B., and Marinescu, I. (2018). The effect of labor market information on community college students' major choice. *Economics of Education Review*, 65, 18-30.

Bates, M. (2020). Public and private employer learning: Evidence from the adoption of teacher value added. *Journal of Labor Economics*, 38(2), 375-420.

Baum, S., and Holzer, H. (2017). Do too many community college students major in liberal arts?. *Urban Institute*. Retrieved at https://www.urban.org/urban-wire/do-too-many-community-college-students-major-liberal-arts

Beede, D., Julian, T., Khan, B., Lehrman, R., McKittrick, G., Langdon, D., and Doms, M. (2011). Education supports racial and ethnic equality in STEM. ESA Issue Brief# 05-11. *US Department of Commerce*.

Bettinger, E. P., and Long, B. T. (2005). Do faculty serve as role models? The impact of instructor gender on female students. *American Economic Review*, 95(2), 152-157.

Boatman, A., and Soliz, A. (2018). Statewide transfer policies and community college student success. *Education Finance and Policy*, 13(4), 449-483.

Bottia, M. C., Stearns, E., Mickelson, R. A., Moller, S., and Valentino, L. (2015). Growing the roots of STEM majors: Female math and science high school faculty and the participation of students in STEM. *Economics of Education Review*, 45, 14-27.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., and Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy*, 1(2), 176-216.

Boyd, D., Lankford, H., Loeb, S., and Wyckoff, J. (2005). The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, 24(1), 113-132.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3), 414-427.

Canes, B. J., and Rosen, H. S. (1995). Following in her footsteps? Faculty gender composition and women's choices of college majors. *ILR Review*, 48(3), 486-504.

Carrell, S. E., Page, M. E., and West, J. E. (2010). Sex and science: How professor gender perpetuates the gender gap. *The Quarterly Journal of Economics*, 125(3), 1101-1144.

Chen, X. (2013). STEM attrition: College students' paths into and out of STEM fields. Statistical Analysis Report. NCES 2014-001. *National Center for Education Statistics*.

Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources*, 41(4), 778-820.

Cullen, J. B., Koedel, C., and Parsons, E. (forthcoming). The compositional effect of rigorous teacher evaluation on workforce quality. *Education Finance and Policy*.

Darolia, R., Koedel, C., Main, J. B., Ndashimye, J. F., and Yan, J. (2020). High school course access and postsecondary STEM enrollment and attainment. *Educational Evaluation and Policy Analysis*, 42(1), 22-45.

Decker, P., Mayer, D. P., and Glazerman, S. (2004). The effects of Teach for America on students: Findings from a national evaluation. *Princeton, NJ: Mathematica Policy Research*.

Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of Economics and Statistics*, 86(1), 195-210.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter?. *American Economic Review*, 95(2), 158-165.

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human resources*, 42(3), 528-554.

Deming, D., Goldin, C., and Katz, L. (2013). For-profit colleges. *The Future of Children*, 137-163.

Donaldson, M. L., and Johnson, S. M. (2011). Teach for America teachers: How long do they teach? Why do they leave? *Phi Delta Kappan*, 93(2), 47-51.

Egalite, A. J., and Kisida, B. (2017). The effects of teacher match on students' academic perceptions and attitudes. *Educational Evaluation and Policy Analysis*, 40(1), 59-81.

Egalite, A. J., Kisida, B., and Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, 45, 44-52.

Ehrenberg, R. G., and Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from high school and beyond. *Economics of education review*, 13(1), 1-17.

Ehrenberg, R. G., Goldhaber, D. D., and Brewer, D. J. (1995). Do teachers' race, gender, and ethnicity matter? Evidence from the National Educational Longitudinal Study of 1988. *ILR Review*, 48(3), 547-561.

Freyaldenhoven, S., Hansen, C., and Shapiro, J. M. (2019). Pre-event trends in the panel event-study design. *American Economic Review*, 109(9), 3307-38.

Gershenson, S., Holt, S. B., and Papageorge, N. W. (2016). Who believes in me? The effect of student-teacher demographic match on teacher expectations. *Economics of Education Review*, 52, 209-224.

Glazerman, S., Protik, A., Teh, B. R., Bruch, J., and Max, J. (2013). Transfer incentives for high-performing teachers: Final results from a multisite randomized experiment. NCEE 2014-4004. *National Center for Education Evaluation and Regional Assistance*.

Goldhaber, D., Cowan, J., and Walch, J. (2013). Is a good elementary teacher always good? Assessing teacher performance estimates across subjects. *Economics of Education Review*, 36(1), 216-228.

Goldhaber, D., Quince, V., and Theobald, R. (2018). Has it always been this way? Tracing the evolution of teacher quality gaps in U.S. public schools. *American Educational Research Journal*, 55(1), 171-201.

Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing (No. w25018). *National Bureau of Economic Research*.

Gross, B., and Goldhaber, D. (2009). Can transfer and articulation policies propel community college students to a bachelor's degree--and is this the only goal?. *Center on Reinventing Public Education, University of Washington Bothell*.

Gross, B., and Goldhaber, D. (2009). Community college transfer and articulation policies: Looking beneath the surface. CRPE Working Paper 2009-1. *Center on Reinventing Public Education*.

Guha, R., Hyler, M. E., and Darling-Hammond, L. (2016). The teacher residency: An innovative model for preparing teachers. *Learning Policy Institute*.

Hoxby, C., and Turner, S. (2013). Expanding college opportunities for high-achieving, low income students. *Stanford Institute for Economic Policy Research Discussion Paper*, (12-014).

Holt, S. B., and Gershenson, S. (2019). The impact of demographic representation on absences and suspensions. *Policy Studies Journal*, 47(4), 1063-1093.

Institute of Medicine (US). Committee on the Robert Wood Johnson Foundation Initiative on the Future of Nursing. (2011). The future of nursing: Leading change, advancing health. *Washington, DC: National Academies Press*.

Jaschik, S. (2015). Community college liberal arts. *Inside Higher Ed*. Retrieved at https://www.insidehighered.com/news/2015/01/20/data-show-steady-growth-humanities-and-liberal-arts-education-community-colleges

Kane, T. J., Rockoff, J. E., and Staiger, D. O. (2008). What does teacher certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, 27(6), 615-631.

Koedel, C., Mihaly, K., and Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review*, 47, 180-195.

Koedel, C., and Parsons, E. (2020). The effect of the community eligibility provision on the ability of free and reduced-price meal data to identify disadvantaged students. *CALDER Working Paper*, No. 234-0320.

Koedel, C., Parsons, E., Podgursky, M., and Ehlert, M. (2015). Teacher preparation programs and teacher quality: Are there real differences across programs?. *Education Finance and Policy*, 10(4), 508-534.

Kofoed, M. S., and McGovney, E. (2019). The effect of same-gender or same-race role models on occupation choice evidence from randomly assigned mentors at West Point. *Journal of Human Resources*, 54(2), 430-467.

Lankford, H., Loeb, S., and Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, 24(1), 37-62.

Lefgren, L., and Sims, D. P. (2012). Using subject test scores efficiently to predict teacher value-added. *Educational Evaluation and Policy Analysis*, 34(1), 109-121.

Lim, J., and Meer, J. (2017). The impact of teacher-student gender matches: Random assignment evidence from South Korea. *Journal of Human Resources*, 52(4), 979-997.

Lindsay, C. A., and Hard, C. M. D. (2017). Exposure to same-race teachers and student disciplinary outcomes for black students in North Carolina. *Educational Evaluation and Policy Analysis*, 39(3), 485-510.

National Association for College Admission Counseling. (2016). The transfer process defined. Retrieved at https://www.nacacnet.org/knowledge-center/transfer/the-transfer-process-defined/

National Girls Collaborative Project. (2015). State of girls and women in STEM. *Ngcproject.org*.

National Science Board. (2016). Science and engineering indicators 2016. *Arlington, VA: National Science Foundation* (NSB-2016-1).

Noonan, R. (2017). STEM jobs: 2017 update (ESA Issue Brief# 02-17). *Washington, DC: US Department of Commerce, Economics and Statistics Administration, Office of the Chief Economist*.

Papageorge, N., Gershenson, S., and Kang, Kyungmin. (2020). Teacher expectations matter. *Review of Economics and Statistics*, 102(2), 234-251.

Papay, J. P., Bacher-Hicks, A., Page, L. C., and Marinell, W. H. (2017). The challenges of teacher retention in urban schools: Evidence of variation from a cross-site analysis. *Educational Researcher*, 46(8), 434-448.

Papay, J. P., West, M. R., Fullerton, J. B., and Kane, T. J. (2012). Does an urban teacher residency increase student achievement? Early evidence from Boston. *Educational Evaluation and Policy Analysis*, 34(4), 413-434.

Parsons, E., Koedel, C., and Tan, L. (2019). Accounting for student disadvantage in value-added models. *Journal of Educational and Behavioral Statistics*, 44(2), 144-179.

Price, J. (2010). The effect of instructor race and gender on student persistence in STEM fields. *Economics of Education Review*, 29(6), 901-910.

Roksa, J., and Keith, B. (2008). Credits, time, and attainment: Articulation policies and success after transfer. *Educational Evaluation and Policy Analysis*, 30(3), 236-254.

Rose, H., and Betts, J. R. (2004). The effect of high school courses on earnings. *Review of Economics and Statistics*, 86(2), 497-513.

Roth, J. (2018). Pre-test with caution event-study estimates after testing for parallel trends. *Working Paper*.

Rouse, C. E. (1995). Democratization or diversion? The effect of community colleges on educational attainment. *Journal of Business & Economic Statistics*, 13(2), 217-224.

Sass, T. R. (2015). Licensure and worker quality: A comparison of alternative routes to teaching. *The Journal of Law and Economics*, 58(1), 1-35.

Shapiro, D., Dundar, A., Huie, F., Wakhungu, P. K., Yuan, X., Nathan, A., and Hwang, Y. (2017). Tracking transfer: Measures of effectiveness in helping community college students to complete bachelor's degrees. *Signature Report*, (13).

Springer, M. G., Swain, W. A., and Rodriguez, L. A. (2016). Effective teacher retention bonuses: Evidence from Tennessee. *Educational Evaluation and Policy Analysis*, 38(2), 199-221.

Swain, W. A., Rodriguez, L. A., and Springer, M. G. (2019). Selective retention bonuses for highly effective teachers in high-poverty schools: Evidence from Tennessee. *Economics of Education Review*, 68, 148-160.

Tennessee Higher Education Commission. (2014). 2014 report card on the effectiveness of teacher training programs. *Tennessee Department of Education*.

Torpey, E. (2018). Employment outlook for bachelor's-level occupations. *Career Outlook, U.S. Bureau of Labor Statistics*.

U.S. Congress Joint Economic Committee. (2012). STEM education: Preparing for the jobs of the future. *Washington DC*.

U.S. Department of Veterans Affairs. (2013). Minneapolis VA and University of Minnesota to educate more nurses. Retrieved at https://www.minneapolis.va.gov/features/Minneapolis-VA-and-University-of-Minnesota-Educate-Nurses.asp

Wiswall, M. (2013). The dynamics of teacher quality. *Journal of Public Economics*, 100, 61-78.

Xu, Z., Hannaway, J., and Taylor, C. (2011). Making a difference? The effects of Teach for America in high school. *Journal of Policy Analysis and Management*, 30(3), 447-469.

VITA

Yang An was born and grew up in Dalian, China. He earned a bachelor's degree in 2011 from Liaoning University and master's degrees in Economics and Mathematics in 2014 from Ohio University. He earned a doctoral degree in Economics in 2020 from the University of Missouri-Columbia.