Neighborhood Effects on STEM Major Choice

Jeonghyeok Kim, Rohit Munshi, and Brian Murphy*

September 8, 2025

Click here for the latest version

Abstract

This paper provides causal evidence that the neighborhoods where students grow up play a significant role in shaping their college major choices, focusing on STEM fields. Using administrative data from Texas and variation in the timing of school moves across counties and districts, we estimate the impact of neighborhood exposure on the likelihood of pursuing a STEM major. We find that students who move to high-STEM neighborhoods—defined by the share of non-moving peers who earn STEM degrees—are increasingly likely to major in STEM with each additional year of exposure. We also show that neighborhood STEM exposure is strongly tied to the local occupational landscape, especially the concentration of residents working in STEM fields, with the highest-STEM areas clustered around major research and technology hubs. This suggests that exposure to local STEM careers is a key mechanism behind the observed effects. Mediation analysis further reveals that the effects operate primarily through behavioral pathways—specifically, increased enrollment in advanced science/math coursework—rather than improvements in academic performance. Importantly, the benefits of STEM-rich neighborhoods extend to underrepresented groups, including students from economically disadvantaged backgrounds, women, and racial minorities. These findings underscore the critical role of neighborhood environments in shaping educational pathways and highlight their importance in addressing educational inequality and strengthening the STEM pipeline.
JEL: I24, J24, R23, J15
Keywords: STEM, neighborhoods, college major choice

*Jeonghyeok Kim: Department of Economics, University of Houston, Houston, Texas, USA (jkim124@uh.edu); Rohit Munshi: Department of Economics, University of Houston, Houston, Texas, USA (rmunshi@cougarnet.uh.edu); Brian Murphy: Department of Economics, University of Houston, Houston, Texas, USA (bmmurphy2@uh.edu). We are grateful to Aimee Chin and Chinhui Juhn for their invaluable guidance and support. For helpful comments, we thank Blake Heller, Fan Wang, Emily Hannum, Yona Rubinstein, Elaine Liu, Vikram Maheshri, Willa Friedman, Steven Craig, Saerom Ahn, Gergely Ujhelyi and participants at the University Houston Micro Workshop. We would also like to thank Jeanette Narvaez and Catherine Horn from the Education Research Center for their help while accessing the data. The research presented here uses confidential data from the State of Texas supplied by the Education Research Center (ERC) at The University of Houston (UH 080). All errors are our own. The conclusions of this research do not necessarily reflect the opinion or official position of Texas Education Research Center, the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission, or the State of Texas.

1 Introduction

Science, Technology, Engineering, and Mathematics (STEM) fields are increasingly central to national competitiveness, economic growth, and social mobility. As the U.S. economy becomes more technology-driven, the demand for workers trained in STEM disciplines continues to grow, raising widespread concerns about the adequacy of the nation's STEM pipeline (Carnevale, Smith, and Melton 2011; Langdon et al. 2011). A series of federal reports have identified persistent shortages of qualified STEM professionals as a potential threat to both economic prosperity and national security (Holden and Lander 2012; National Science and Technology Council 2000; National Science Board 2003, 2024; Xue and Larson 2015). Alongside these concerns, long-standing disparities in STEM participation persist (Maltese and Tai 2011; Rozek et al. 2019; Xie, Fang, and Shauman 2015). Black and Hispanic individuals account for 28% of the U.S. workforce, yet hold only 17% of STEM jobs, and women account for just a quarter or less of the workforce in computing and engineering sectors (Fry, Kennedy, and Funk 2021). These disparities not only limit access to economic opportunity but also constrain the nation's ability to draw on its full talent pool in addressing complex technological challenges.

A substantial body of research has sought to explain these STEM choices and disparities by examining individual and family-level determinants of students' educational and career choices. Differences in ability, preferences, expectations, and early academic preparation have all been found to influence the decision to pursue a STEM degree (Altonji, Arcidiacono, and Maurel 2016; Patnaik, Wiswall, and Zafar 2021). Other studies have highlighted the importance of schools and educational environments, including exposure to advanced coursework, teacher quality, and peer influences (Beilock et al. 2010; Clotfelter, Ladd, and Vigdor 2005; Lavy and Sand 2015; Legewie and DiPrete 2014; Wang 2013). However, far less attention has been given to the role of neighborhoods as determinants of students' educational trajectories, particularly their field of study in college.

This paper investigates whether, and to what extent, the neighborhoods in which students grow up shape their likelihood of obtaining a STEM degree. We use large-scale administrative data from Texas, focusing on students who enrolled in kindergarten (grade 0) between 1994 and 2003, which allows us to track their educational trajectories from early childhood through college. We present new quasi-experimental evidence on the effects of neighborhoods on college STEM choices, which we defined as attaining an undergraduate degree in STEM majors. Our empirical strategy builds on the approach developed by Chetty and Hendren (2018), which shows that the neighborhoods in which children grow up shape their adulthood outcomes, including earnings and college attendance rates. This paper extends this by going beyond better-worse comparisons to the characteristics of neighborhoods. In particular, we explore whether neighborhood exposure to STEM fields shapes the choice of field of study.

We begin by documenting disparities in STEM choice across student groups and geographic areas. Black, Hispanic, and economically disadvantaged students are less likely to earn STEM degrees compared to White, non-disadvantaged peers. While differences in academic ability—measured by math scores—explain part of the gap, a substantial portion remains even after adjusting for test score distributions. We also find wide variation in STEM attainment across counties, with the highest rates concentrated near major tech and research hubs like Austin and NASA Space Center. These patterns are stable over time and persist even when limiting the sample to undergraduate degree holders. Together, these findings suggest that contextual factors may play an important role in shaping students' STEM pathways.

We focus on students who move across counties and school districts during their school years to estimate the causal effects of neighborhood exposure on STEM major choice. By comparing students who move in earlier versus later grades, we capture the cumulative effects of exposure to neighborhoods with different STEM rates. We define STEM rates as the share of peers from the same racial and economic background who major in STEM, measured among those who never move across neighborhoods (non-movers). This measure reflects the idea that students' decisions to pursue STEM degrees aggregate the various influences they encounter in their local environments. Our main finding is that the length of time students spend in high-STEM neighborhoods has a substantial and linear effect on their likelihood of choosing a STEM major in college. Specifically, each additional year spent in a county with a 1 percentage point higher STEM rate increases the probability of majoring in STEM by approximately 0.03 percentage points.

The findings suggest that neighborhoods have a meaningful influence on students' STEM major choices. The estimates indicate that about 40% of the variation in STEM major choices across different counties can be attributed to the causal effects of place in their school years. For example, a student who moves in kindergarten from a county at the 25th percentile (2.7%) to one at the 75th percentile (4.4%)—a 1.02 standard deviation difference—would experience a 0.67 percentage point increase in their probability of majoring in STEM. This effect amounts to roughly 16% of the average STEM major attainment rate (4.18%).

To interpret the 3% convergence rate as a causal effect of neighborhood exposure, we assume that the likelihood of majoring in STEM does not systematically vary by the grade at which students move. We assess this assumption using several tests: (i) overidentification tests using different student groups and majors, (ii) restricting the comparison group to students from highly similar contexts, (iii) analyses based on displacement shocks and predicted destinations, and (iv) coefficient stability tests with respect to unobservables. The results from these tests support the conclusion that the 3% annual convergence reflects a causal effect of neighborhood exposure on STEM major choice.

We investigate one important neighborhood dimension underlying the convergence in college major choices. Our primary measure of exposure—the STEM major rate among non-movers—captures the combined influence of neighborhood characteristics. Because non-movers remain in the same community throughout schooling, their college major choices reflect the cumulative impact of local conditions, including peer environments, schools, and neighborhood labor markets. To examine one important channel more directly, we focus on the local occupational composition, specifically, the share of STEM workers in a neighborhood, as a measure for the non-school environment and examine its relationship to students' college major choices. We find a strong correlation between the proportion of STEM occupations in a neighborhood and the proportion of students majoring in STEM, regardless of whether the sample is restricted to bachelor's degree holders. In contrast, we do not observe a similar relationship for other majors. We then replicate the baseline analysis using the STEM occupation composition of the county as the exposure measure. We find consistent neighborhood effects of STEM occupation exposure. Furthermore, an overidentification test shows that this effect is specific to the gender of the student and STEM workers, highlighting the importance of same gender exposure for their future college major. These findings shed light on the non-academic channels through which place matters. They suggest that beyond schools, the local labor market—through its composition and representation—can shape students' perceptions of what career paths are possible or attainable.

An important question that remains is whether high-STEM neighborhoods also encourage underrepresented groups to major in STEM. For example, are students from economically disadvantaged families more likely to pursue STEM majors if they move to STEM-oriented areas, even when their parents do not work in STEM fields? This question is critical to understanding persistent disparities in STEM participation across racial, gender, and socioeconomic lines. To address this, we examine heterogeneity in exposure effects by economic disadvantage status, gender, and race. We find that living in a high-STEM neighborhood increases the likelihood of choosing a STEM major across all groups, including those traditionally underrepresented in STEM. However, the level of exposure is not evenly distributed. Students from disadvantaged backgrounds or underrepresented racial groups are less likely to live in high-STEM neighborhoods, limiting the cumulative benefit they receive from such environments.

To explore the consequences of this unequal exposure, we conduct a simple counterfactual analysis estimating how much the STEM major gaps would narrow if all students had equal neighborhood exposure. White students experience about 30–65% higher STEM exposure than Black and Hispanic students. Equalizing exposure would close about 20–40% of the racial gap in STEM major rates. Similarly, about 20% of the STEM gap between economically disadvantaged and non-disadvantaged students would be closed. For gender, although male and female students live in similar neighborhoods, women have less exposure to female STEM role models. Simulating equal gender-specific exposure would reduce the gender gap by about 10%.

Lastly, we explore the mechanisms underlying the exposure effects—specifically, whether they operate through academic performance or behavioral channels. While students in high-STEM neighborhoods are more likely to take advanced science and math courses and achieve higher math scores, our mediation analysis shows that the increased STEM major selection is primarily influenced by greater enrollment in advanced science/math courses (accounting for 30% of the effect), rather than by improvements in test scores. This suggests that neighborhood exposure influences STEM engagement mainly through behavioral responses rather than gains in academic performance. This finding aligns with prior research highlighting the importance of academic preparedness for STEM choices (Justman and Méndez 2018; Maltese and Tai 2011; Sadler et al. 2014; Speer 2023), suggesting that preparedness is not only a source of group differences in STEM but also a consequence of the environments students experience. Taken together, our results underscore the potential of place-based policies—such as expanding access to advanced coursework and connecting students with STEM professionals—to promote STEM participation and reduce disparities.

In this paper, we contribute to three strands of the literature (see Appendix B for a more detailed review of the literature). First, we offer a novel perspective on the factors that influence college (STEM) major choice. Existing studies emphasize preferences (Beffy, Fougère, and Maurel 2012; Wiswall and Zafar 2015), expectations (Arcidiacono et al. 2020; Stinebrickner and Stinebrickner 2014), family background (Hanushek et al. 2023), and school characteristics (Arold 2024; Elsner, Isphording, and Zölitz 2021). In their review of the literature on STEM education, Xie, Fang, and Shauman (2015) point out that "little is known about other potential contextual factors, such as local labor market characteristics or proximity to science-focused industry." Building on this observation, this paper extends the literature by introducing the concept of contextual factors that causally promote students' engagement with STEM education and by identifying the occupational composition of neighborhoods as a key mechanism underlying neighborhood effects.

The second contribution is to shed light on gender and racial gaps in STEM (Xie, Fang, and Shauman 2015). Previous literature has identified multiple channels contributing to these gaps, including abilities or preparedness (Arcidiacono, Aucejo, and Hotz 2016; Ellison and Swanson 2010), preferences (Ngo and Dustan 2024; Wiswall and Zafar 2018), and school environments (Carrell, Page, and West 2010; Landaud, Ly, and Maurin 2020; Philippis 2023). We add another important dimension—neighborhoods—which have received limited attention. We find that the neighborhoods students grow up in during their school years significantly influence their likelihood of choosing STEM majors, regardless of gender, race, or socioeconomic background. Our counterfactual analysis shows that a substantial portion of the gaps between groups could be narrowed if students had similar neighborhood experiences during their formative years.

Lastly, this study contributes to the understanding of developmental neighborhood effects on children by emphasizing the multidimensional nature of neighborhoods. A large body of research has documented how the neighborhood characteristics correlate with a wide range of long-term outcomes (Brooks-Gunn et al. 1993; Jencks, Mayer, et al. 1990; Sampson, Morenoff, and Gannon-Rowley 2002; Sharkey and Faber 2014). More recent studies that leverage quasi-experimental variation provide causal evidence that exposure to better neighborhoods during childhood can have substantial positive impacts (Chetty and Hendren 2018; Chetty, Hendren, and Katz 2016; Chyn 2018; Nakamura, Sigurdsson, and Steinsson 2022). Building on the empirical strategy of Chetty and Hendren (2018), we show that neighborhood environments also shape students' educational trajectories and, potentially, their career paths, primarily through influencing behavior rather than improving academic performance. In doing so, we extend the findings of Bell et al. (2019), who show that neighborhood exposure to inventors is crucial for fostering future inventors. Our work highlights how the neighborhood context can influence a broader set of academic choices, underscoring the importance of local environments in shaping children's aspirations and opportunities.

2 Data

We use individual-level Texas administrative data sets, which include following sources: K–12 education records from the Texas Education Agency (TEA), which cover public school students beginning in the 1994–1995 academic year and allow us to track students' neighborhoods; and all public and most private post-secondary education data in Texas from the Texas Higher Education Coordinating Board (THECB)[1], allowing us to construct the individual STEM choice and neighborhood STEM exposure measures.

One limitation of the THECB data is that it only captures postsecondary outcomes for students attending college within Texas. Thus, if a student moves out of state, we cannot determine

[1] The THECB data contain all public community and technical colleges; all public universities and health-related institutions; almost all independent colleges and universities (available from 2003 onward); and career schools and colleges (available from 2004 onward). See http://www.txhighereddata.org/Interactive/CBMStatus/ for additional information on participating institutions.

whether they attended college elsewhere or did not attend college at all. However, this issue is unlikely to introduce significant bias, given the relatively low rate of out-migration from Texas (Foote and Stange 2022) and our empirical strategies. To further address this limitation, we supplement our analysis using data from the National Student Clearinghouse (NSC) for Texas high school graduating cohorts beginning in 2009. Although this dataset does not perfectly overlap with our primary sample, it allows us to examine whether incorporating out-of-state college attendance meaningfully changes our findings. Our analyses using the NSC data show minimal impact on the estimates. Detailed information on the NSC data and estimation results is provided in Appendix C.

The analysis sample includes all students who started kindergarten (grade 0) between 1994-2003.[2] Students also need to be observed for at least 10 school years to be included (Appendix Table A.1 presents the robustness of the baseline estimate to the year restriction). We divide the sample into two parts: non-movers and movers. We define the non-movers of each county p as the subset of students who reside in a single county in all school years. The movers sample consists of students who move from one county to another. In our baseline analysis, we focus on the subset of individuals in counties where the number of non-mover students of the same economic disadvantage status and race exceeds 2,000 (i.e., an average of 200 students per school cohort) to ensure sufficiently large samples for precisely estimating the outcomes of non-moving students, which are the key independent variables in our analysis (see Appendix Figure A.1 for robustness to alternative enrollment size cutoffs). Racial groups are categorized as Black, Hispanic, and White; students of other racial backgrounds are excluded due to their small sample size. The economic disadvantage status is determined by eligibility for free or reduced-price lunch. The baseline sample includes 1,781,130 students, of whom 1,575,342 are non-moving students.

STEM choice is defined as obtaining an undergraduate degree in STEM majors from institutions in Texas. STEM majors are classified following the STEM major list from the Department of Homeland Security in 2016 (see Appendix J for the full list of STEM majors). Among STEM degrees awarded in Texas between 1994 and 2024, 51% are in engineering (CIP code

[2] While kindergarten was not mandatory in Texas during the sample year, kindergarten enrollment was about 95% of first grade enrollment.

14), 23% in biological and biomedical sciences (26), 11% in mathematics and statistics (27) or physical sciences (40), with the remainder in other STEM fields. Our baseline measure of STEM exposure is defined as among non-moving students in a county (p) by group of race and economic status (g), the proportion of students who obtain an undergraduate degree in STEM majors by age 26:

$$\overline{STEM}_{pg} = \frac{\text{STEM Degree}_{pg}}{\text{All students}_{pg}}.$$

This measure reflects the idea that students' decisions to pursue STEM degrees aggregate the various influences they experience in their local environments, including factors such as peer aspirations, school resources, teacher quality, and community attitudes toward STEM. By focusing on non-moving students, we capture the sustained exposure to neighborhood characteristics.

Table 1 presents the summary statistics for the analysis sample. The first two columns show statistics for non-moving students who live in counties with an average cohort size of more than 2000 non-moving students of the same economic status and race. The next two columns focus on students who moved once between counties. Although our analysis does not require movers to be comparable to non-movers, we find that non-moving students have slightly higher educational attainment than students who moved once. For instance, the average STEM rate is 3.9% for non-moving students, compared to 3.8% for one-time movers.

3 STEM Choice Across Student Groups and Places

We begin by presenting gaps in STEM degree attainment by race and economic disadvantage status, measured by eligibility for free or reduced-price lunch (FRL). Figure 1 (a) shows the share of students earning a STEM degree between 1994 and 2003, broken down by race and economic status. Two patterns emerge: students from racial minority groups and those who are economically disadvantaged are underrepresented in STEM majors. While 6% of White students who are not economically disadvantaged earn a STEM degree, only 1–2% of disadvantaged students do so, regardless of race. Even among non-disadvantaged students, the rates are lower for Black and Hispanic students: about 3% of Black and 4% of Hispanic students in this group earn a STEM degree.

To assess the role of baseline ability in these gaps, we control for third-grade math scores by non-parametrically reweighting each group's math score distribution to match that of White students who are not economically disadvantaged, following the methodology of DiNardo, Fortin, and Lemieux (1996). The second set of bars in Figure 1 (a) presents the resulting reweighted STEM attainment rates. While this adjustment reduces the magnitude of the gaps, it does not eliminate them. Substantial differences remain compared to non-disadvantaged White students. We find similar patterns when comparing male and female students in Appendix Figure A.2.

Figure 1 (b) further examines how much of the STEM attainment gap between economically disadvantaged and non-disadvantaged students can be explained by differences in math test scores as students advance through school. A clear upward trend emerges. As grade level increases, test scores account for a growing share of the gap. By 8th grade, approximately 42% of the gap is attributable to differences in test performance. A linear regression across grades 3 through 8 indicates that, on average, an additional 3.7 percentage points of the gap is explained by test scores with each advancing grade. Extrapolating from this trend, up to 57% of the gap could potentially be explained by test scores by 12th grade. Nonetheless, a substantial portion of the gap remains unexplained. Importantly, these test score differences may not be innate or solely the result of family background. Rather, they likely reflect the cumulative impact of contextual and environmental influences.

To better understand what accounts for this residual gap, we turn to geographic variation by presenting the STEM rates of students who attended schools in the same county throughout their school years (non-movers). Figure 2 (a) illustrates the average STEM rates by county, with the counties divided into seven quantiles based on their $\overline{STEM}_p$ values. Darker shades indicate counties with higher average STEM rates. There is significant variation in STEM attainment across counties. Importantly, the areas with the highest STEM rates are located near major research and tech hubs.[3] The pattern is persistent even when we only consider students who earned an undergraduate degree, seen in Figure 2 (b). The correlation between county-level

[3] For example, the top five school districts with the highest STEM rates are near the NASA Space Center (Friendswood, Clear Creek) or in the Austin area (Eanes, Lake Travis, Round Rock).

STEM rates among all students and undergraduate degree holders is 0.66. Thus, the variation in STEM choice across counties is not just because of differences in general educational attainment.

Moreover, the variation in STEM rates across different counties is notably persistent over time. Figures 2 (c) and (d) compare STEM rates for counties between the 1994-1997 and 2001-2003 cohorts of non-moving students. Counties with high STEM rates in earlier cohorts tend to have higher STEM rates in later cohorts as well. This persistence is evident even when the analysis is limited to students with an undergraduate degree. This enduring trend could be attributed to two potential factors: neighborhood effects influencing students' field of study choices or specific types of families (or students) gravitating towards areas where STEM majors are more prevalent. In the following analysis, we explore how much of this difference across neighborhoods is attributed to the causal effects of neighborhoods on STEM choices.

4 Estimating Neighborhood Exposure Effects on STEM Choice

We conceptualize "neighborhood" effects as the combined influence of various factors, including schools, teachers, peers, neighbors, and local workers. In the main text, we focus on variation across counties, and in Appendix D, we replicate the analysis at the school district level, finding consistent results.

We aim to assess the potential increase in student STEM choices if they were raised in environments with high STEM exposure. To investigate this, we analyze students who move between counties with varying STEM rates in different grades to estimate the impact of exposure to neighborhoods. We generally follow Chetty and Hendren (2018) for estimation designs, which has been widely applied to recover causal effects in related contexts (Baran, Chyn, and Stuart 2024; Chyn, Collinson, and Sandler 2025; Deutscher 2020; Finkelstein, Gentzkow, and Williams 2021; Laliberté 2021; Nakamura, Sigurdsson, and Steinsson 2022; Schwank 2024).

First, consider the cohort of students who move once at grade m, from grades 1 to 11 (we obtain consistent results for students who move across counties multiple times, as shown in Appendix E). We analyze the relationship between these students' STEM choices in college and those of non-moving students in their origin and destination counties by using the following regression model:

$$STEM_i = \alpha_{osgm} + \sum_{m=1}^{11} b_m I(m_i = m)\Delta STEM_{odg} + \eta X_{odg} + \varepsilon_{1i}, \quad (1)$$

where $STEM_i$ denotes the students' STEM undergraduate degree attainment by age 26; α_{osgm} is a fixed effect for the interaction of origin county o, school cohort s, racial–economic status group g (Black, White, or Hispanic, crossed with economically disadvantaged or not), and grade at move m; $I(m_i = m)$ is an indicator function that is 1 when $m_i = m$ and 0 otherwise; and $\Delta STEM_{odg} = \overline{STEM}_{dg} - \overline{STEM}_{og}$ represents the difference in STEM attainment between the destination and origin counties, based on non-moving students from the same racial and economic status group. X_{odg} denotes a control vector including the undergraduate degree attainment rate of the origin and destination for the same racial and economic status group.

The coefficients b_m are the coefficients of interest, which capture the effect of moving at grade m, allowing the length of exposure to vary.[4] We assume that selection into destination is not systematically related to the timing of the move. That is, while more advantaged or education-oriented families may be more likely to relocate to high-STEM areas, they do not disproportionately do so when their children are younger. Under this assumption, we can identify the causal effect of neighborhood exposure using observational data: if neighborhoods have a causal impact, the effect should be stronger the earlier the exposure occurs since the duration of exposure is longer. We define the marginal exposure effect at grade m as $\gamma_m = b_m - b_{m+1}$, which estimates the impact of spending one additional year in the destination neighborhood at grade m. Given our identifying assumption, this provides an unbiased estimate of the neighborhood exposure effect (see Appendix F for a detailed discussion of the identifying assumption).

Before presenting the estimation results, we first construct a non-parametric binned scatter plot for students who moved in grade 4. After demeaning both $STEM_i$ and $\Delta STEM_{odg}$ within each origin (o) by school cohort (s) by racial and economic status group (g), we divide the residuals into twenty equal-sized groups and plot the mean value for each bin. Figure 3 shows that students who move to areas where non-moving students are more (less) likely to choose

[4] To address concerns that grade retention might bias these estimates, we estimate the model using a sample of students who never repeated a grade and find consistent results in Appendix Figure A.3.

STEM degrees are more (less) likely to major in STEM degrees. Importantly, moving to a high- or low-STEM area has a symmetric but opposite impact. The relationship between $STEM_i$ and $\Delta STEM_{odg}$ appears linear. The estimated regression coefficient of $b = 0.605$ indicates that a 1 percentage point increase in $\Delta STEM_{odg}$ is associated with a 0.605 percentage point increase in $STEM_i$. While this relationship reflects both neighborhood effects and selection, it suggests that the environment in which students grow up plays an important role in shaping their college major choices.

Figure 4 (a) presents the estimates of b_m from equation (1). These estimates reveal important patterns. Notably, the values of b_m decrease consistently with the grade at which students move. Under the identification assumption that selection effects do not vary with the student's grade at move, this declining trend provides evidence of an exposure effect, indicating that moving to a higher STEM area earlier in a student's school years results in a larger increase in the probability of majoring in STEM. The linearity observed between b_m and the grade at move m suggests that the exposure effect, defined as γ_m, remains approximately constant with respect to the grade at move m. By regressing $\hat{b}_m$ on m, we estimate an average annual exposure effect of $\gamma = 0.032$ (s.e. = 0.003). This implies that the outcomes of students who move converge toward those of non-movers in the destination area at a rate of 3.2% per year of exposure during the school years. Additionally, the estimates of b in the higher grades approach 0.1 by eleventh grade, indicating that most of the positive effects observed in moving in early grades to high STEM areas can be attributed to exposure effects rather than selection effects.

Equation (1) includes over 20,000 fixed effects (α_{osgm}), which makes estimation challenging in later parts that require smaller samples. To overcome this difficulty, we estimate a model where we control for the STEM rates of the origin, rather than relying on origin fixed effects. This leads to the following regression:

$$
\begin{aligned}
STEM_i = &\sum_{m=1}^{11} b_m I(m_i = m)\Delta STEM_{odg} + \eta X_{odg} \\
&+ \sum_{m,s,g} I(m_i = m, s_i = s, g_i = g)(\alpha^1_{msg} + \alpha^2_{msg}\overline{STEM}_{og}) + \varepsilon_{2i}.
\end{aligned}
\tag{2}
$$

Figure 4 (b) presents the coefficients b_m from estimating equation (2). These coefficients closely resemble those shown in Figure 4 (a). By regressing the b_m coefficients on m, we estimate an average annual exposure effect of $\gamma = 0.031$ (s.e. = 0.002). This estimate is consistent with that obtained from the fixed effects specification, indicating that controlling for the STEM rates in the origin county from non-moving students sufficiently accounts for differences between origin counties. In other words, the STEM choices of movers can be effectively modeled as a weighted average of the STEM rates of non-moving students in both the origin and destination, with the number of school years spent in each place.

Considering approximately constant exposure effects with respect to the grade at move m, we parameterize the exposure effects shown in Figure 4 linearly, replacing the semi-parametric $\sum_{m=1}^{11} b_m I(m_i = m)\Delta STEM_{odg}$ term in equation (2) with a linear interaction:

$$\begin{aligned} STEM_i = &\gamma(12 - m_i)\Delta STEM_{odg} + \theta_1 \Delta STEM_{odg} + \theta_2 X_{odg} \\ &+ \sum_{m,s,g} I(m_i = m, s_i = s, g_i = g)(\alpha^1_{msg} + \alpha^2_{msg}\overline{STEM}_{og}) + \varepsilon_{3i}. \end{aligned} \tag{3}$$

Estimating this specification yields an average annual exposure effect of $\gamma = 0.030$ (s.e. = 0.004), as reported in column (1) of Table 2. This estimate remains robust across various specifications and sample definitions. Column (2) extends the baseline model by adding fixed effects for race and economic status interacted with the destination STEM rate. Column (3) adds further controls, including general labor market characteristics (such as employment rates, median income, and poverty rate) as well as third-grade reading and math scores of moving students. Column (4) incorporates additional fixed effects to the baseline race-by-economic status-by-origin catchment area-by-destination catchment area-by-year of move. Catchment areas are constructed based on observed student movement patterns across schools. We identify approximately 1,800 catchment areas in the sample period in Texas, with each student being compared to roughly two other students (see Appendix H.2 for more details). Columns (5) through (8) replicate the specifications in columns (1) through (4), replacing the origin and destination STEM rates with place fixed effects. Column (7) additionally adds origin county-by-destination county-by-race-by-economic status fixed effects instead of adding further controls.

While the linear specification fits the data well, Figure 4 also shows a steeper slope in higher grades, suggesting stronger effects of neighborhood exposure for older students. To assess whether exposure effects vary by the timing of the move, we explore heterogeneity across grade levels in Appendix Section G. The results indicate that exposure during high school years has a greater impact on STEM major choice than exposure in earlier grades: $\gamma_{g<8} = 0.16$ and $\gamma_{g\geq 8} = 0.66$. These patterns suggest that late school years may be particularly formative in shaping students' academic trajectories. This is consistent with the idea that high school is when students choose advanced coursework and form clearer career aspirations. Early exposure may help build a foundation, but exposure during high school is especially influential because it coincides with key decision-making years. These findings also support our identifying assumption that selection effects are relatively stable across grades: it is unlikely that families systematically time moves based on unobserved STEM-related factors in a way that would produce the differences we observe by grade. Thus, the stronger effects for older movers likely reflect greater sensitivity to STEM exposure rather than selection. While we do not adopt the grade-split in the main analysis due to sample size limitations, the results highlight the importance of exposure timing.

The findings suggest that neighborhoods play a substantial role in shaping students' decisions to pursue STEM majors in college. Based on our estimates, the cumulative effect of neighborhood exposure over a student's K–12 years—assuming a consistent annual effect of 0.03 percentage points—explains approximately 39% ($13 \times 3\%$) of the total variation in STEM major attainment across counties.[5] This highlights that almost half of the geographic differences in STEM outcomes can be attributed to the environments students are exposed to during their school years, rather than to differences in student or family characteristics. To illustrate the magnitude of this effect, consider a student who moves in kindergarten from a county-by-group ($\overline{STEM}_{pg}$) at the 25th percentile of the STEM exposure distribution (1.8%) to one at the 75th percentile (6.4%), representing a 1.3 standard deviation increase. This student would be expected to experience a 1.79 percentage point increase in their probability of majoring in STEM, an effect size equivalent to roughly 43% of the average STEM attainment rate in the sample (4.18%).

[5] Accounting for heterogeneous exposure impacts explains approximately 46% of the total variation in STEM major attainment across counties (calculated as $8 \times 1.6\% + 5 \times 6.6\%$).

4.1 Validation of Baseline Design

We assess the key identifying assumption—that potential outcomes for students who move to higher versus lower STEM areas are not systematically influenced by the grade at which they move—using multiple tests to detect selection bias and omitted variable bias. First, we conduct outcome-based overidentification (or placebo) tests using variation in STEM rates among non-movers across cohorts and genders, and find group-specific convergence patterns that are unlikely to reflect selection, as it is implausible that parents could foresee such future differences in STEM outcomes when deciding to move. Moreover, we implement another overidentification test using alternative exposure measures based on non-STEM majors and find that STEM choice responds only to STEM-major exposure, suggesting that the estimated effects are driven by STEM-specific influences rather than general educational improvements. Second, we restrict comparisons to students of the same race and economic disadvantage status who move from the same origin to the same destination school catchment in the same year—effectively controlling for unobserved family and contextual factors—and find consistent results. Third, we analyze moves likely driven by displacement shocks and use predicted STEM exposure changes after the move (shift-share IV), which reduce omitted variable bias related to endogenous moving decisions, and again find similar estimates. Lastly, we conduct a stability test of coefficients with respect to unobservables and find that our estimates are unlikely to be driven by omitted variable bias (see Appendix H for more details).

The results consistently show that the evidence for exposure effects remains robust across various refinements of the baseline design, suggesting that the observed effects are not driven by selection. Based on the validity tests we conducted, we conclude that our estimate of $\gamma = 0.03$ represents an unbiased estimate of the annual exposure effect.

5 Occupation Composition in the Neighborhood

We explore one important neighborhood dimension underlying the observed STEM exposure effects. Our baseline measure of STEM exposure captures the combined influence of all neighborhood factors that shape STEM choices. These factors broadly fall into school and non-school environments, which often interact. While the effects of school environments, such as curriculum quality and teacher characteristics, are well-documented (Beilock et al. 2010; Clotfelter, Ladd, and Vigdor 2005; Lavy and Sand 2015; Legewie and DiPrete 2014; Wang 2013), the role of non-school environments has received less attention (see the relevant discussion in Xie, Fang, and Shauman 2015). To address this gap, we focus on one key non-school factor: the occupational composition of the neighborhood, and examine how it influences students' college major choices.

We use the occupational composition of neighborhoods as a proxy for the non-school environment. This measure captures students' potential exposure to working adults in STEM careers, whether through direct interactions, role modeling, or general visibility in the local economy. Although we do not have access to students' addresses, we can approximate neighborhood occupational composition by using the school addresses to estimate the school attendance zones in each county. Using detailed occupational data from the 2000 Census at the Census tract level (Manson 2020), we first construct a measure of STEM occupation exposure for each tract.[6] Based on the list of STEM occupations from O*NET (O*NET 2025),[7] we then calculate the STEM proportion in each Census tract ($STEM_c$) as the ratio of STEM workers to the total population (see Appendix Figure A.5 for Census tract level distribution). The STEM occupation exposure measure for each racial and economic status group in each county ($STEMOcc_{pg}$) is then defined as the average STEM proportion of the Census tracts where the schools (s) are located, along with nearby tracts, weighted by the number of students in each group enrolled in each school ($Enroll_{sg}$):

$$STEMOcc_{pg} = \sum_{s \in p} \left(\frac{Enroll_{sg}}{\sum_{s \in p} Enroll_{sg}} \times STEM_{c(s)} \right).$$

We begin by plotting the overall enrollment-weighted STEM occupational exposure measure

[6] Census tracts are small, relatively permanent statistical subdivisions of a county, with an average population of about 4,000 inhabitants.

[7] Specifically, our list of STEM occupations includes "Computer specialists," "Mathematical science occupations," "Architects, surveyors, and cartographers," "Engineers," "Drafters, engineering, and mapping technicians," "Life and physical scientists," "Life, physical, and social science technicians," "Health diagnosing and treating practitioners and technical occupations," and "Health technologists and technicians" among employed civilian population 16 years and over by sex by detailed occupation type. Additional documentation on NHGIS data sources is available at: https://www.nhgis.org/documentation/tabular-data, under 2000 Census: SF 4.

($STEMOcc_p$) across counties in Figure 5 (a). Notably, the distribution of STEM occupation rates closely aligns with our baseline STEM exposure measure. The two measures are highly correlated, with a correlation of 0.79. Even when the STEM rate is calculated only among undergraduate degree holders (i.e., STEM majors as a share of all undergraduate degrees), the correlation remains substantial at 0.47. As shown in Figure 5 (b), the relationship is approximately linear: a one-percentage-point increase in the STEM occupation rate is associated with a 0.3 percentage point increase in the baseline STEM exposure measure. In contrast, we find no comparable relationship for other majors. For instance, the correlation between liberal arts major rates and STEM occupation rates is 0.42 for all students but drops to -0.15 among STEM degree holders (Appendix Figure A.6). These patterns suggest that non-school environments, such as the local occupational landscape, play an important role in shaping students' decisions to pursue STEM fields.

Parallel to Figure 4, we estimate equations (2) and (3), replacing the baseline STEM exposure with the difference in STEM occupational exposure ($\Delta STEMOcc_{odg} = STEMOcc_{dg} - STEMOcc_{og}$). As shown in Figure 6, the estimates of b_m decline with the grade at which the move occurs. Moreover, Table A.2 shows that the estimated exposure effects are robust across various specifications. Specifically, we find that for each additional year spent in a neighborhood with a 1 percentage point higher STEM occupational share, the likelihood of choosing a STEM major increases by 0.017 percentage points. The estimation results are robust across different definitions of the STEM occupation exposure measure (Appendix Table A.3). It implies that a child moving in kindergarten from the 25th percentile (0.0463) of neighborhood with STEM occupation exposure ($STEMOcc_{pg}$) to one at the 75th percentile (0.0945)—a 1.5 standard deviation difference—would see an increase in the probability of majoring in STEM by 1.02 percentage points. This represents 24% of the average STEM major choice rate (4.18%).

We conduct analogous overidentification tests to those for the baseline STEM exposure measure using the gender of workers. We test for overidentification across genders, separately analyzing the STEM worker composition among male and female populations. As shown in Table A.4, the coefficient for the own-gender occupation exposure effect is 0.021, while the coefficient for the other-gender occupational composition indicates no effect. These results highlight that the exposure effects are gender-specific, emphasizing the importance of exposure to female STEM workers in increasing STEM participation among female students.

These findings highlight the non-academic channels through which place shapes educational outcomes. Beyond schools, the local labor market—through its occupational structure and representation—can influence how students perceive what careers are possible or attainable. Our overidentification tests suggest that the observed patterns are unlikely to be driven by selection (i.e., it is implausible that parents choose neighborhoods based on the gender composition of local STEM workers to match their child's gender). In fact, the areas with the highest STEM major rates tend to be located near major research and technology hubs, such as those surrounding NASA and Austin. For instance, 14 of the top 20 STEM Census tracts are located near the NASA Space Center, Austin, or the Houston Medical Center. This suggests that high neighborhood STEM occupation rates are more likely to drive higher student STEM attainment, rather than the pattern being explained by selection into particular educational or occupational environments. Given the empirical design, it is also unlikely that the result is driven by parents' own occupations.

6 Heterogeneity Across Groups

An important question that remains is whether high-STEM neighborhoods also encourage underrepresented groups to major in STEM. This question is particularly important to understand the gaps across demographic and socioeconomic status in STEM choices. To investigate this, we examine the heterogeneity in exposure effects across economic disadvantage status, genders, and races.

Focusing first on economic disadvantage status, we estimate equation (2) separately for economically disadvantaged and non-disadvantaged students. Figure 7 presents the coefficients b_m for both groups, showing a steady decline with the grade at which the move occurs, regardless of the economic disadvantage status. Table 3 presents linearly parameterized estimation results for exposure effects following equation (3). We find an average annual exposure effect of γ = 0.031 for non-disadvantaged students and 0.17 for disadvantaged students. This suggests that the benefits of living in high-STEM areas extend to students from both disadvantaged and non-disadvantaged backgrounds, while there is a difference in magnitude. Moreover, the overall likelihood of choosing STEM majors differs between the groups. The figure further shows that among students who moved in grade 11, the estimated effect for non-disadvantaged students is close to 0.2, whereas for disadvantaged students it is near zero. Given that a single year of exposure—especially in the final year before college—is unlikely to drive substantial changes on its own, this disparity may reflect differences in selection. Specifically, non-disadvantaged families may be more strategic in choosing high-STEM neighborhoods, with unobserved factors influencing STEM choice more closely aligned with local STEM rates. In contrast, we find little evidence of such selection among disadvantaged families.

We next examine heterogeneity by gender and race. As shown in Figure A.4 (a) and (b), exposure effects are present across all groups. The estimated annual exposure effects are $\gamma = 0.040$ for male students and 0.018 for female students; for racial groups, $\gamma = 0.025$ for White students, 0.051 for Black, and 0.024 for Hispanic students. These findings indicate that neighborhood STEM exposure positively affects students across gender and racial groups. Similar to the pattern by economic status, we observe a gap in late-grade exposure effects by race, with White students showing evidence of strategic moves. In contrast, we do not find a similar gap between male and female students, likely because there is no systematic difference in how families of boys and girls sort into high-STEM neighborhoods.

While all student groups benefit from STEM exposure, the level of exposure varies significantly depending on the neighborhoods where students of different racial and economic backgrounds reside. To assess the impact of these disparities, we conduct a simple counterfactual analysis that estimates how much the gap in STEM major choice would shrink if students from different backgrounds experienced the same neighborhood environments. Based on occupation exposure rates ($STEMOcc_{pg}$), we calculate the average STEM exposure: White students, on average, experience 29 and 66% higher neighborhood STEM exposure than Black and Hispanic students, respectively. If Black and Hispanic students had the same exposure levels as White students throughout their school years, their probability of majoring in STEM would increase by 0.7 and 0.9 percentage points, respectively, closing 22% of the Black–White gap and 41% of the

Hispanic–White gap.[8] Similarly, students who are not economically disadvantaged experience neighborhood STEM exposure rates that are 54% higher than those of disadvantaged students. Equalizing exposure would increase disadvantaged students' STEM major rates by 0.7 percentage points, narrowing 18% of the gap. We also find consistent results using the baseline STEM exposure measure (see Appendix I for more details).

Moreover, although male and female students live in the same neighborhoods, the share of STEM workers differs by gender, which is important given our findings on own-gender convergence. To assess the role of gender-specific exposure, we conduct a counterfactual analysis in which female students are exposed to the same neighborhood environments as male students, measured through same-gender STEM exposure. Under this scenario, the likelihood of women choosing a STEM major increases by 0.46 percentage points, closing 11% of the gender gap.

7 Exploring Underlying Mechanisms

Our analysis demonstrates that the neighborhood in which students grow up plays a meaningful role in shaping their decision to pursue STEM fields in college. In particular, we identify the local occupational structure—namely, the share of STEM workers in a given area—as a key contextual factor driving this effect. To better understand how such neighborhood characteristics influence student choices, we examine two potential mechanisms: academic performance and course-taking behavior. By analyzing these intermediate outcomes, we aim to determine whether the observed exposure effects operate through improvements in academic ability or through behavioral changes that reflect increased access to and engagement with STEM opportunities.

We begin by describing the characteristics of high-STEM neighborhoods. Figures 8 (a) and (b) present the relationship between our baseline STEM exposure measure and two educational outcomes at the county level: the number of Advanced Placement (AP) science and math courses taken and 11th grade math test scores.[9] To isolate the STEM-specific component of these

[8] Based on the heterogeneous exposure effects reported in Table 3, the expected reduction in the STEM attainment gap is approximately $(0.088-0.068)\times 13\times 2.7 = 0.70$ percentage points for Blacks and White students, $(0.088-0.053)\times 13\times 2.0 = 0.91$ for Hispanic and White students, and $(0.088-0.057)\times 13\times 1.7 = 0.69$ for economically disadvantaged and non-disadvantaged students.

[9] AP courses also include International Baccalaureate (IB) courses. We conduct a similar analysis using non-

outcomes, we residualize science AP course-taking by non-science AP course-taking, and math scores by reading scores. This residualization helps account for general academic achievement and isolates the variation more directly associated with STEM engagement. Consistent with our expectations, we find that students living in high-STEM-exposure areas are more likely to take science AP courses and perform better in math, relative to non-science/math AP courses and reading scores.

Turning to causal analysis, Figures 8 (c) and (d) examine how years of exposure to high-STEM areas affect intermediate outcomes using equation (2). Students who move to high-STEM areas earlier are likely to take more advanced science/math courses than those who move later, but no similar pattern emerges for math test scores. This suggests that course-taking behavior is more sensitive to exposure than test scores, reinforcing the idea that neighborhood effects operate primarily through behavioral channels rather than academic performance. This result is consistent with prior research on the role of academic preparedness in STEM choices (Justman and Méndez 2018; Maltese and Tai 2011; Sadler et al. 2014; Speer 2023), indicating that preparedness is not only a source of group differences in STEM but also a consequence of the environments students experience.

To formally assess whether these intermediate outcomes explain the overall exposure effect on STEM major choice, we conduct a mediation analysis in Table 4. Controlling for baseline (3rd-grade) or end-of-high-school (11th-grade) test scores has no effect on the estimated exposure effect. In contrast, controlling for the number of AP courses taken reduces the effect by about 30%. Additionally, math scores and science AP courses have more than ten times the predictive power for STEM choice compared with reading scores and non-science AP courses. This attenuation indicates that increased enrollment in advanced science and math coursework—rather than improvements in academic performance—is the primary pathway through which neighborhood exposure influences STEM major selection.

These findings highlight the importance of behavioral mechanisms in shaping educational

science AP courses and reading scores and find contrasting patterns, as shown in Appendix Figure A.7. The standardized test administered in 11th grade (TAKS) takes place between 2003 and 2013, likely covering cohorts of students who entered kindergarten between 1994 and 2001. As a result, the analysis using 11th-grade test scores does not include the full sample used in our main analysis. In Appendix Table A.5, we present results using 8th-grade test scores instead—which are available for all years—and find consistent patterns.

outcomes. They suggest that neighborhood environments influence students' choices by shaping the academic paths they follow, potentially signaling what is possible or expected. Policies aimed at increasing STEM participation may therefore benefit from expanding access to advanced coursework, especially in communities where such opportunities are limited. Place-based interventions that connect students with local STEM professionals, improve school offerings in STEM subjects, or promote a STEM-supportive culture could play a valuable role in closing participation gaps and fostering more equitable STEM engagement.

8 Conclusion

This paper provides causal evidence that neighborhoods influence students' college STEM choices by comparing students who move across counties and school districts at different grade levels. The estimated neighborhood effects suggest that nearly half of the geographic variation in STEM outcomes can be attributed to the environments students experience during their school years, rather than to differences in individual characteristics alone. We also find that the local occupational composition—particularly the presence of STEM-related jobs—is a key mechanism underlying these effects. Importantly, the benefits of high-STEM exposure are not limited to groups traditionally overrepresented in STEM; the effects extend across gender, race, and economic status. Counterfactual simulations based on our estimates indicate that 15–40% of the STEM participation gaps across these groups could be closed if they had equal access to high-STEM environments.

Our findings on STEM choice generalize to other college majors and levels of educational attainment. We observe similar patterns of major-specific convergence for a broad set of majors, which together account for over 70% of four-year college graduates. The estimated exposure effects decline with age at move across all major categories, with effect sizes varying by field (Appendix Figure A.8). We also find comparable exposure effects when defining STEM choice differently—such as enrollment in STEM programs or attainment of a STEM graduate degree—suggesting that early neighborhood environments shape not only whether students complete STEM degree, but also their broader post-secondary educational trajectories (Appendix Figure A.9). These results reinforce the view that students' academic and potentially career decisions are shaped not just by preferences or ability, but also by the educational and occupational context of the neighborhoods in which they grow up.

An important implication of this paper is the spillover effects of research hubs and agglomeration economies. Previous studies have analyzed the direct effects of agglomerations and local business booms (Allcott and Keniston 2018; Black, McKinnish, and Sanders 2005; Cust and Poelhekke 2015; Feyrer, Mansur, and Sacerdote 2017; Greenstone, Hornbeck, and Moretti 2010). While these hubs and economies generate numerous job opportunities for adults within neighborhoods, an additional potential benefit lies in their spillover impact on the human capital development of the next generation. These spillover effects have the potential to transcend income and socioeconomic disparities, ultimately benefiting children from disadvantaged households. As this research establishes causal links between neighborhood characteristics—shaped in part by research hubs and agglomeration—and students' STEM outcomes, the strategic placement of such hubs should be viewed not only as an economic development tool but also as a lever for expanding educational opportunity.

References

Allcott, Hunt, and Daniel Keniston. 2018. "Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America." *The Review of Economic Studies* 85, no. 2 (April): 695–731.

Altonji, Joseph G, Peter Arcidiacono, and Arnaud Maurel. 2016. "The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects." In *Handbook of the Economics of Education,* 5:305–396. Elsevier.

Arcidiacono, Peter, Esteban M Aucejo, and V Joseph Hotz. 2016. "University Differences in the Graduation of Minorities in Stem Fields: Evidence from California." *American Economic Review* 106 (3): 525–562.

Arcidiacono, Peter, V Joseph Hotz, Arnaud Maurel, and Teresa Romano. 2020. "Ex Ante Returns and Occupational Choice." *Journal of Political Economy* 128 (12): 4475–4522.

Arold, Benjamin W. 2024. "Evolution Vs. Creationism in the Classroom: The Lasting Effects of Science Education*." *The Quarterly Journal of Economics* (June).

Baran, Cavit, Eric Chyn, and Bryan A. Stuart. 2024. "The Great Migration and Educational Opportunity." *American Economic Journal: Applied Economics* 16, no. 3 (July): 354–398.

Beffy, Magali, Denis Fougère, and Arnaud Maurel. 2012. "Choosing the Field of Study in Postsecondary Education: Do Expected Earnings Matter?" *The Review of Economics and Statistics* 94, no. 1 (February): 334–347.

Beilock, Sian L, Elizabeth A Gunderson, Gerardo Ramirez, and Susan C Levine. 2010. "Female Teachers' Math Anxiety Affects Girls' Math Achievement." *Proceedings of the National Academy of Sciences* 107 (5): 1860–1863.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen. 2019. "Who Becomes an Inventor in America? the Importance of Exposure to Innovation." *The Quarterly Journal of Economics* 134 (2): 647–713.

Black, Dan, Terra McKinnish, and Seth Sanders. 2005. "The Economic Impact of the Coal Boom and Bust." *The Economic Journal* 115 (503): 449–476.

Brooks-Gunn, Jeanne, Greg J Duncan, Pamela Kato Klebanov, and Naomi Sealand. 1993. "Do Neighborhoods Influence Child and Adolescent Development?" *American journal of sociology* 99 (2): 353–395.

Carnevale, Anthony P, Nicole Smith, and Michelle Melton. 2011. "Stem: Science Technology Engineering Mathematics." *Georgetown University Center on Education and the Workforce.*

Carrell, Scott E, Marianne E Page, and James E West. 2010. "Sex and Science: How Professor Gender Perpetuates the Gender Gap." *The Quarterly journal of economics* 125 (3): 1101–1144.

Chetty, Raj, and Nathaniel Hendren. 2018. "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *The Quarterly Journal of Economics* 133 (3): 1107–1162.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review* 106 (4): 855–902.

Chyn, Eric. 2018. "Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children." *American Economic Review* 108 (10): 3028–3056.

Chyn, Eric, Robert Collinson, and Danielle H Sandler. 2025. "The Long-Run Effects of America's Largest Residential Racial Desegregation Program: Gautreaux." *The Quarterly Journal of Economics* (February): qjaf011.

Clotfelter, Charles T, Helen F Ladd, and Jacob Vigdor. 2005. "Who Teaches Whom? Race and the Distribution of Novice Teachers." *Economics of Education review* 24 (4): 377–392.

Cust, James, and Steven Poelhekke. 2015. "The Local Economic Impacts of Natural Resource Extraction." *Annual Review of Resource Economics* 7, no. Volume 7, 2015 (October): 251–268.

Deutscher, Nathan. 2020. "Place, Peers, and the Teenage Years: Long-Run Neighborhood Effects in Australia." *American Economic Journal: Applied Economics* 12, no. 2 (April): 220–249.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64 (5): 1001–1044.

Ellison, Glenn, and Ashley Swanson. 2010. "The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions." *Journal of Economic Perspectives* 24 (2): 109–128.

Elsner, Benjamin, Ingo E Isphording, and Ulf Zölitz. 2021. "Achievement Rank Affects Performance and Major Choices in College." *The Economic Journal* 131 (640): 3182–3206.

Feyrer, James, Erin T. Mansur, and Bruce Sacerdote. 2017. "Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution." *American Economic Review* 107, no. 4 (April): 1313–1334.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2021. "Place-based Drivers of Mortality: Evidence from Migration." *American Economic Review* 111, no. 8 (August): 2697–2735.

Foote, Andrew, and Kevin M Stange. 2022. "Attrition from Administrative Data: Problems and Solutions with an Application to Postsecondary Education. Working Paper 30232." *National Bureau of Economic Research.*

Fry, Richard, Brian Kennedy, and Cary Funk. 2021. "Stem Jobs See Uneven Progress in Increasing Gender, Racial and Ethnic Diversity." *Pew Research Center,* 1–28.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings." *Journal of political economy* 118 (3): 536–598.

Hanushek, Eric A, Babs Jacobs, Guido Schwerdt, Rolf Van der Velden, Stan Vermeulen, and Simon Wiederhold. 2023. *Where Do Stem Graduates Stem from? the Intergenerational Transmission of Comparative Skill Advantages.* Technical report. National Bureau of Economic Research.

Holden, J, and Eric Lander. 2012. "Report to the President: Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics." *Executive Office of the President President's Council of Advisors on Science and technology.*

Jencks, Christopher, Susan E Mayer, et al. 1990. "The Social Consequences of Growing Up in a Poor Neighborhood." *Inner-city poverty in the United States* 111:186.

Justman, Moshe, and Susan J Méndez. 2018. "Gendered Choices of Stem Subjects for Matriculation Are Not Driven by Prior Differences in Mathematical Achievement." *Economics of education review* 64:282–297.

Laliberté, Jean-William. 2021. "Long-Term Contextual Effects in Education: Schools and Neighborhoods." *American Economic Journal: Economic Policy* 13 (2): 336–377.

Landaud, Fanny, Son Thierry Ly, and Éric Maurin. 2020. "Competitive Schools and the Gender Gap in the Choice of Field of Study." *Journal of Human Resources* 55, no. 1 (January): 278–308.

Langdon, David, George McKittrick, David Beede, Beethika Khan, and Mark Doms. 2011. "Stem: Good Jobs Now and for the Future. Esa Issue Brief# 03-11." *US Department of Commerce.*

Lavy, Victor, and Edith Sand. 2015. *On the Origins of Gender Human Capital Gaps: Short and Long Term Consequences of Teachers' Stereotypical Biases.* Technical report. National Bureau of Economic Research.

Legewie, Joscha, and Thomas A DiPrete. 2014. "The High School Environment and the Gender Gap in Science and Engineering." *Sociology of education* 87 (4): 259–280.

Maltese, Adam V, and Robert H Tai. 2011. "Pipeline Persistence: Examining the Association of Educational Experiences with Earned Degrees in Stem Among Us Students." *Science education* 95 (5): 877–907.

Manson, Steven M. 2020. "Ipums National Historical Geographic Information System: Version 15.0."

Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson. 2022. "The Gift of Moving: Intergenerational Consequences of a Mobility Shock." *The Review of Economic Studies* 89 (3): 1557–1592.

National Science and Technology Council. 2000. *Ensuring a Strong Us Scientific, Technical, and Engineering Workforce in the 21st Century.* Executive Office of the President.

National Science Board. 2003. *The Science and Engineering Workforce: Realizing America's Potential.* National Science Foundation.

———. 2024. *The State of U.s. Science and Engineering 2024.* National Science Foundation.

Ngo, Diana, and Andrew Dustan. 2024. "Preferences, Access, and the Stem Gender Gap in Centralized High School Assignment." *American Economic Journal: Applied Economics* 16 (4): 257–287.

O*NET. 2025. "All Stem Occupations." https://www.onetonline.org/find/stem?t=0.

Patnaik, Arpita, Matthew Wiswall, and Basit Zafar. 2021. "College Majors 1." In *The Routledge Handbook of the Economics of Education.* Routledge.

Philippis, Marta De. 2023. "STEM Graduates and Secondary School Curriculum: Does Early Exposure to Science Matter?" *Journal of Human Resources* 58, no. 6 (November): 1914–1947.

Rozek, Christopher S, Gerardo Ramirez, Rachel D Fine, and Sian L Beilock. 2019. "Reducing Socioeconomic Disparities in the Stem Pipeline Through Student Emotion Regulation." *Proceedings of the National Academy of Sciences* 116 (5): 1553–1558.

Sadler, Philip M, Gerhard Sonnert, Zahra Hazari, and Robert Tai. 2014. "The Role of Advanced High School Coursework in Increasing Stem Career Interest." *Science Educator* 23 (1): 1–13.

Sampson, Robert J, Jeffrey D Morenoff, and Thomas Gannon-Rowley. 2002. "Assessing "neighborhood Effects": Social Processes and New Directions in Research." *Annual review of sociology* 28 (1): 443–478.

Schwank, Hanna. 2024. "Childhood Migration and Educational Attainment: Evidence from Indonesia." *Journal of Development Economics* 171 (October): 103338.

Sharkey, Patrick, and Jacob W Faber. 2014. "Where, When, Why, and for Whom Do Residential Contexts Matter? Moving Away from the Dichotomous Understanding of Neighborhood Effects." *Annual review of sociology* 40:559–579.

Speer, Jamin D. 2023. "Bye Bye Ms. American Sci: Women and the Leaky Stem Pipeline." *Economics of Education Review* 93:102371.

Stinebrickner, Ralph, and Todd R. Stinebrickner. 2014. "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout." *The Review of Economic Studies* 81, no. 1 (January): 426–472.

Wang, Xueli. 2013. "Why Students Choose Stem Majors: Motivation, High School Learning, and Postsecondary Context of Support." *American Educational Research Journal* 50 (5): 1081–1121.

Wiswall, Matthew, and Basit Zafar. 2015. "Determinants of College Major Choice: Identification Using an Information Experiment." *The Review of Economic Studies* 82 (2): 791–824.

———. 2018. "Preference for the Workplace, Investment in Human Capital, and Gender*." *The Quarterly Journal of Economics* 133, no. 1 (February): 457–507.

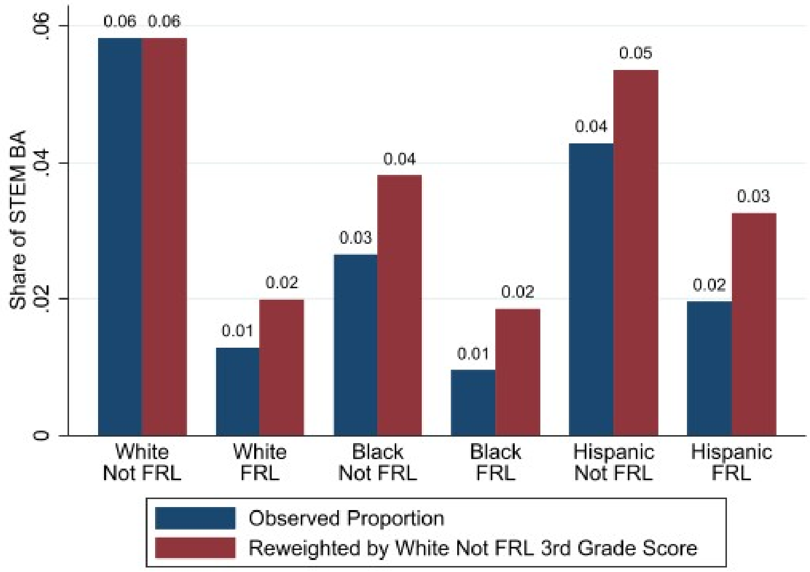
Xie, Yu, Michael Fang, and Kimberlee Shauman. 2015. "Stem Education." *Annual review of sociology* 41:331–357.

Xue, Yi, and Richard C Larson. 2015. "Stem Crisis or Stem Surplus? Yes and Yes." *Monthly labor review* 2015.

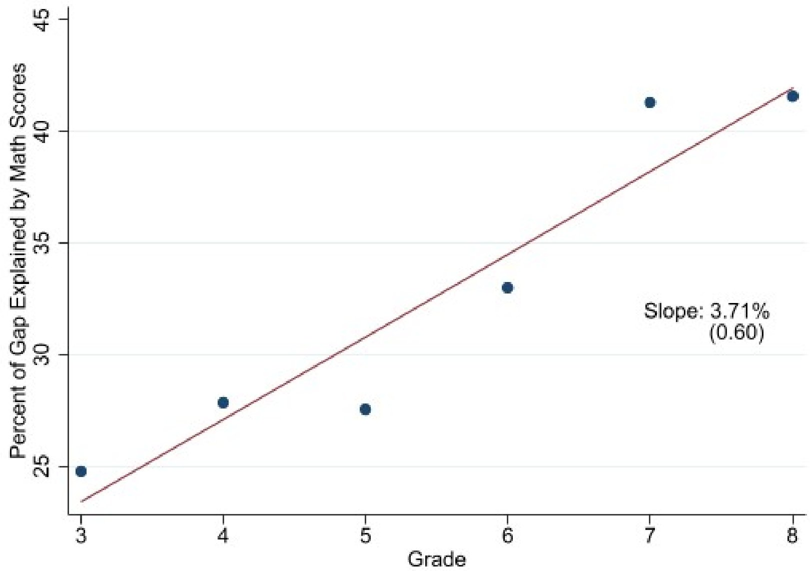
9 Figures and Tables

Fig. 1. STEM Attainment and Gaps Across Groups

(a) STEM Rates by Race and Economic Status

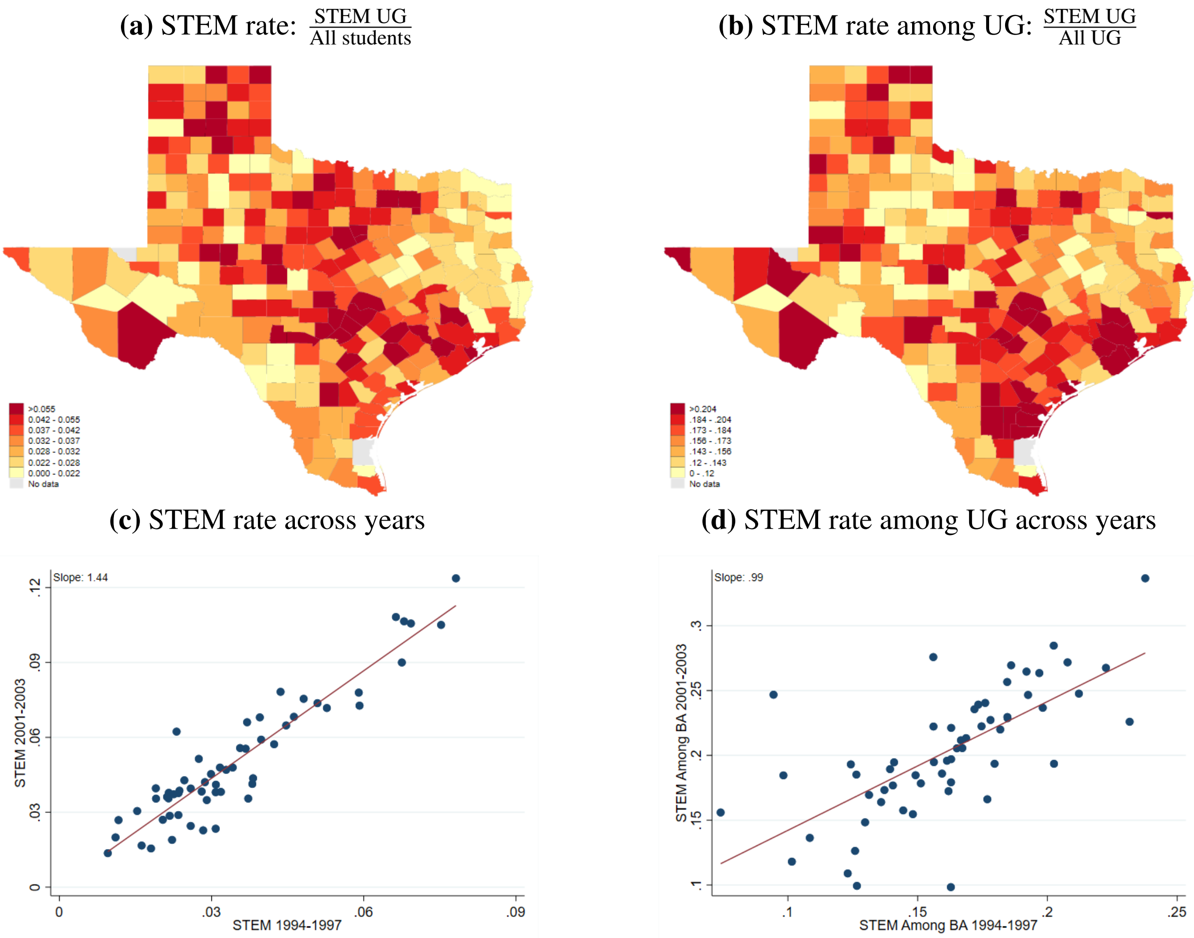


(b) Explained Gap in STEM Attainment by Test Scores Grades 3–8



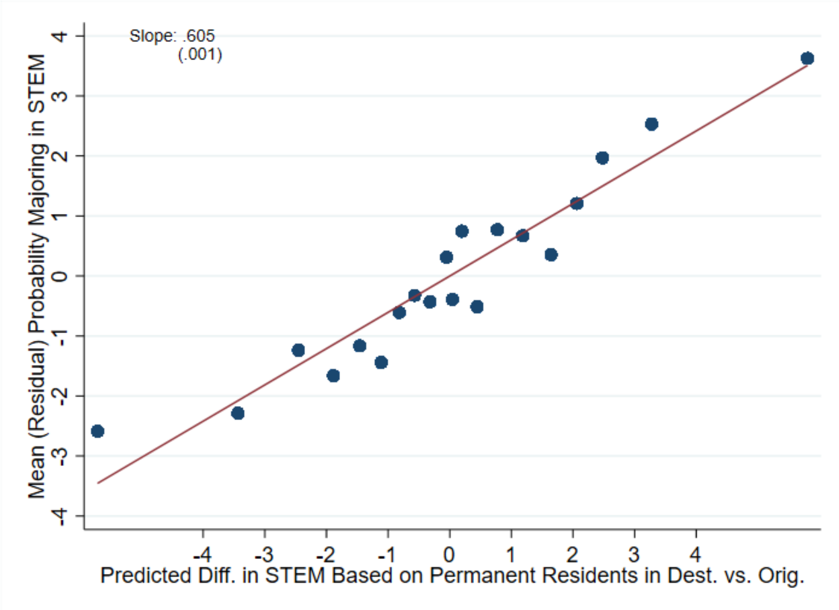
Notes: Figure (a) shows the share of students who obtain a STEM degree following the 2016 list of STEM majors from the Department of Homeland Security, broken down by race and economic disadvantage status, measured by eligibility for free or reduced-price lunch (FRL). The navy bars represent the share of students in each group who earn a STEM undergraduate degree by age 26. The red bars show the reweighted rates for each group, assuming they had the same third-grade math score distribution as White students who are not economically disadvantaged (Not FRL), following the methodology of DiNardo, Fortin, and Lemieux (1996). Specifically, we calculate the mean STEM attainment for each race and economic status group conditional on third-grade math scores and reweight these using the math score distribution of non-disadvantaged White students. Figure (b) illustrates how much of the STEM degree attainment gap between economically disadvantaged and non-disadvantaged students can be explained by differences in math scores. For each grade, students are divided into 20 bins based on their math scores. Within each bin, we compute the average STEM attainment rate for both groups, then reweight the distribution of disadvantaged students to match that of their non-disadvantaged peers. The sample for both panels includes all students who began kindergarten between 1994 and 2003.

Fig. 2. STEM Rate and Its Persistence Across Counties



Notes: These maps plot the proportion of non-moving students obtaining STEM undergraduate degrees (UG) by age 26 among all students in Figure (a) and among undergraduate degree holders in Figure (b). The maps are constructed by grouping counties into seven quantiles and shading the areas so that darker colors correspond to higher outcomes for students. Areas with no data are shaded with a gray color. The sample includes all students in the kindergarten in 1994-2003 who did not move across counties in their school years. Figures (c) and (d) plot the proportion of non-moving students obtaining an undergraduate degree in STEM majors in Texas counties from 1994-1997 school cohorts starting kindergarten in that corresponding year, compared to 2001-2003 school cohorts. The proportion in Figure (c) is calculated among all students in each county, and (d) is among all undergraduate degree holders. STEM majors are defined following the Department of Homeland Security's 2016 STEM major list.

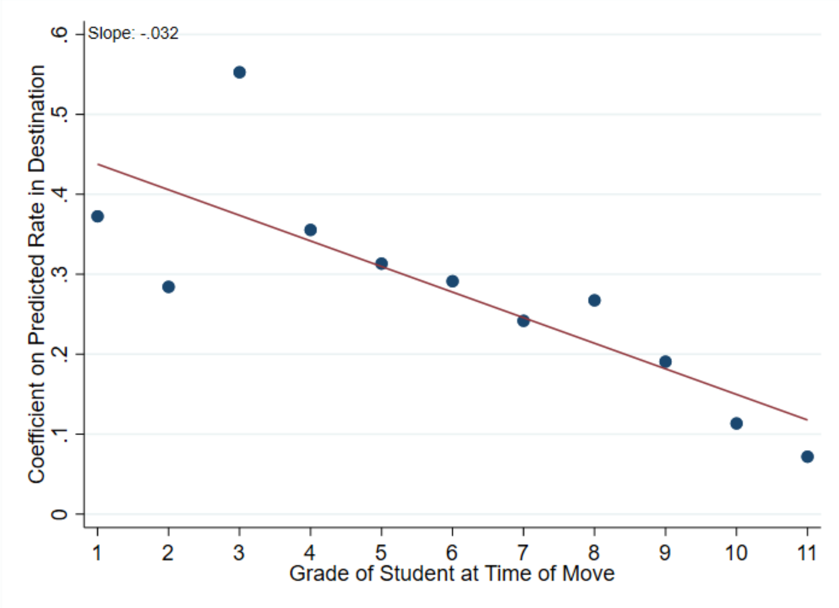
Fig. 3. Comparing Movers' Outcomes to Predicted Outcomes from Non-Movers' STEM Rates in Origin and Destination Counties



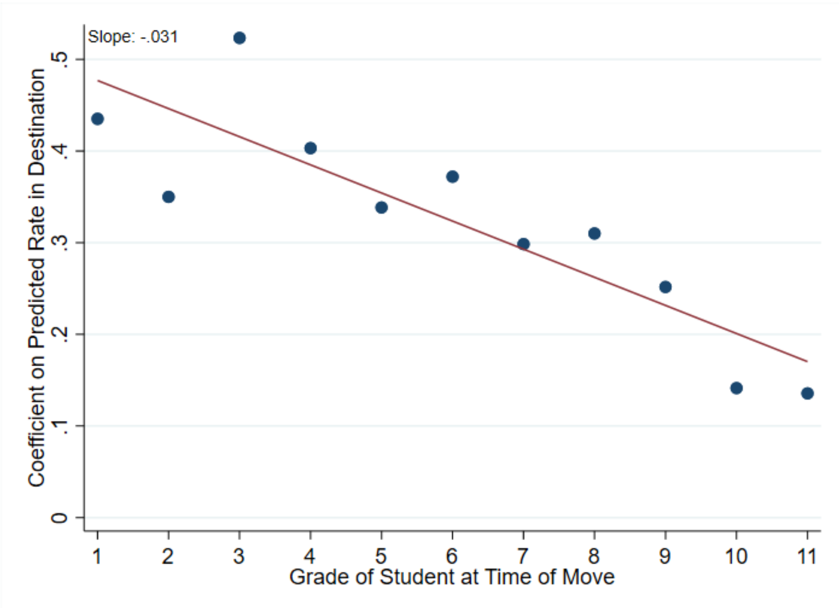
Notes: This figure presents a binned scatter plot depicting the relationship between the STEM probability of students who moved to a different county at grade 4 and the differences in the average probability of majoring in STEM of non-moving students in the destination vs. origin counties. The sample includes all students in kindergarten in 1994-2003 who moved when the student was in grade 4 and moved only once in school years. Students' STEM choices $STEM_i$ are measured as obtaining a STEM degree by age 26. Neighborhood STEM rate ($\overline{STEM}_{pg}$) is defined as the average STEM attainment among non-moving students in county p and racial-economic status group g. STEM rate changes ($\Delta STEM_{odg}$) for movers are calculated as the difference in STEM rates between the destination and origin counties: $\overline{STEM}dg - \overline{STEM}og$. To construct the figure, we first demean both the outcome and the exposure variables within cells defined by origin county (o), school-entry cohort (s), and race and economic status (g) among grade-4 movers. Specifically, we compute residuals: $STEM_i^r = STEM_i - E[STEM_i|o,s,g]$ and $\Delta STEM_{odg}^r = \Delta STEM_{odg} - E[\Delta STEM_{odg}|o,s,g]$. We then divide $\Delta STEM_{odg}^r$ into 20 equal-sized bins and plot the average value of $STEM_i^r$ against the average value of $\Delta STEM_{odg}^r$ within each bin.

Fig. 4. Exposure Effects on STEM Choice in College

(a) Semi-parametric Estimates



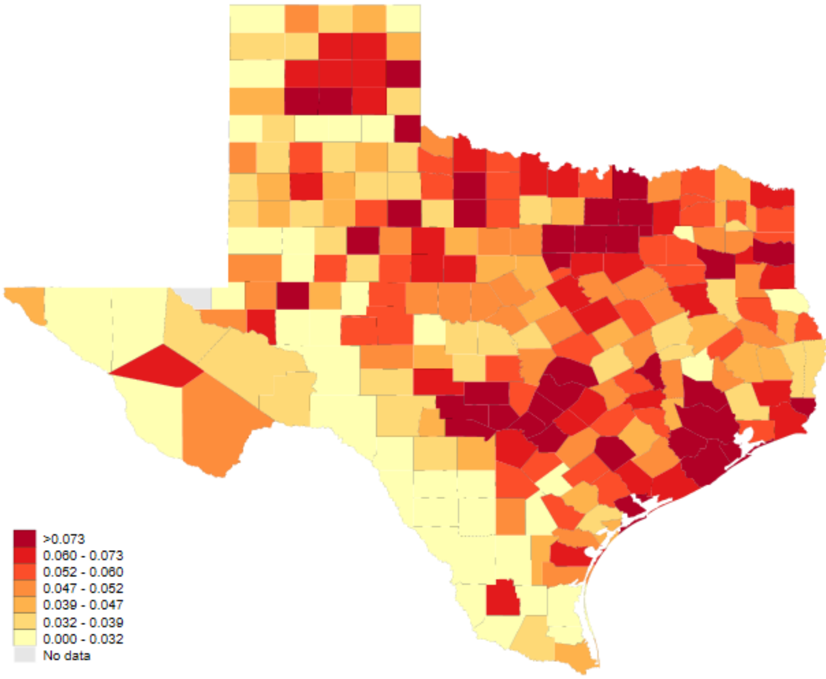
(b) Parametric Estimates



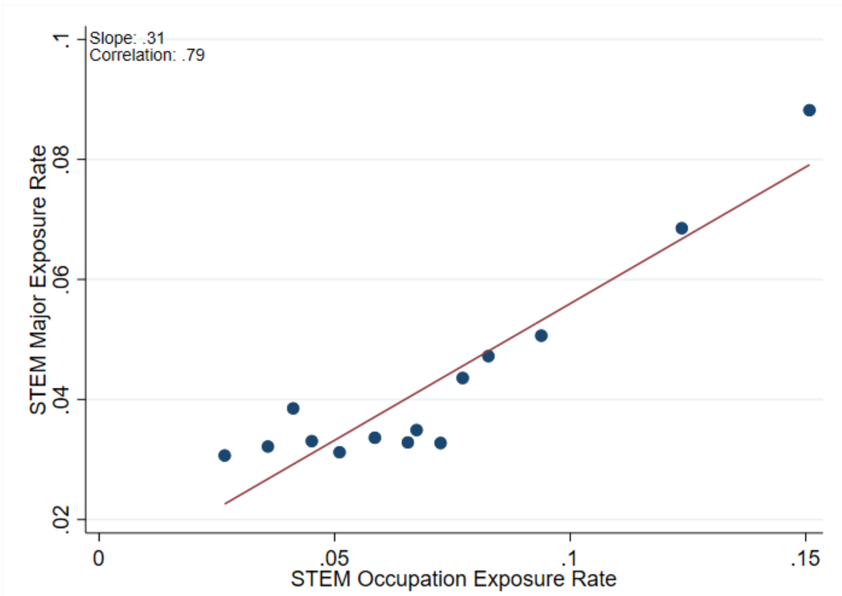
Notes: Figure (a) plots the estimated coefficients b_m by students' grade at the time of their move (m), using the semi-parametric specification in equation (1). The outcome variable, $STEM_i$, measures whether a student obtained an undergraduate degree in a STEM field by age 26. The sample includes all students who entered kindergarten between 1994 and 2003 and moved only once during their school years. The b_m coefficients represent the effect of moving at grade m to a county where the STEM rate among non-moving students in the same racial and economic status group increases from 0 to 1. These coefficients are estimated by regressing $STEM_i$ on $\Delta STEM_{odg} = \overline{STEMdg} - \overline{STEMog}$—the difference in predicted STEM attainment between destination and origin counties—interacted with the grade of the move. The model includes fixed effects for origin county-by-race-by-economic status-by-cohort-by-grade at move. Figure (b) presents estimates from the parametric specification in equation (2), which mirrors the approach in Figure (a) but replaces the origin fixed effects with the STEM rate for non-movers in the origin county. In both figures, best-fit lines are generated by unweighted OLS regressions of the b_m coefficients on m, and the estimated slopes, which are shown on the left side of each figure. The magnitudes of the slope represent estimates of annual exposure effects.

Fig. 5. STEM Occupation Exposure and College STEM Major Attainment Across Counties

(a) STEM Occupation Share by County



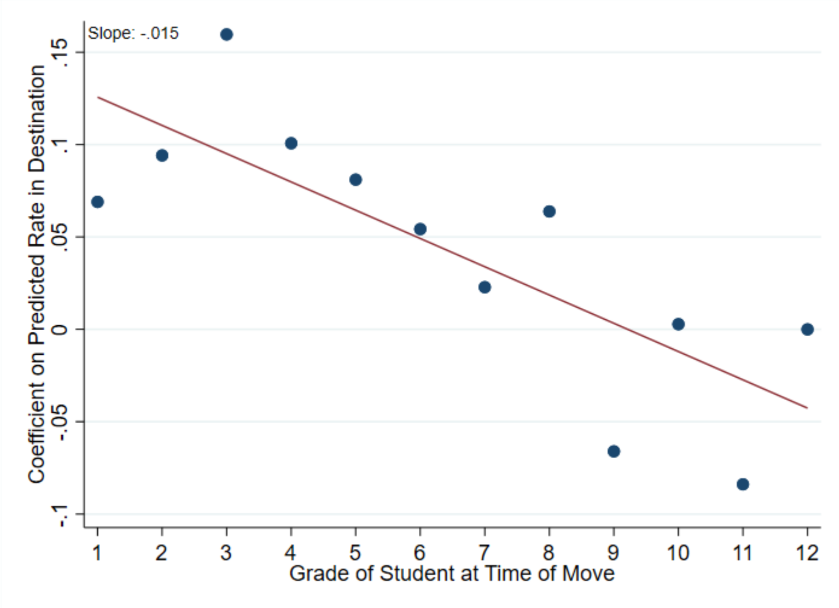
(b) Relationship Between STEM Occupation Exposure and College STEM Major Rates



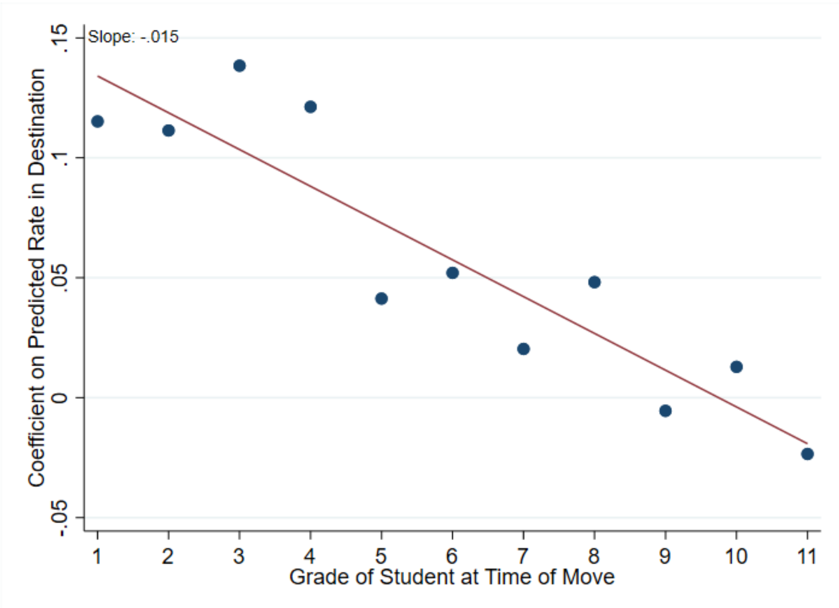
Notes: Figure (a) presents the proportion of STEM workers among the population in each county. To construct a county-level STEM occupation measure, we first construct a measure of STEM occupation exposure for each Census tract using occupational data from the 2000 Census at the Census tract level. Based on the list of STEM occupations from ONET, we calculate the STEM proportion in each Census tract as the ratio of STEM workers to the population 18-54. We then aggregate this to the county level by weighting each tract's STEM proportion by the number of students associated with school addresses in that tract. See Section 5 for more details. The map is constructed by grouping counties into seven quantiles and shading the areas so that darker colors correspond to a higher proportion of STEM workers. Areas with no data are shaded with a gray color. Figure (b) presents a binned scatter plot depicting the relationship between the STEM occupation exposure rate ($STEMOcc_p$) and baseline STEM major exposure rate ($\overline{STEM}_p$). The sample includes all students in kindergarten from 1994-2003 who never moved across counties during school years. We divide the $STEMOcc_p$ into twenty equal-sized groups among non-movers and plot the mean value of $\overline{STEM}_p$ vs. the mean value of $STEMOcc_p$ in each bin.

Fig. 6. Exposure Effects on STEM Choice in College: STEM Occupation Exposure

(a) Semi-parametric Estimates

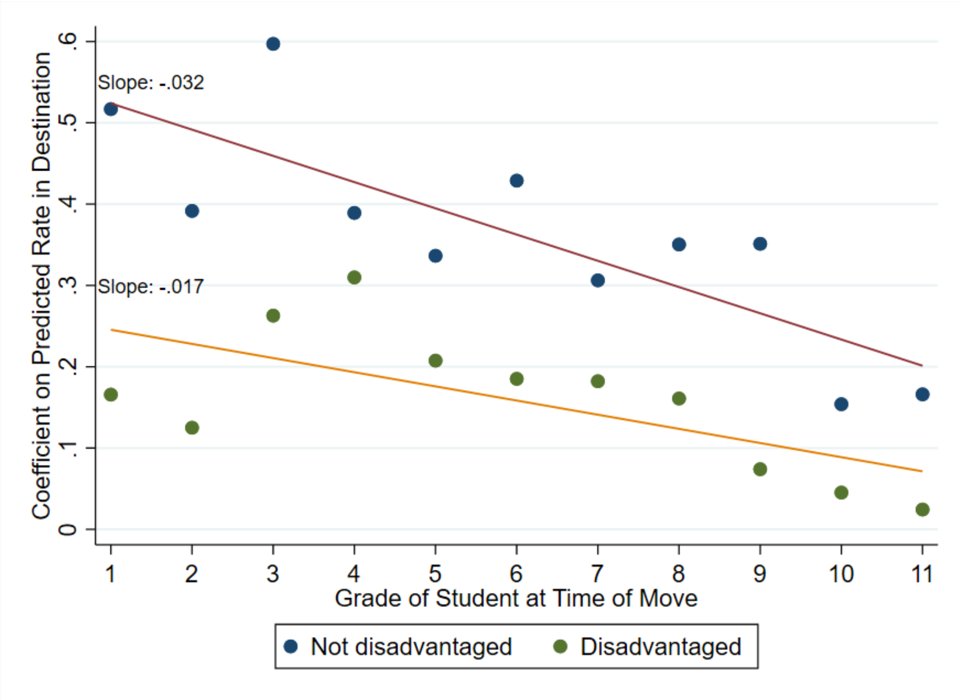


(b) Parametric Estimates



Notes: Figure (a) plots estimates of the coefficients b_m across the students' grade when they move (m) using the parametric specification in equation (1), replacing $\Delta STEM_{odg}$ to $\Delta STEMOcc_{odg} = STEMOcc_{dg}$ - $STEMOcc_{og}$, the difference in STEM worker proportions between destination and the origin. See Section 5 for more details on STEM occupation exposure measure construction. The sample includes all students in kindergarten in 1994-2003 who moved only once in their school years. Students' STEM choices $STEM_i$ are measured as obtaining a STEM undergraduate degree by age 26. The b_m coefficients can be interpreted as the effect of moving to an area where the neighborhood STEM occupation rate increases from 0 to 1 at grade m. They are estimated by regressing the students' STEM degree attainment in college $STEM_i$ on $\Delta STEMOcc_{odg}$, the difference between STEM worker proportion in the destination vs. the origin, interacted with each grade of the student at the time of the move m. We include origin county-by-race-by-economic disadvantage status-by-school cohort-by-grade at move fixed effects when estimating this specification. Figure (b) plots estimates from the parametric specification in equation (2). This specification replicates the specification used in Figure (a), replacing the origin fixed effects with the STEM occupation rate in the origin. Best-fit lines are estimated using unweighted OLS regressions of the b_m coefficients on m. The slopes of the regression line are reported on the left side of each figure.

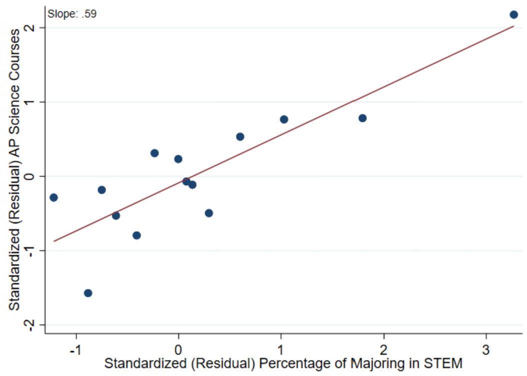
Fig. 7. Heterogeneous Exposure Effects by Economic Disadvantage Status



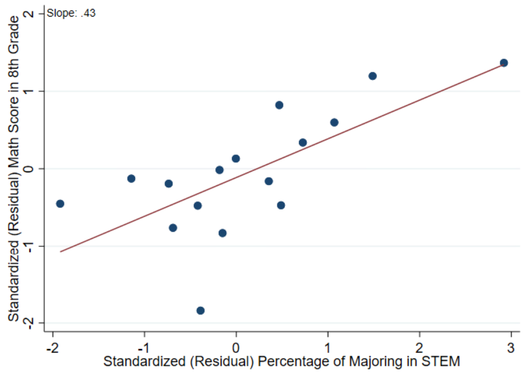
Notes: This figure plots the estimated coefficients b_m by students' grade at the time of their move (m), using the parametric specification in equation (1), estimated separately by students' economic disadvantage status. Students are classified as economically disadvantaged if they received free or reduced-price lunch in kindergarten. The blue dots and red line represent estimates for non-disadvantaged students, while the green dots and orange line represent estimates for disadvantaged students. The sample includes all students who entered kindergarten between 1994 and 2003 and moved only once during their school years. The outcome variable, $STEM_i$, measures whether a student earned a STEM undergraduate degree by age 26. The b_m coefficients reflect the effect of moving at grade m to a county where the STEM rate among non-moving students of the same racial and economic status group increases from 0 to 1. These are estimated by regressing $STEM_i$ on $\Delta STEM_{odg} = \overline{STEM}_{dg} - \overline{STEM}_{og}$—the difference in predicted STEM attainment between the destination and origin counties—interacted with the grade of the move. The model includes fixed effects for race-by-economic status-by-school cohort-by-grade at move, interacted with predicted STEM rate for non-moving students from the same racial and economic status group in the origin. Best-fit lines are estimated by regressing the b_m coefficients on m using unweighted OLS. The slopes of these regression lines, shown on the left of each line, represent annual exposure effects.

Fig. 8. County Characteristics and Exposure Effects on Intermediate Outcomes

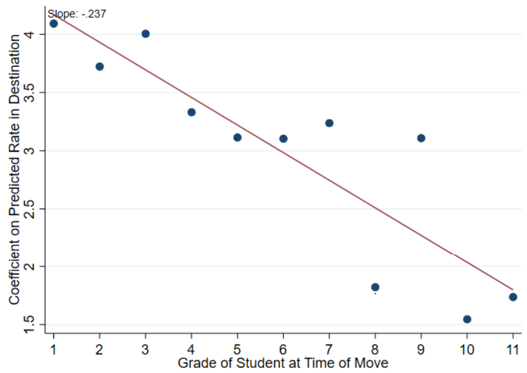
(a) County Characteristics: Science AP courses



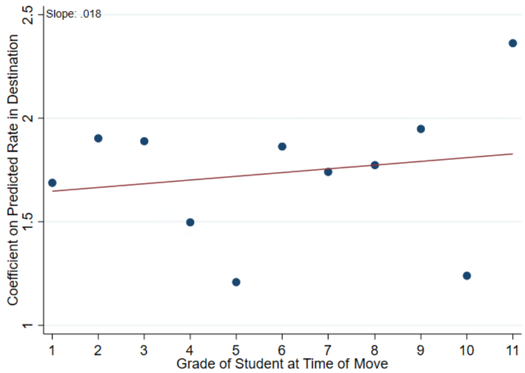
(b) County Characteristics: Math scores



(c) Exposure Effects: Science AP Courses



(d) Exposure Effects: Math Scores



Notes: The figures display county-level characteristics and the neighborhood exposure effect on intermediate outcomes: the number of science AP courses taken and math scores. Science AP courses include both math and science AP and IB courses. Math scores refer to standardized 11th-grade math test results. Figures (a) and (b) present a binned scatter plot depicting the relationship between the STEM exposure rate ($STEM_p$) and county-level educational characteristics. The sample includes all students in kindergarten in 1994-2003 who never moved across counties during their school years. To construct the figure (a), we first regress $STEM_p$ on the number of non-science AP courses and obtain residual $STEM_p^r$ and regress the number of science AP courses on the number of non-science AP courses and obtain residual AP_{sci}^r. Then, we standardize both AP_{sci}^r and $STEM_p^r$ by county. We then divide the standardized $STEM_p^r$ into twenty equal-sized groups and plot the mean value of demeaned AP_{sci}^r vs. the mean value of standardized $STEM_p^r$ in each bin. Figures (b) follow a similar procedure, residualizing both the 11th-grade math score and the STEM exposure rate by reading scores Figures (c) and (d) plot the estimated coefficients b_m by students' grade at the time of their move (m), using the parametric specification in equation (1), with intermediate outcomes. The sample includes all students who entered kindergarten between 1994 and 2003 and moved only once during their school years. The b_m coefficients reflect the effect of moving at grade m to a county where the STEM rate among non-moving students of the same racial and economic status group increases from 0 to 1. These are estimated by regressing the intermediate outcome variable on $\Delta STEM_{odg} = \overline{STEM}_{dg} - \overline{STEM}_{og}$—the difference in predicted STEM attainment between the destination and origin counties—interacted with the grade of the move. The model includes fixed effects for race-by-economic status-by-school cohort-by-grade at move, interacted with predicted STEM rate for non-moving students from the same racial and economic status group in the origin. Best-fit lines are estimated by regressing the b_m coefficients on m using unweighted OLS. The slopes of these regression lines, shown on the left of each line, represent annual exposure effects.

Table 1: Summary Statistics for Non-Movers and Movers

	Non-moving Students		Moving once students	
	(1) Mean	(2) Std. Dev.	(3) Mean	(4) Std. Dev.
Attaining Bachelor's degree	0.214	0.41	0.209	0.406
Attaining STEM degree	0.039	0.195	0.038	0.192
Female students	0.49	0.50	0.49	0.50
Black students	0.14	0.35	0.15	0.36
Hispanic students	0.48	0.50	0.34	0.47
White students	0.38	0.49	0.51	0.50
Economically disadvantaged students	0.56	0.50	0.45	0.50
Observations	1,575,342		205,788	

Notes: The table presents summary statistics for the samples used in the baseline analyses. The sample consists of all students in the data who attended kindergarten between 1994-2003 and are observed for at least 10 years within Texas public schools. We report summary statistics for two subsets of this sample. Non-moving students represent students who do not move across counties throughout their school years. Moving-once students represent students who move once across counties in their school years. Attaining an undergraduate degree is defined as having an undergraduate degree from an institution in Texas by age 26. Attaining a STEM degree is defined as having an undergraduate degree in STEM majors defined by the STEM degree list from the Department of Homeland Security in an institution in Texas by age 26. Economic disadvantage is defined based on eligibility for free or reduced-price lunch. See Section 2 for further details on variable and sample definitions.

Table 2: Exposure Effects in STEM Major Choice

	Parametric				Semi-parametric			
	(1) Baseline	(2) + Destination info	(3) + Test scores & Labor char.	(4) Catchment FE	(5) Baseline	(6) + Destination info	(7) + Test scores & Orig.×dest. FE	(8) Catchment FE
Exposure effects	0.030***	0.027***	0.030***	0.034**	0.031***	0.028***	0.029***	0.037*
	(0.005)	(0.005)	(0.005)	(0.013)	(0.006)	(0.007)	(0.008)	(0.021)
Observations	205777	205777	182835	87000	199502	199497	172057	77607

Notes: This table reports estimates of annual exposure effects on students' STEM degree attainment by age 26. The coefficients can be interpreted as the impact of spending an additional school year in a county where non-moving students have a one percentage point higher STEM attainment rate. Each column presents results from a regression of students' STEM degree attainment on the difference in STEM rates between destination and origin counties for non-moving students of the same race and economic status group, interacted with the grade at which the students move (m). Column (1) reports the estimate of γ and its standard error from equation (3), using the sample of one-time movers defined in the notes to Table 1. Column (2) adds to column (1) fixed effects for race-by-economic status interacted with the destination STEM rate of non-movers from the same race and economic status group. Column (3) further includes labor market characteristics (median income, employment rate, and poverty rate in origin and destination counties, constructed using the same method described in Section 5) and moving students' third-grade math and reading scores. Column (4) adds to column (1) finer fixed effects: race-by-economic status-by-origin school catchment area-by-destination catchment area-by-year of move. Columns (1) through (4) all include race-by-economic status-by-cohort-by-grade-of-move fixed effects, interacted with the origin STEM rate of non-movers from the same race and economic status group. Columns (5) through (8) replicate the specifications in Columns (1) through (4), respectively, but replace the origin STEM rate of non-movers—which was interacted with race-by-economic status-by-cohort-by-grade-of-move fixed effects—with origin county fixed effects. Additionally, Column (7) includes origin-by-destination county fixed effects by race and economic status instead of labor market controls. Some observations are dropped due to the inclusion of fixed effects. For results based on a consistent sample across specifications, see Appendix Table H.4. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

Table 3: Heterogeneous Effects Across Different Groups

Panel A: using the baseline STEM exposure measure

	(1) Not Econ.	(2) Econ.	(3) Black	(4) Hispanic	(5) White	(6) Male	(7) Female
Exposure effects	0.031***	0.017***	0.051***	0.024**	0.025***	0.040***	0.018**
	(0.007)	(0.006)	(0.007)	(0.011)	(0.007)	(0.006)	(0.008)
Obs.	112,374	93,403	30,607	69,606	105,564	105,549	100,228

Panel B: using STEM occupation exposure measure

	(1) Not Econ.	(2) Econ.	(3) Black	(4) Hispanic	(5) White	(6) Male	(7) Female
Exposure effects	0.015***	0.017***	0.027***	0.020***	0.014**	0.022***	0.013**
	(0.004)	(0.004)	(0.005)	(0.008)	(0.006)	(0.005)	(0.005)
Obs.	112,371	93,400	30,606	69,604	105,561	105,547	100,224

Notes: The table reports estimates of annual exposure effects γ using different groups of samples. In all columns, the dependent variable is the student's STEM degree attainment by age 26. The sample consists of a group of students of one-time movers, defined in the notes to Table 1. In Panels A and B, each column replicates column (1) of Table 2 and A.2, using the sample group indicated in the column header. The labels "Not Econ." and "Econ." denote economically non-disadvantaged and disadvantaged students, respectively. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

Table 4: Mediation Analysis

	(1) Baseline	(2) +3rd test score	(3) +11th test score	(4) +AP Courses
Exposure effects	0.030***	0.030***	0.031***	0.021***
	(0.007)	(0.007)	(0.007)	(0.006)
Math score 3rd grade		0.022***		
		(0.001)		
Reading score 3rd grade		0.002		
		(0.001)		
Math score 11th grade			0.040***	
			(0.001)	
Reading score 11th grade			-0.003**	
			(0.001)	
Non-science AP courses				0.002***
				(0.000)
Science AP courses				0.038***
				(0.001)
Observations	127,942	127,942	127,942	127,942

Notes: The table reports estimates of annual exposure effects γ with different intermediate outcome variables. In all columns, the dependent variable is the student's STEM degree attainment by age 26. The sample consists of a group of students of one-time movers, defined in the notes to Table 1. Each column replicates column (1) of Table 2. Students with both 3rd and 11th grade test scores are included. See Appendix Table A.5 for results based on a sample with varying observation counts. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

ONLINE APPENDIX

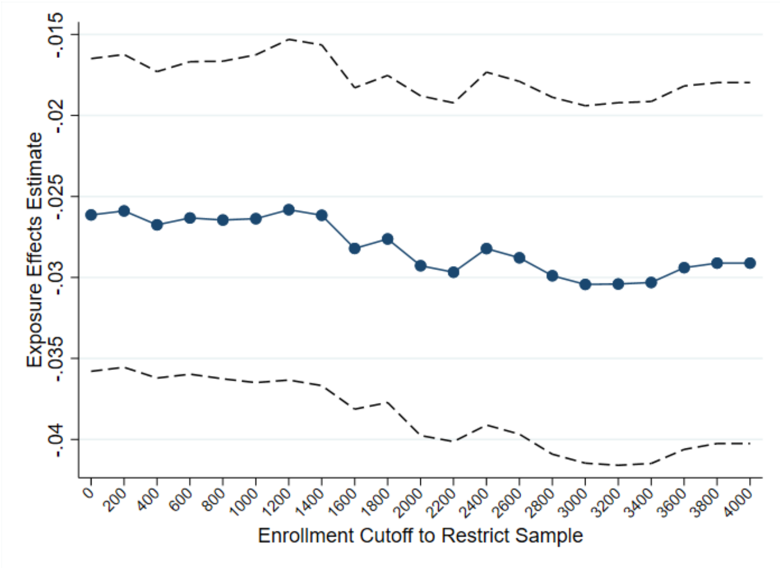
Neighborhood Effects on STEM Major Choice

Jeonghyeok Kim   Rohit Munshi   Brian Murphy

Table of Contents

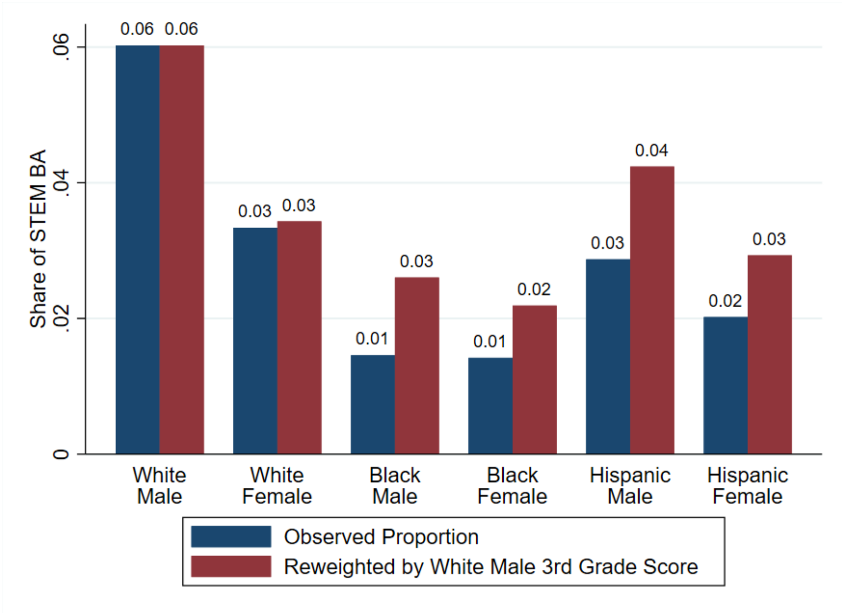
A Additional Results

Fig. A.1. Exposure Effects Estimates Using Different Enrollment Cutoffs



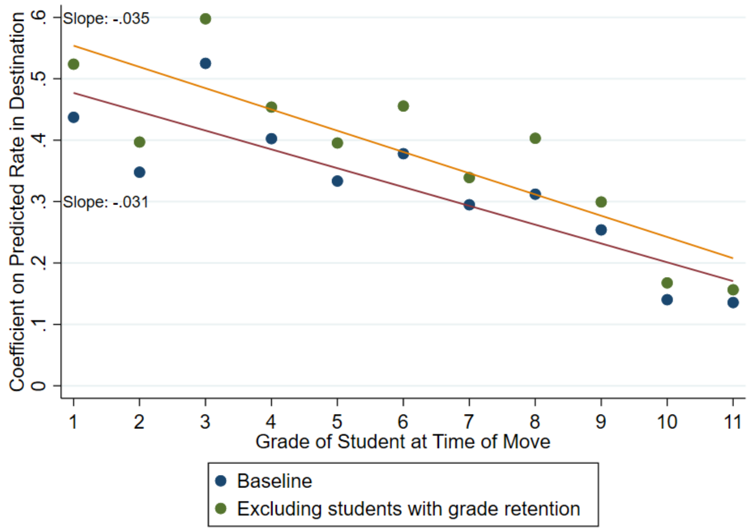
Notes: The figure presents estimates of annual exposure effects on students' STEM degree attainment by age 26 with 95% confidence intervals, using different enrollment cutoffs ranging from no cutoff to a minimum of 4,000 non-moving students with the same race and economic status in school cohorts attending kindergarten between 1994-2003 within a county. Each dot reports the estimate of γ from equation (3) using all students of one-time movers, defined in the notes to Table 1. The dashed lines show the corresponding 95% confidence intervals for each estimate.

Fig. A.2. STEM Rates by Race and Gender



Notes: The figure shows the share of students who obtain a STEM degree following the 2016 list of STEM majors from the Department of Homeland Security, broken down by race and gender. The navy bars represent the share of students in each group who earn a STEM undergraduate degree by age 26. The red bars show the reweighted rates for each group, assuming they had the same third-grade math score distribution as White male students, following the methodology of DiNardo, Fortin, and Lemieux (1996). Specifically, we calculate the mean STEM attainment for each race and gender group conditional on third-grade math scores and reweight these using the math score distribution of White male students.

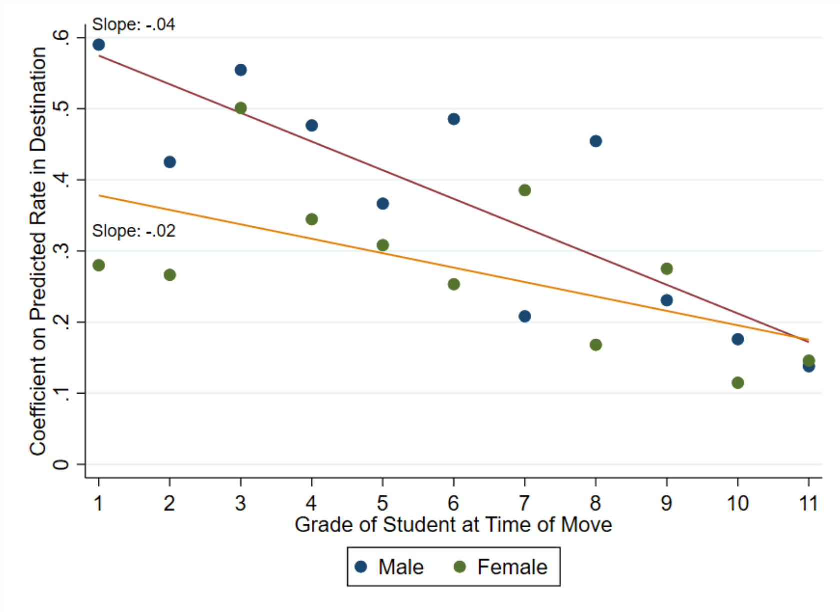
Fig. A.3. Exposure Effects on STEM Choice in College Excluding Sample with Grade Retention



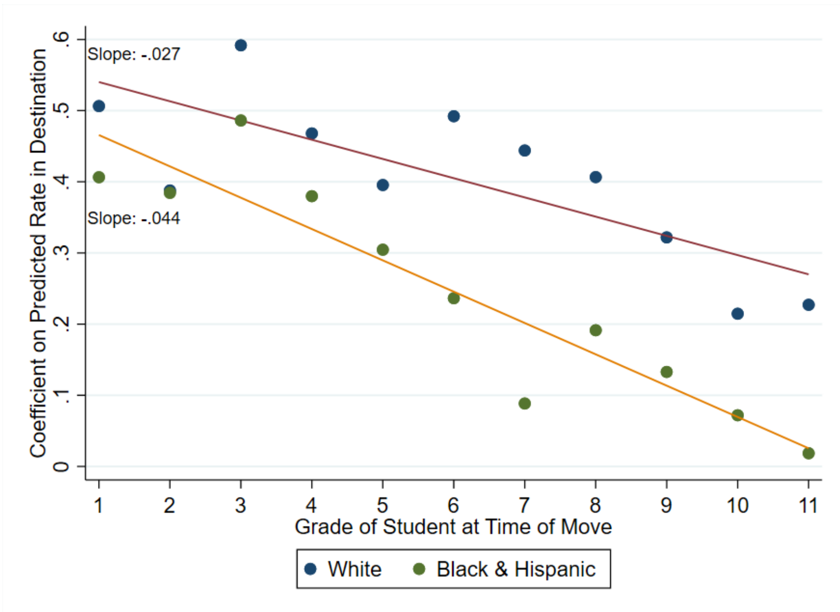
Notes: This figure plots the estimated coefficients b_m by students' grade at the time of their move (m), using the parametric specification in equation (1), with the sample of students who never repeated their grade throughout their school years. The blue dots and red line represent estimates for baseline estimates, while the green dots and orange line represent estimates with the sample of students who never repeated their grade throughout school years. The sample includes all students who entered kindergarten between 1994 and 2003 and moved only once during their school years. The outcome variable, $STEM_i$, measures whether a student earned a STEM undergraduate degree by age 26. The b_m coefficients reflect the effect of moving at grade m to a county where the STEM rate among non-moving students of the same racial and economic status group increases from 0 to 1. These are estimated by regressing $STEM_i$ on $\Delta STEM_{odg} = \overline{STEM}_{dg} - \overline{STEM}_{og}$—the difference in predicted STEM attainment between the destination and origin counties, interacted with the grade of the move. The model includes fixed effects for race-by-economic status-by-school cohort-by-grade at move, interacted with predicted STEM rate for non-moving students from the same racial and economic status group in the origin. Best-fit lines are estimated by regressing the b_m coefficients on m using unweighted OLS. The slopes of these regression lines, shown on the left of each line, represent annual exposure effects.

Fig. A.4. Heterogeneous Exposure Effects by Gender and Race

(a) Gender



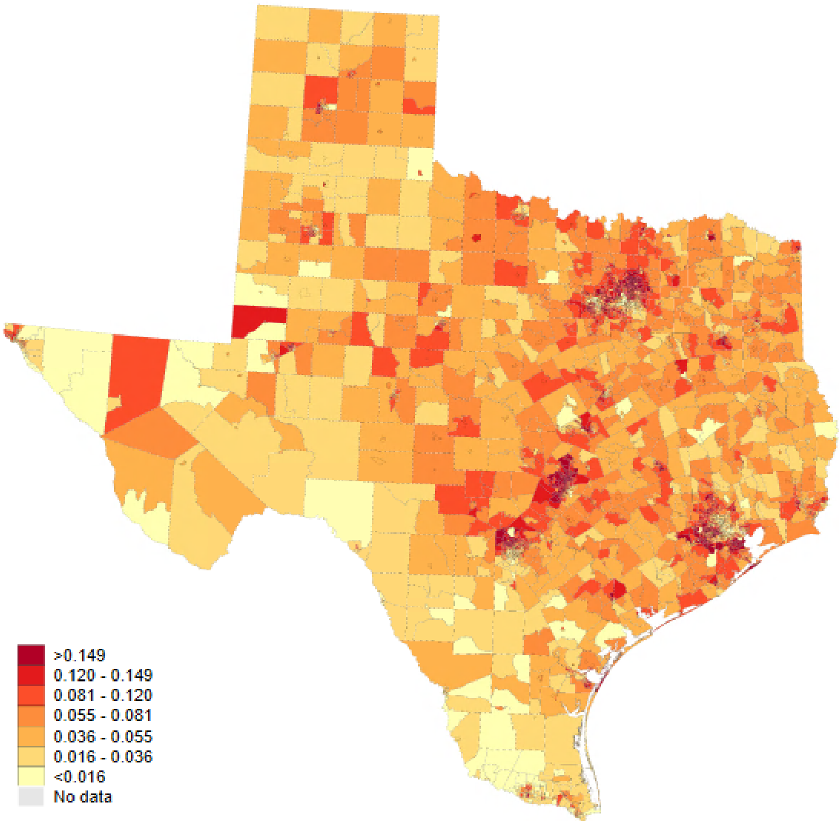
(b) Race/Ethnicity



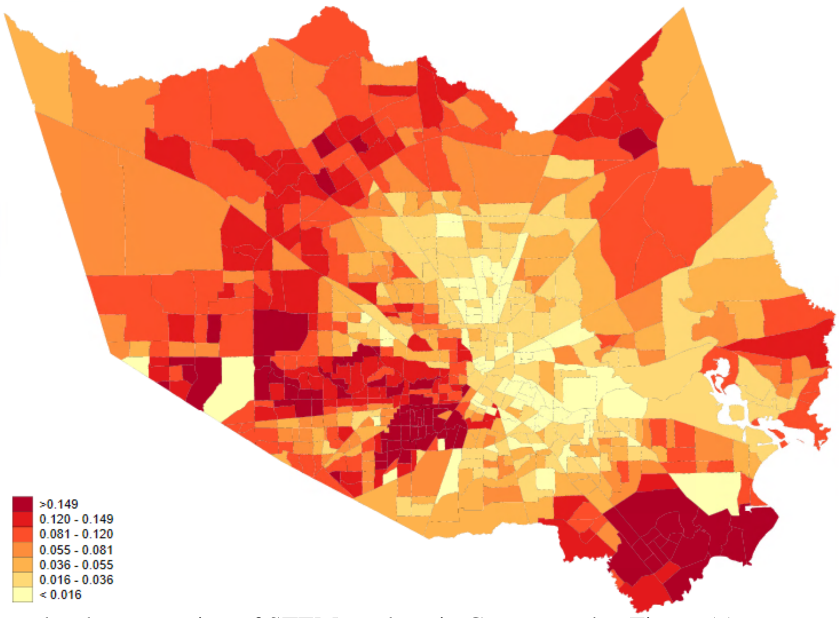
Notes: This figure plots the estimated coefficients b_m by students' grade at the time of their move (m), using the parametric specification in equation (1), estimated separately by gender or race/ethnicity of students. The blue dots and the red line are from the sample of male students for Figure (a) and White students for Figure (b), and the green dots and the orange line are from the sample of female students for (a) and Black/Hispanic students for (b). The sample includes all students who entered kindergarten between 1994 and 2003 and moved only once during their school years. The outcome variable, $STEM_i$, measures whether a student earned a STEM undergraduate degree by age 26. The b_m coefficients reflect the effect of moving at grade m to a county where the STEM rate among non-moving students of the same racial and economic status group increases from 0 to 1. These are estimated by regressing $STEM_i$ on $\Delta STEM_{odg} = \overline{STEM}_{dg} - \overline{STEM}_{og}$—the difference in predicted STEM attainment between the destination and origin counties—interacted with the grade of the move. The model includes fixed effects for race-by-economic status-by-school cohort-by-grade at move, interacted with predicted STEM rate for non-moving students from the same racial and economic status group in the origin. Best-fit lines are estimated by regressing the b_m coefficients on m using unweighted OLS. The slopes of these regression lines, shown on the left of each line, represent annual exposure effects.

Fig. A.5. STEM Occupation Proportion Across Census Tracks

(a) STEM Occupation Proportion



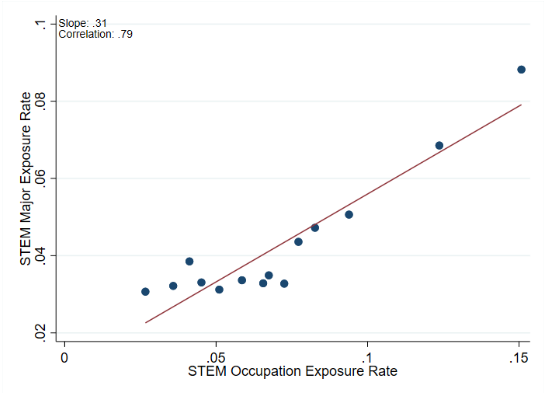
(b) STEM Occupation Proportion - Harris County



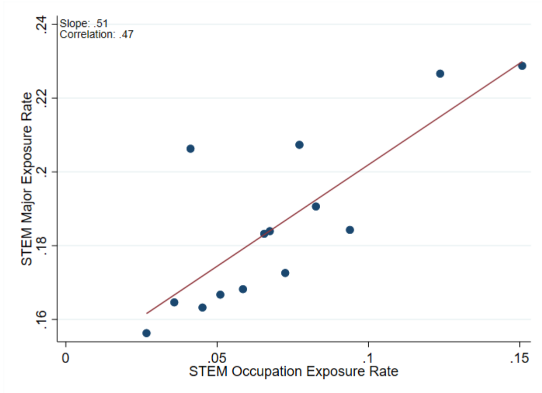
Notes: These maps plot the proportion of STEM workers in Census tracks. Figure (a) presents the overall Texas and Figure (b) focuses on Harris County. The maps are constructed by grouping Census tracks into seven quantiles and shading the areas so that darker colors correspond to a higher proportion of STEM workers. Areas with no data are shaded with a gray color. These maps are based on the 2000 Census, and the STEM occupation follows the list of STEM occupations from ONET. Specifically, our list of STEM occupations includes "Computer specialists," "Mathematical science occupations," "Architects, surveyors, and cartographers," "Engineers," "Drafters, engineering, and mapping technicians," "Life and physical scientists," "Life, physical, and social science technicians," "Health diagnosing and treating practitioners and technical occupations," and "Health technologists and technicians" among employed civilian population 16 years and over by sex by detailed occupation type. See more details in Section 5.

Fig. A.6. STEM or Liberal Arts Major Rates and STEM Occupation Rates Across Counties

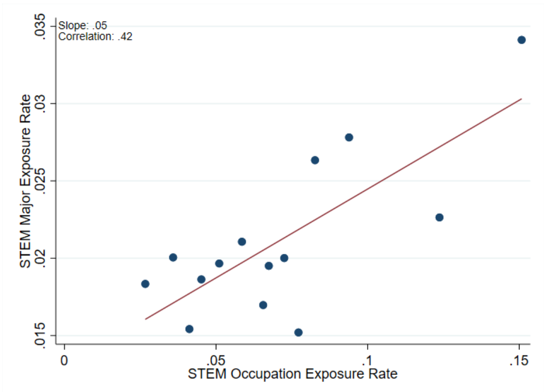
(a) STEM Major Rate and STEM Occupation Rate



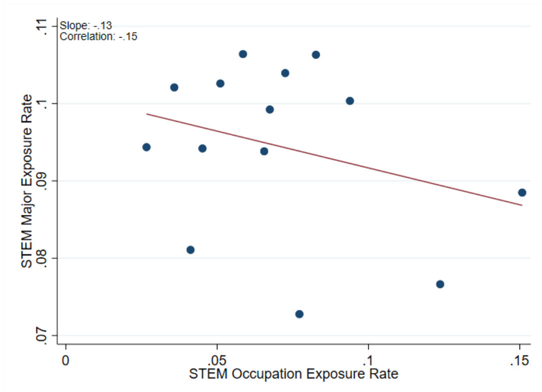
(b) STEM Major Rate among BAs and STEM Occupation Rate



(c) Liberal Arts Major Rate and STEM Occupation Rate



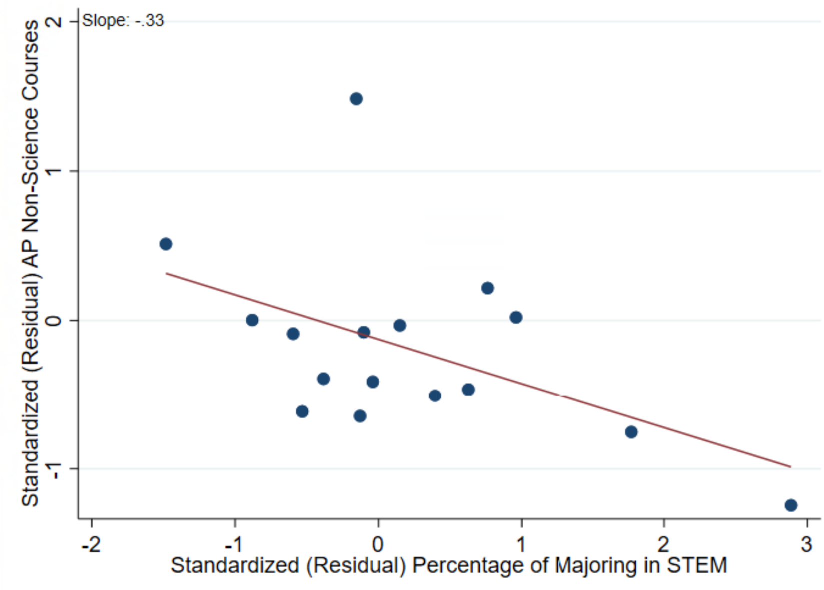
(d) Liberal Arts Major Rate among BAs and STEM Occupation Rate



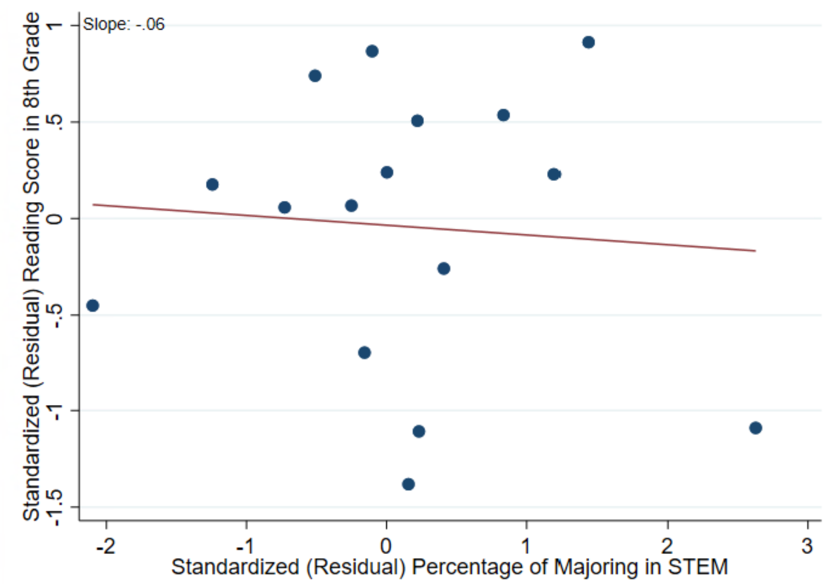
Notes: The figures present a binned scatter plot depicting the relationship between the STEM occupation exposure rate ($STEMOcc_p$) and STEM or liberal arts major exposure rate ($\overline{STEM}_p$ or $\overline{LibArts}_p$). The STEM and liberal arts major exposure rates are calculated based on the proportion of students majoring in STEM or liberal arts fields. STEM majors are defined according to the DHS STEM major list (Appendix J). Liberal arts majors are broadly defined to include education (CIP code 13); foreign languages, literatures, and linguistics (16); English language and literature/letters (23); liberal arts and sciences, general studies, and humanities (24); philosophy and religious studies (38); theology and religious vocations (39); social sciences (45); communication, journalism, and related programs (09); and history (54). To construct a county-level STEM occupation measure, we first construct a measure of STEM occupation exposure for each Census tract using occupational data from the 2000 Census at the Census tract level. Based on the list of STEM occupations from ONET, we calculate the STEM proportion in each Census tract as the ratio of STEM workers to the population 18-54. We then aggregate this to the county level by weighting each tract's STEM proportion by the number of students associated with school addresses in that tract. See Section 5 for more details. The sample includes all students in kindergarten from 1994-2003 who never moved across counties during their school years. We divide the $STEMOcc_p$ into twenty equal-size groups among non-movers and plot the mean value of $\overline{STEM}_p$ (or $\overline{LibArts}_p$) vs. the mean value of $STEMOcc_p$ in each bin.

Fig. A.7. STEM Rate and County Characteristics: Non-science AP Courses and Reading Scores

(a) Non-science AP courses



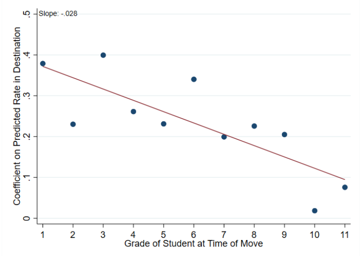
(b) Reading scores



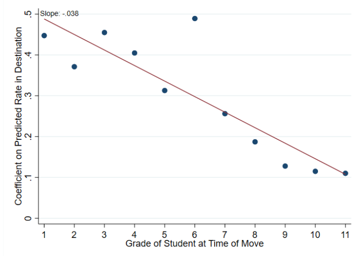
Notes: The figures present a binned scatter plot depicting the relationship between the STEM exposure rate ($STEM_p$) and county-level educational characteristics: the average number of AP non-science courses and average 11th-grade reading scores. The sample includes all students in kindergarten in 1994-2003 who never moved across counties during their school years. To construct the figure (a), we first regress $STEM_p$ on the number of science AP courses and obtain residual $STEM_p^r$ and regress the number of non-science AP courses on the number of science AP courses and obtain residual AP_{nonsci}^r. Then, we standardize both AP_{nonsci}^r and $STEM_p^r$ by county. We then divide the standardized $STEM_p^r$ into twenty equal-sized groups and plot the mean value of demeaned AP_{nonsci}^r vs. the mean value of standardized $STEM_p^r$ in each bin. Figures (b) follow a similar procedure, plotting the residualized correlation between the STEM exposure rate and reading score. Reading scores and the STEM rate are residualized by math scores.

Fig. A.8. Exposure Effects on Other Major Choice in College

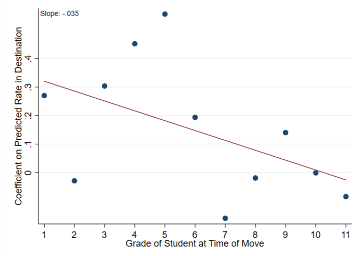
(**a**) Engineering



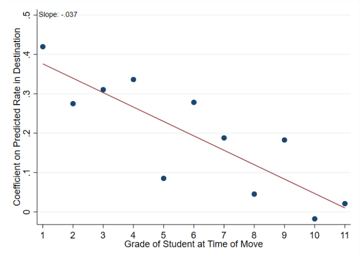
(**b**) Bio-science



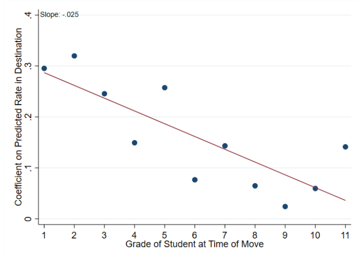
(**c**) Physical Sciences



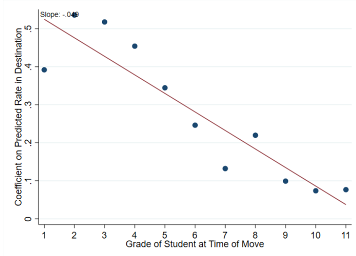
(**d**) Liberal Arts



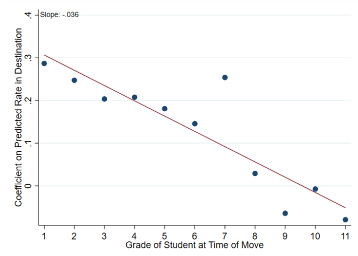
(**e**) Business



(**f**) Health Profession



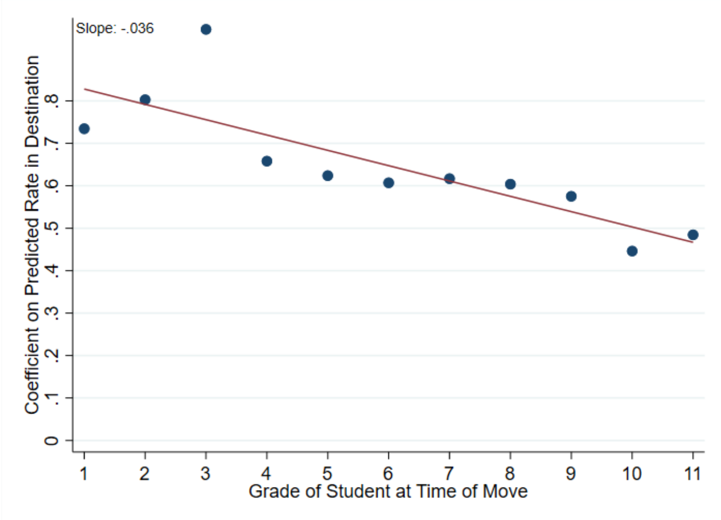
(**g**) Arts



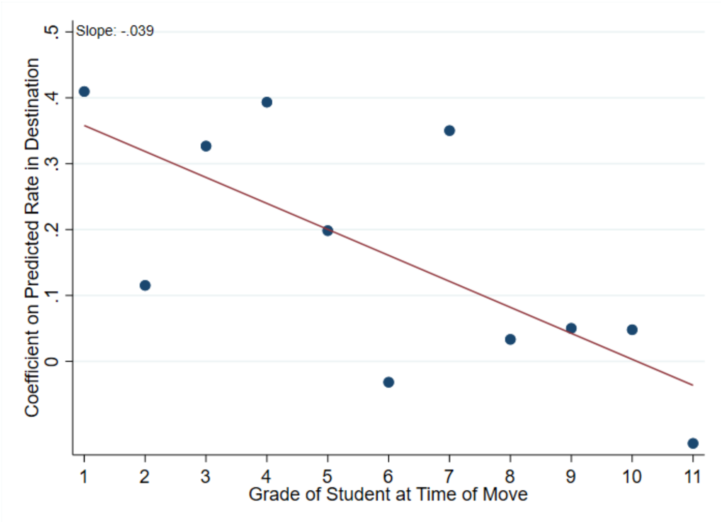
Notes: The figures plot estimates of the coefficients b_m across the students' grade when they move (m) using the parametric specification in equation (1), measuring students' degree attainment in the specified major on the label. We classify majors based on the 2010 two-digit CIP program codes as follows: engineering includes computer and information sciences and support services (11) and engineering (14); bio-science includes biological and biomedical sciences (26); physical science includes physical sciences (40) and mathematics and statistics (27); liberal arts include education (13), foreign languages, literatures, and linguistics (16), English language and literature/letters (23), liberal arts and sciences, general studies, and humanities (24), philosophy and religious studies (38), theology and religious vocations (39), social sciences (45), communication, journalism, and related programs (09), and history (54); business includes business, management, marketing, and related support services (52); health professions include health professions and related programs (51); and arts include visual and performing arts (50). The sample includes all students in kindergarten in 1994-2003 who moved only once in their school years. Students' major choice is measured as obtaining a undergraduate degree in the specified major by age 26. They are estimated by regressing the students' degree attainment in college $Major_i$ on $\Delta Major_{odg} = \overline{Major}_{dg}$ - $\overline{Major}_{og}$, the difference in predicted major attainment between the destination and origin counties, based on non-moving students from the same racial and economic status group, interacted with each grade of the student at the time of the move m. We include race-by-economic status-by-school cohort-by-grade at move fixed effects, interacted with predicted major rate for non-moving students from the same racial and economic status group in the origin. Best-fit lines are estimated using unweighted OLS regressions of the b_m coefficients on m. The slopes of the regression line are reported on the left side of each line.

Fig. A.9. Exposure Effects on STEM Choice: Enrollment and Graduate Degree

(a) Enrolling in Any College



(b) Obtaining in Graduate Degree



Notes: The figures plot estimates of the coefficients b_m across the students' grade when they move (m) using the parametric specification in equation (1), defining STEM choice differently. STEM choice is defined as enrolling in a STEM major in any college by age 26 for Figure (a) and obtaining a STEM graduate degree in Figure (b). This specification replicates the specification used in Figure 4 (b), replacing the STEM rates of counties and STEM choice of movers at the specified levels. The sample includes all students in kindergarten in 1994-2003 who moved only once in their school years. The b_m coefficients can be interpreted as the effect of moving to an area where non-moving students' STEM rate increases from 0 to 1 at grade m. They are estimated by regressing the students' STEM enrollment or graduate degree $STEM_i$ on $\Delta STEM_{odg}$, the difference between STEM enrollment rates or STEM graduation degree rates the destination vs. the origin, based on non-moving students from the same racial and economic status group, interacted with each grade of the student at the time of the move m. We include race-by-economic status-by-school cohort-by-grade at move fixed effects, interacted with predicted STEM rate for non-moving students from the same racial and economic status group in the origin. Best-fit lines are estimated using unweighted OLS regressions of the b_m coefficients on m. The slopes of the regression line are reported on the left side of each line.

Table A.1: Exposure Effects on STEM Major Choice: Robustness to Minimum Years of Observation

	(1) $\geq$ 6 years	(2) $\geq$ 7 yrs	(3) $\geq$ 8 yrs	(4) $\geq$ 9 yrs	(5) $\geq$ 10 yrs	(6) $\geq$ 11 yrs
Exposure effects	0.028***	0.029***	0.029***	0.029***	0.030***	0.031***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	217,048	212,306	205,777	196,169	177,906	146,120

Notes: The table reports estimates of annual exposure effects on students' STEM degree attainment by age 26, using different sample definitions based on the minimum number of years each student is observed (baseline: 10 years). The estimates can be interpreted as the impact of spending an additional school year in a county where the share of non-moving students pursuing STEM degrees is one percentage point higher. Each column reports estimates from a regression of a student's STEM degree attainment by 26 on the difference between neighborhood STEM rates in the destination vs. the origin, interacted with the grade of the student at the time of the move (m). Each regression also includes race-by-economic status-by-school cohort-by-grade at move fixed effects, interacted with STEM rate and additional controls specified in equation (3). Each column reports the estimate of γ from column (1) of Table 2, with the minimum number of years of observation indicated in the column header. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

Table A.2: Exposure Effects on STEM Major Choice: STEM Occupation Exposure Measure

	Parametric				Semi-parametric			
	(1) Baseline	(2) + Destination info	(3) + Test scores & Labor char.	(4) Catchment FE	(5) Baseline	(6) + Destination info	(7) + Test scores & Orig.×dest. FE	(8) Catchment FE
Exposure effects	0.017***	-0.017***	0.021***	0.026**	0.017**	-0.017**	0.018**	0.039**
	(0.003)	(0.004)	(0.004)	(0.010)	(0.007)	(0.007)	(0.008)	(0.017)
Observations	205,771	205,771	182,830	86,998	199,496	199,491	172,051	77,605

Notes: The table reports estimates of annual exposure effects on students' STEM degree attainment by age 26. The estimates can be interpreted as the impact of spending an additional school year in a county with one percentage point more STEM workers in a neighborhood. See Section 5 for more details on STEM occupation exposure measure construction. Each column reports estimates from a regression of a student's STEM degree attainment by 26 on the difference between neighborhood STEM occupation rates in the destination vs. the origin, interacted with the grade of the student at the time of the move (m). Column (1) reports the estimate of γ and its standard error from equation (3), using the sample of one-time movers defined in the notes to Table 1. Column (2) adds to column (1) fixed effects for race-by-economic status interacted with the destination STEM occupation rate. Column (3) further includes labor market characteristics (median income, employment rate, and poverty rate in origin and destination counties, constructed using the same method described in Section 5) and moving students' third-grade math and reading scores. Column (4) adds to column (1) finer fixed effects: race-by-economic status-by-origin school catchment area-by-destination catchment area-by-year of move. Columns (1) through (4) all include race-by-economic status-by-cohort-by-grade-of-move fixed effects, interacted with the origin STEM occupation rate. Columns (5) through (8) replicate the specifications in Columns (1) through (4), respectively, but replace the origin STEM occupation rate—which was interacted with race-by-economic status-by-cohort-by-grade-of-move fixed effects—with origin county fixed effects. Additionally, Column (7) includes origin-by-destination county fixed effects by race and economic status instead of labor market controls. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

Table A.3: Exposure Effect Estimates to Neighborhoods Occupational Composition: Robustness

	(1) Base	(2) +$\leq$ 0.5 mile	(3) +$\leq$ 1	(4) +$\leq$ 2	(5) +$\leq$ 3	(6) +$\leq$ 5	(7) Only Elementary
Exposure effects	0.017***	0.017***	0.018***	0.018***	0.018***	0.016***	0.019***
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)
Observations	205,771	205,777	205,777	205,777	205,777	205,777	205,773

Notes: The table reports estimates of annual exposure effects on students' STEM degree attainment by age 26, with different neighborhood STEM occupation measures. The estimates can be interpreted as the impact of spending an additional school year in a county with one percentage point more STEM workers in a neighborhood. Each column reports estimates from a regression of a student's STEM degree attainment by 26 on the difference between neighborhood STEM occupation rates in the destination vs. the origin, interacted with the grade of the student at the time of the move (m). Each regression also includes race-by-economic status-by-school cohort-by-grade at move fixed effects, interacted with STEM occupation rate and additional controls specified in equation (3). Column (1) reports the baseline estimate of γ from Table A.2 column (1), where the STEM occupation rate of school ($STEM_{c(s)}$) is defined as the STEM rate of the Census track school is located. Columns (2)–(6) define the school's STEM rate as the STEM rate of the census tracts within the specified radius (in miles) in each column, including the tract where the school is located. Column (7) includes only elementary schools to calculate the district occupation exposure measure. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

Table A.4: Occupation Exposure Effect Estimates: Gender Specific Convergence

	(1)	(2)	(3)
Exposure effects-own gender	0.018***		0.021***
	(0.003)		(0.004)
Exposure effects-other gender		0.013***	-0.004
		(0.003)	(0.004)
Observations	197,950	197,961	197,939

Notes: The table reports estimates of annual exposure effects γ using gender-specific STEM occupational rates. The estimates can be interpreted as the impact of spending an additional school year in a county with one percentage point more gender-specific STEM workers in a neighborhood. In all columns, the dependent variable is the student's STEM degree attainment by age 26. The sample consists of all students of one-time movers, defined in the notes to Table 1. The table shows gender-specific convergence to the STEM worker proportion. Column (1) replicates column (1) of Table A.2, replacing the STEM occupation rate based on all workers in the origin and destination with STEM occupation rates of the same gender as the student who moves. Column (2) replicates column (1), replacing the own-gender STEM occupation rate with the STEM occupation rate of the opposite gender. Column (3) combines the variables in columns (1) and (2), including both the own-gender and other-gender predictions. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

Table A.5: Mediation Analysis: Unbalanced Table

	Sample one			Sample two		Sample three	
	(1) Baseline	(2) +3rd test score	(2) +8th test score	(1) Baseline	(3) +11th test score	(1) Baseline	(4) +AP Courses
Exposure effects	0.031***	0.031***	0.029***	0.030***	0.031***	0.029***	0.020***
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)	(0.004)
Math score 3rd grade		0.019***					
		(0.001)					
Reading score 3rd grade		0.002**					
		(0.001)					
Math score 8th grade			0.039***				
			(0.002)				
Reading score 8th grade			-0.008***				
			(0.001)				
Math score 11th grade				0.039***			
				(0.001)			
Reading score 11th grade				-0.003***			
				(0.001)			
Non-science AP courses							0.002***
						(0.000)	
Science AP courses							0.038***
							(0.001)
Observations	176,388	176,388	176,388	141,597	141,597	205,777	205,777

Notes: The table reports estimates of annual exposure effects γ with different intermediate outcome variables. In all columns, the dependent variable is the student's STEM degree attainment by age 26. The sample consists of a group of students of one-time movers, defined in the notes to Table 1. Each column replicates column (1) of Table 2. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

B Previous Papers on College Major Choices, STEM Gaps, and Developmental Neighborhood Effects

B.1 College Major Choices

Empirical research on college major choice has identified several key mechanisms shaping students' decisions. A substantial body of research highlights the importance of expected earnings (Beffy, Fougère, and Maurel 2012; Long, Goldhaber, and Huntington-Klein 2015; Montmarquette, Cannings, and Mahseredjian 2002; Weinstein 2022). Closely related, several papers examine how providing students with information mainly about expected earnings and employment prospects affects their choices, showing that access to better information can lead to shifts in major selection toward more lucrative or secure fields (Ajzenman et al. 2021; Baker et al. 2018; Christian, Ronfeldt, and Zafar 2024; Conlon 2021; Ding et al. 2021; Wiswall and Zafar 2015).

Another line of research investigates the importance of individual beliefs and expectations about future outcomes, including confidence in academic ability or perceived fit with a major (Arcidiacono et al. 2020; Stinebrickner and Stinebrickner 2014; Wiswall and Zafar 2021). Other studies focus on non-pecuniary factors, such as personal interests or perceived social value of different fields (Astorne-Figari and Speer 2019; Ngo and Dustan 2024; Wiswall and Zafar 2015; Zafar 2013).

A separate strand of research examines the importance of educational environments. One set of papers focuses on the influence of peers on college major choice (Elsner, Isphording, and Zölitz 2021; Griffith and Main 2019; Ost 2010; Rask 2010), finding that students' relative performance, compared to their peers, can significantly affect their choice of major. Another set of papers studies the role of course availability, showing that offering sufficient coursework options prior to starting college is critical for students to pursue specific majors (Darolia et al. 2020; De Giorgi, Pellizzari, and Redaelli 2010; Gottfried and Bozick 2016; Liu, Conrad, and Blazar 2024).

In the context of STEM specifically, a growing literature has documented the role of preference and preparation, including high school course-taking and curriculum rigor, as a determinant of later STEM enrollment and persistence (Arold 2024; Darolia et al. 2020; De Philippis 2023; Jia 2021; Joensen and Nielsen 2016; Tai et al. 2006; Wang 2013). Family background and intergenerational transmission are also crucial, with multiple studies finding that parental skills, educational attainment, and occupational fields strongly correlate with children's propensity to pursue STEM degrees (Aina and Nicoletti 2018; Chise, Fort, and Monfardini 2021; Hanushek et al. 2023; Oguzoglu and Ozbeklik 2016).

B.2 STEM Gaps

A body of research exists seeking to explain the persistent underrepresentation of women and minority groups in STEM fields. This research highlights several key mechanisms operating across individual and educational dimensions. One prominent line of work focuses on differences in personal preferences and interests between men and women (Ceci et al. 2014; Ngo and Dustan 2024; Patnaik et al. 2022; Reuben, Wiswall, and Zafar 2017; Wiswall and Zafar 2018; Zafar 2013). These studies suggest that gender differences in certain values, like flexibility or stability, can steer students toward STEM or non-STEM fields.

Another set of research emphasizes disparities in STEM preparedness, documenting how differences in math performance, course selection, and high school rigor contribute to later gender and racial gaps in STEM college enrollment (Arcidiacono, Aucejo, and Hotz 2016; Card and Payne 2017; Delaney and Devereux 2021; Ellison and Swanson 2010; Justman and Méndez 2018; Saltiel 2023; Sovero, Buchinsky, and Baird 2021; Speer 2023). Another strand of literature highlights the role of non-cognitive factors, such as willingness to compete, attitudes toward risk, and confidence. Studies find that women, on average, report less willingness to enter competitive environments and greater aversion to risk, which can reduce their likelihood of pursuing STEM majors where these attributes are often perceived as important (Astorne-Figari and Speer 2019; Buser, Niederle, and Oosterbeek 2014; Delaney and Devereux 2021; Kugler, Tinsley, and Ukhaneva 2021).

Beyond individual characteristics, the educational environment has emerged as a critical factor in shaping STEM gaps. Research shows that having same-gender or same-race teachers can act as powerful role models, increasing students' engagement and persistence in STEM fields (Breda et al. 2023; Canaan and Mouganie 2021; Carrell, Page, and West 2010; Griffith and Main 2021; Hoffmann and Oreopoulos 2009; Oliver et al. 2021; Porter and Serra 2020; Price 2010). For example, Carrell, Page, and West (2010) finds that female students randomly assigned to female professors are significantly more likely to enroll in STEM courses and persist in STEM majors. Conversely, teacher biases can reinforce stereotypes and discourage underrepresented students from pursuing STEM tracks, as documented by Lavy and Sand (2015).

The competitive climate of educational settings can also influence STEM trajectories. For example, Landaud, Ly, and Maurin (2020) shows that highly competitive environments disproportionately deter girls from pursuing STEM pathways. Similarly, studies on peer composition and within-class ranking highlight that being surrounded by high-achieving peers or experiencing lower relative ranking can discourage female students from STEM majors, suggesting that local social comparison dynamics matter (Anelli, Shih, and Williams 2023; Bostwick and Weinberg 2022; Fischer 2017; Goulas, Griselda, and Megalokonomou 2024; Zölitz and Feld 2021).

Other research has examined socioeconomic disparities in STEM major choice. First, Castleman, Long, and Mabel (2018) finds that financial assistance is critical for boosting STEM choices among economically disadvantaged students. Moreover, Cohodes, Ho, and Robles (2022) finds that summer STEM enrichment programs significantly increase STEM degree completion rates for students from underrepresented backgrounds. Finally, Bleemer and Mehta (2021) finds that college major restrictions prevent STEM major choices among racial minority students.

As we have presented, there has been a long and extensive discussion on college (STEM) major choices and the persistence of STEM gaps across groups of students, with prior research highlighting individual preferences, beliefs, preparedness, family background, peer influences, teacher characteristics, and institutional factors. However, as noted by Xie, Fang, and Shauman (2015), there remains a critical gap in understanding the role of neighborhoods, particularly non-educational environments, in shaping students' trajectories into STEM. Our paper extends this literature by explicitly examining how neighborhood characteristics, such as local occupational composition, affect students' likelihood of choosing STEM majors. By bringing neighborhoods into the analysis, we offer new evidence that place-based contextual factors play a substantial role in perpetuating or narrowing gender, racial, and socio-economic disparities in STEM participation. These findings underscore that interventions to reduce STEM gaps cannot focus solely on schools or individual characteristics, but must also address broader neighborhood contexts that shape educational and occupational decisions.

B.3 Developmental Neighborhood Effects

Researchers have long documented how neighborhood characteristics shape individuals' life outcomes, including educational attainment, criminal behavior, health, and labor market success (Brooks-Gunn et al. 1993; Jencks, Mayer, et al. 1990; Sampson, Morenoff, and Gannon-Rowley 2002; Sampson, Raudenbush, and Earls 1997; Sharkey and Faber 2014; Wilson 2012; Wodtke, Harding, and Elwert 2011).

More recent studies have shifted toward estimating the causal impacts of place, employing quasi-experimental and randomized field experiments to disentangle neighborhood effects from selection bias. Some papers exploit unique geographic shocks. For example, research using the Moving to Opportunity experiment has demonstrated substantial positive impacts of improved neighborhood environments during childhood on outcomes such as education, health, and earnings (Chetty, Hendren, and Katz 2016; Katz, Kling, and Liebman 2001; Ludwig et al. 2013). Similarly, Nakamura, Sigurdsson, and Steinsson (2022) exploit volcanic lava destruction as an instrument for forced migration, finding a significant increase in earnings among children displaced by eruptions. Baran, Chyn, and Stuart (2024) leverage selective migration patterns in the 1940 Census, showing that Black children whose families moved North during the Great Migration completed more years of schooling than their peers who stayed in the South.

Other studies use public housing programs to estimate neighborhood impacts. Oreopoulos (2003) finds little evidence of neighborhood effects, whereas Chyn (2018) documents significant positive impacts on education, crime reduction, and earnings. The Chicago Gautreaux placement program, analyzed by Chyn, Collinson, and Sandler (2025), employs sibling fixed effects to reveal large gains in earnings and interracial marriage rates among children relocated to white suburbs. In addition, Barrios-Fernández (2022) uses student loan eligibility cutoffs to show how neighbors' choices influence an individual's likelihood of enrolling in college.

A growing body of work employs "mover" designs, tracing variation in outcomes to the age at which children relocate. Chetty and Hendren (2018) identifies significant effects of neighborhood exposure on intergenerational mobility in the US, while Deutscher (2020) matches Australian tax records with census data and finds consistent results. Laliberté (2021) combines a school-boundary regression discontinuity with an exposure design, showing that more than half of place effects operate through primary school quality. Similar positive impacts have also been documented in developing countries: in Africa, Alesina et al. (2021) and Maarseveen (2021) find significant benefits of moving to better neighborhoods, focusing on children's educational outcomes. Similarly, in Indonesia, Schwank (2024) reports convergence of children's outcomes toward those of their new neighborhoods. Lastly, Bell et al. (2019) find that children who move to high-innovation areas are more likely to become inventors.

While these studies provide compelling evidence on the importance of neighborhood environments for various outcomes, they primarily focus on better-versus-worse neighborhood comparisons in broad domains such as income or education levels. Our paper extends this literature by shifting the lens to individual decision-making, specifically examining how neighborhood characteristics influence college major choices, with a focus on STEM selection. This approach offers new insights into how place-based factors shape not just overall outcomes but also critical educational and occupational trajectories.

C Estimates Including Out-of-State STEM Attainment

To test whether out-of-state college attainment affects our estimates, we use data from the National Student Clearinghouse (NSC), which covers approximately 98% of higher education enrollment in the United States. We access two types of NSC data. First, we use data on out-of-state college graduates from high school cohorts between 2009 and 2016, which includes information on college majors and detailed CIP codes.[C.1] However, this dataset does not specify the level of degree earned (e.g., certificate, associate, bachelor's, or graduate).

Second, we use data for high school cohorts graduating between 2015 and 2018, which includes information on college graduates both within and outside of Texas. For roughly half of these graduates, we observe both their degree type and major. However, the reported majors are not detailed enough to be directly matched to the DHS-defined list of STEM fields. To address this, we focus on students who graduated from colleges in Texas, for whom we have detailed major information and STEM classification from the THECB data. Using this subset, we calculate the STEM share within each combination of degree type and reported major. If more than half of the students in a given category are classified as STEM majors based on the DHS definition, we categorize that combination as STEM.

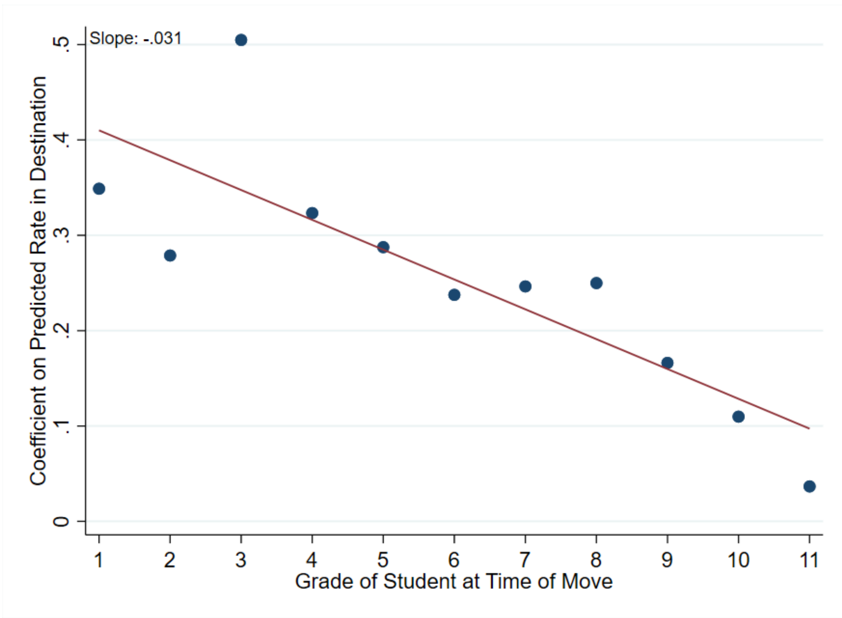
While the data does not fully cover our sample of students who entered kindergarten between 1994 and 2003, it likely includes those who started between 1996 and 2003—corresponding to high school graduation cohorts from 2009 to 2016. Thus, we believe the majority of students who earned out-of-state STEM degrees are captured. We update our STEM exposure measure ($\overline{STEM}_{pg}$) and STEM choice variable ($STEM_i$) by incorporating students who obtained a STEM degree outside of Texas. These out-of-state STEM graduates account for 4.7% of total STEM degree attainment among non-moving students and 4.2% of total STEM degree attainment among students who moved once, which corresponds to the average out-of-state college enrollment 5.3% (THECB 2017).

Appendix Figure C.1 and Table C.1 report results analogous to Figure 4 and Table 2, respectively. While the estimated effect is slightly smaller at $\gamma = 0.028$ (s.e. = 0.004), compared to the baseline estimate of $\gamma = 0.030$, the results remain overall consistent with the baseline findings.

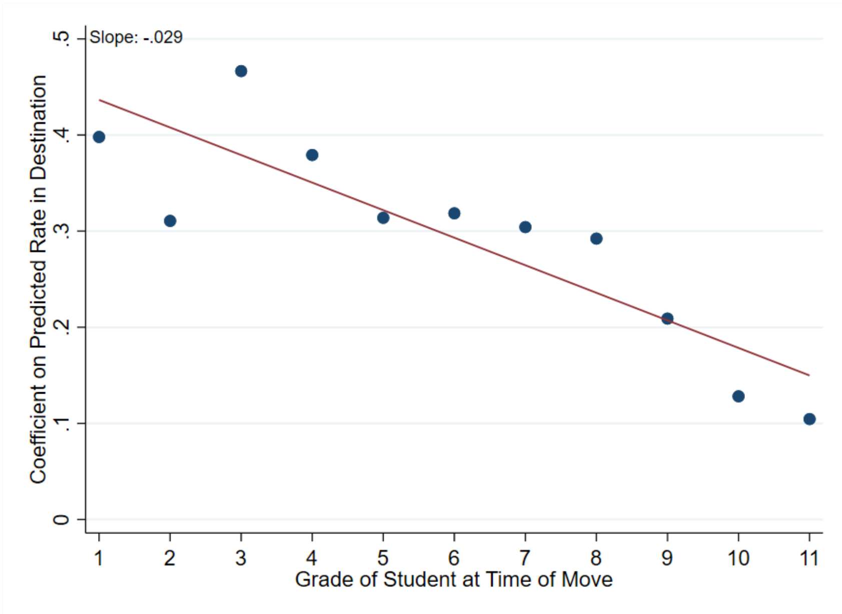
[C.1] About 20% of CIP codes are missing among college graduates who report their majors. Because college majors are broader categories than CIP codes, we impute the missing codes by assigning the most frequent CIP code associated with each reported major. This method allows us to impute CIP codes for approximately 65% of the missing cases.

Fig. C.1. Exposure Effects on STEM Choice in College: Including Out-of-State STEM Choice

(a) Semi-parametric Estimates



(b) Parametric Estimates



Notes: Figure (a) plots estimates of the coefficients b_m across the students' grade when they move (m) using the semi-parametric specification in equation (1), measuring students' STEM degree attainment including degrees earned both within and outside of Texas. The sample includes all students in kindergarten in 1994-2003 who moved only once in their school years. Students' STEM choice $STEM_i$ is measured as obtaining an undergraduate degree in STEM majors by age 26. The b_m coefficients can be interpreted as the effect of moving to an area where non-moving students' STEM rate increases from 0 to 1 at age m. They are estimated by regressing the students' STEM degree attainment in college $STEM_i$ on $\Delta STEM_{odg} = \overline{STEM}_{dg} - \overline{STEM}_{og}$, the difference in predicted STEM attainment between the destination and origin counties, based on non-moving students from the same racial and economic status group, interacted with each grade of the student at the time of the move m. We include origin county-by-race-by-economic disadvantage status-by-school cohort-by-grade at move fixed effects when estimating this specification. Figure (b) plots estimates from the parametric specification in equation (2). This specification replicates the specification used in Figure (a), replacing the origin fixed effects with the STEM rate for non-movers in the origin. Best-fit lines are estimated using unweighted OLS regressions of the b_m coefficients on m. The slopes of the regression line are reported on the left side of each panel. The magnitudes of the slope represent estimates of annual exposure effects.

Table C.1: Exposure Effects in STEM Major Choice: Including Out-of-State STEM Choice

	Parametric				Semi-parametric			
	(1) Baseline	(2) + Destination info	(3) + Test scores & Labor char.	(4) Catchment FE	(5) Baseline	(6) + Destination info	(7) + Test scores & Orig.×dest. FE	(8) Catchment FE
Exposure effects	0.028***	0.026***	0.028***	0.034***	0.029***	0.027***	0.028***	0.037*
	(0.004)	(0.004)	(0.004)	(0.011)	(0.005)	(0.005)	(0.006)	(0.019)
Observations	205,777	205,777	182,835	87,000	199,502	199,497	172,057	77,607

Notes: This table reports estimates of annual exposure effects on students' STEM degree attainment by age 26, including STEM attainment both within and outside of Texas. The coefficients can be interpreted as the impact of spending an additional school year in a county where non-moving students have a one percentage point higher STEM attainment rate. Each column presents results from a regression of students' STEM degree attainment on the difference in STEM rates between destination and origin counties for non-moving students of the same race and economic status group, interacted with the grade at which the student moved (m). Column (1) reports the estimate of γ and its standard error from equation (3), using the sample of one-time movers defined in the notes to Table 1. Column (2) adds to column (1) fixed effects for race-by-economic status interacted with the destination STEM rate of non-movers from the same race and economic status group. Column (3) further includes labor market characteristics (median income, employment rate, and poverty rate in origin and destination counties, constructed using the same method described in Section 5) and moving students' third-grade math and reading scores. Column (4) adds to column (1) finer fixed effects: race-by-economic status-by-origin school catchment area-by-destination catchment area-by-year of move. Columns (1) through (4) all include race-by-economic status-by-cohort-by-grade-of-move fixed effects, interacted with the origin STEM rate of non-movers from the same race and economic status group. Columns (5) through (8) replicate the specifications in Columns (1) through (4), respectively, but replace the origin STEM rate of non-movers—which was interacted with race-by-economic status-by-cohort-by-grade-of-move fixed effects—with origin county fixed effects. Additionally, Column (7) includes origin-by-destination county fixed effects by race and economic status instead of labor market controls. Standard errors are shown in parentheses and clustered at the origin county level. *** $p<0.01$, ** $p<0.05$, * $p<0.10$

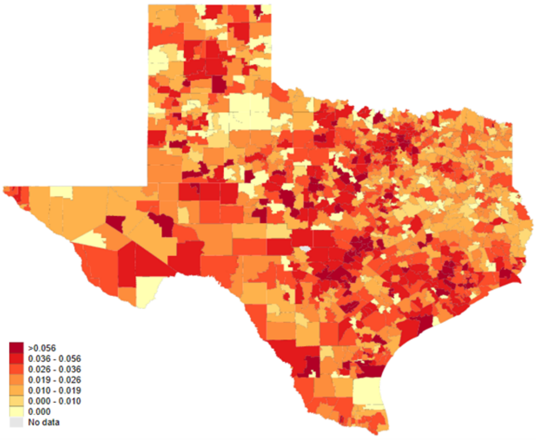
D School District Level Estimates

We present estimates of exposure effects using school districts as the geographic unit of analysis, instead of counties. The school districts-level STEM exposure, denoted as $STEM_p$, is calculated using the same method described in Section 2. We then regress students' STEM choices on the difference in STEM exposure between their destination and origin school districts among the same race and economic status (g), $\Delta STEM_{odg} = \overline{STEM}_{dg} - \overline{STEM}_{og}$, interacted with the grade at which they moved. This analysis follows a specification similar to that of Equations (1, 2, and 3), but applied at the school district level.

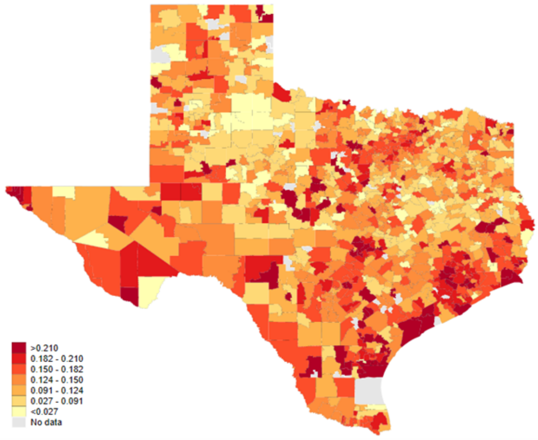
As in the baseline analysis, we focus on students who moved once between school districts. The sample is further restricted to origin and destination school districts with at least 2,000 non-moving students from the same race and economic status group within a 10-year school cohort. Appendix Figures D.1 and D.2 and Table D.1 report results analogous to Figures 2, 4, and Table 2, respectively. We do not conduct the catchment-area-level analysis here, as school district boundaries often overlap with catchment areas in smaller districts, making it difficult to separately identify the exposure effects. While the estimated effect is slightly smaller at $\gamma = 0.027$ (s.e. = 0.007), compared to the baseline estimate of $\gamma = 0.029$, the results remain overall consistent with the baseline findings.

Fig. D.1. STEM Rate and Its Persistence Across School Districts

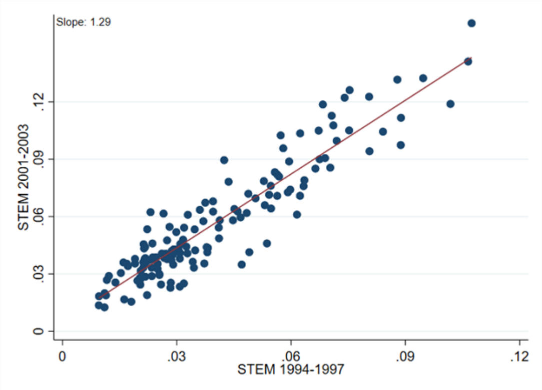
(a) STEM rate: $\frac{\text{STEM BAs}}{\text{All students}}$



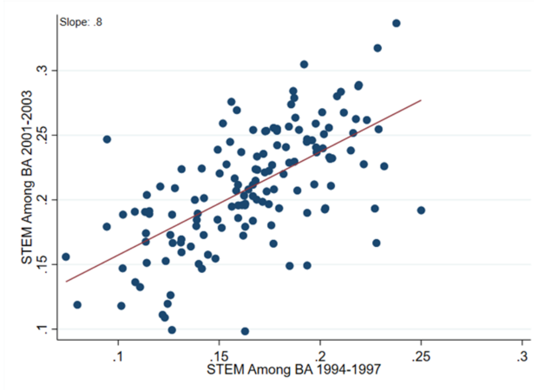
(b) STEM rate among BAs: $\frac{\text{STEM BAs}}{\text{All BAs}}$



(c) STEM rate across years



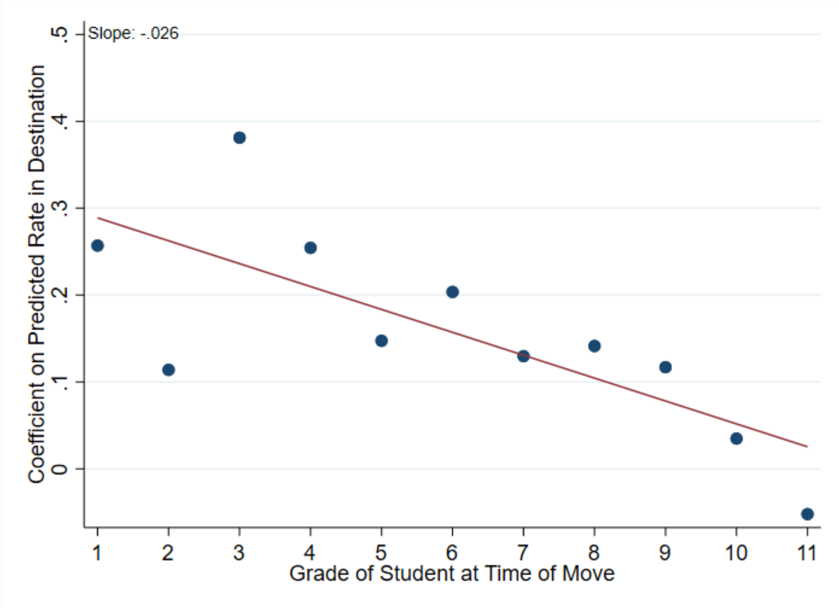
(d) STEM rate among BAs across years



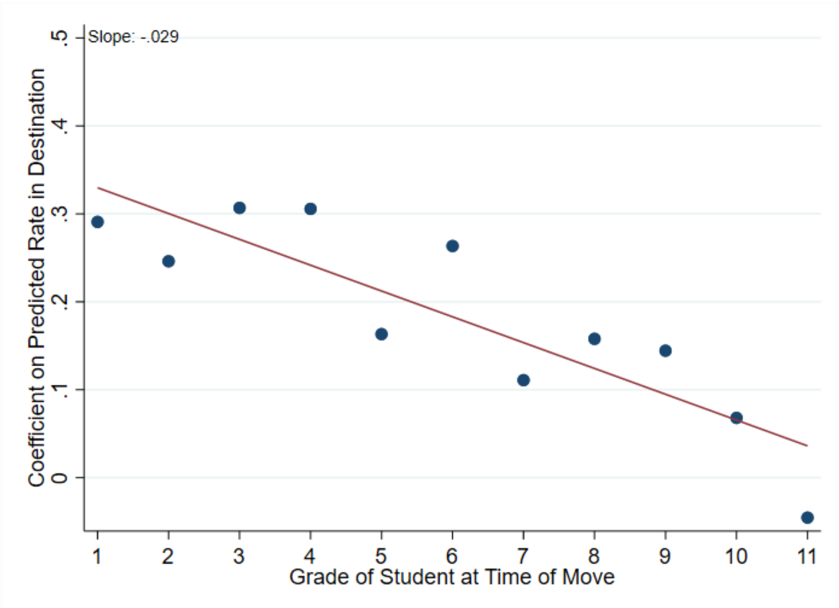
Notes: These maps plot the proportion of non-moving students obtaining STEM undergraduate degrees by age 26 among all students in Figure (a) and among undergraduate degree holders in Figure (b). The maps are constructed by grouping school districts into seven quantiles and shading the areas so that darker colors correspond to higher outcomes for students. Areas with no data are shaded with a gray color. The sample includes all students in the kindergarten in 1994-2003 who did not move across school districts in their school years. Figures (c) and (d) plot the proportion of non-moving students obtaining an undergraduate degree in STEM majors in Texas school districts from 1994-1997 school cohorts starting kindergarten in that corresponding year, compared to 2001-2003 school cohorts. The proportion in Figure (c) is calculated among all students in each county, and (d) is among all undergraduate degree holders. STEM majors are defined following the Department of Homeland Security's 2016 STEM major list.

Fig. D.2. Exposure Effects on STEM Choice in College: School District Level

(a) Semi-parametric Estimates



(b) Parametric Estimates



Notes: Figure (a) plots the estimated coefficients b_m by students' grade at the time of their move (m), using the semi-parametric specification in equation (1). The outcome variable, $STEM_i$, measures whether a student obtained an undergraduate degree in a STEM field by age 26. The sample includes all students who entered kindergarten between 1994 and 2003 and moved only once during their school years. The b_m coefficients represent the effect of moving at grade m to a school district where the STEM rate among non-moving students in the same racial and economic status group increases from 0 to 1. These coefficients are estimated by regressing $STEM_i$ on $\Delta STEM_{odg} = \overline{STEM}dg - \overline{STEM}og$—the difference in predicted STEM attainment between destination and origin school districts—interacted with the grade of the move. The model includes fixed effects for origin school district-by-race-by-economic status-by-cohort-by-grade at move. Figure (b) presents estimates from the parametric specification in equation (2), which mirrors the approach in Figure (a) but replaces the origin fixed effects with the STEM rate for non-movers in the origin school district. In both figures, best-fit lines are generated by unweighted OLS regressions of the b_m coefficients on m, and the estimated slopes, which are shown on the left side of each figure. The magnitudes of the slope represent estimates of annual exposure effects.

Table D.1: Exposure Effects in STEM Major Choice: School District Level

	Parametric			Semi-parametric		
	(1) Baseline	(2) + Destination info	(3) + Test scores & Labor char.	(4) Baseline	(5) + Destination info	(6) + Test scores & Orig.×dest. FE
Exposure effects	0.027*** (0.007)	0.024*** (0.007)	0.026*** (0.007)	0.022** (0.010)	0.018* (0.010)	0.023** (0.011)
Observations	184,071	184,071	152,862	178,054	178,022	140,952

Notes: This table reports estimates of annual exposure effects on students' STEM degree attainment by age 26. The coefficients can be interpreted as the impact of spending an additional school year in a school district where non-moving students have a one percentage point higher STEM attainment rate. Each column presents results from a regression of students' STEM degree attainment on the difference in STEM rates between destination and origin school districts for non-moving students of the same race and economic status group, interacted with the grade at which the student moved (m). Column (1) reports the estimate of γ and its standard error from equation (3), using the sample of one-time movers defined in the notes to Table 1. Column (2) adds to column (1) fixed effects for race-by-economic status interacted with the destination STEM rate of non-movers from the same race and economic status group. Column (3) further includes labor market characteristics (median income, employment rate, and poverty rate in origin and destination school districts, constructed using the same method described in Section 5) and moving students' third-grade math and reading scores. Column (4) adds to column (1) finer fixed effects: race-by-economic status-by-origin school catchment area-by-destination catchment area-by-year of move. Columns (1) through (4) all include race-by-economic status-by-cohort-by-grade-of-move fixed effects, interacted with the origin STEM rate of non-movers from the same race and economic status group. Columns (5) through (8) replicate the specifications in Columns (1) through (4), respectively, but replace the origin STEM rate of non-movers—which was interacted with race-by-economic status-by-cohort-by-grade-of-move fixed effects—with origin school district fixed effects. Additionally, Column (7) includes origin-by-destination school district fixed effects by race and economic status instead of labor market controls. Standard errors are shown in parentheses and clustered at the origin school district level. *** p<0.01, ** p<0.05, * p<0.10

E Multiple Movers

In this appendix, we extend our baseline analysis by estimating exposure effects for students who move across counties multiple times during their school years. This analysis allows us to assess whether the main results remain stable when we analyze students with more complex residential histories.

As in the baseline estimation, we restrict the sample to students who began kindergarten between 1994 and 2003, were observed for at least 10 school years, and for whom there are at least 2,000 non-moving students in both the origin and each destination county within the same race and economic status group. Among these students, we focus on those who moved across counties up to three times. Our goal is to examine the robustness of the estimated exposure effects when accounting for multiple residential transitions.

To formalize this, let $d(j)$ denote the jth destination county, with $j \in \{1,2,3\}$ indicating the sequence of moves. For each destination j, define e_{ij} as the number of school years that student i spent in county $d(j)$.

For each county the student resides in, we construct the exposure measure $\Delta^j STEM_{od(j)g}$, which captures the difference in STEM major rates between the destination county $d(j)$ and the student's origin county o, within the same race and economic status group g. This setup allows us to estimate a separate exposure effect for each move, accounting for the number of years spent in each county.

We estimate the following specification, which generalizes the baseline linear exposure model to accommodate multiple moves:

$$
\begin{aligned}
STEM_i = & \sum_{j=1}^{3} \gamma^j e_{ij} \Delta STEM^j_{od(j)g} + \sum_{j=1}^{3} \theta_1^j \Delta STEM_{od(j)g} + \sum_{j=1}^{3} \theta_2^j X_{od(j)g} \\
& + \sum_{j=1}^{3} \sum_{m(j),s,g} I(m_i = m(j),\, s_i = s,\, g_i = g) \left(\alpha^1_{m(j)sg} + \alpha^2_{m(j)sg} \overline{STEM}_{og} \right) + \varepsilon_i,
\end{aligned}
\tag{E.1}
$$

The parameters of interest in equation (E.1) are the coefficients γ^j, which represent the marginal effect of additional year exposure to higher-STEM counties for each move j. We include a comprehensive set of controls: fixed effects for race-by-economic status-by-cohort-by-grade at move, interacted with the average STEM rate in the student's origin county from the same race and economic status group ($\overline{STEM}_{og}$) and a vector of county characteristics $X_{od(j)g}$, including the undergraduate degree attainment rates for the origin and destination counties for each move j.

Table E.1 presents the estimation results. Column (1) reports estimates from the specification in equation (E.1). We find consistent exposure effects for each move: γ^1 = 0.018 (s.e. = 0.006), γ^2 = 0.033 (s.e. = 0.005), and γ^3 = 0.032 (s.e. = 0.006). These results suggest that the effect of neighborhood STEM exposure remains strong and consistent even when students move multiple times during their school years. In Column (2), we impose the restriction that the exposure effect is constant across moves by setting $\gamma^1 = \gamma^2 = \gamma^3 = \gamma$, and replacing the term $\sum_{j=1}^{3} \gamma^j e_{ij} \Delta STEM^j_{od(j)g}$ with $\gamma \sum_{j=1}^{3} e_{ij} \Delta STEM^j_{od(j)g}$. Under this constraint, the estimated exposure effect is γ = 0.028 (s.e. = 0.003).

Columns (3) and (4) replicate the specifications in Columns (1) and (2), respectively, but modify the control strategy by replacing the origin STEM rate with origin county fixed effects, as outlined in equation (1). Using this alternative specification, we estimate γ^1 = 0.023 (s.e. =

0.009), $\gamma^2 = 0.035$ (s.e. = 0.006), and $\gamma^3 = 0.028$ (s.e. = 0.007) in Column (3). In Column (4), under the constraint of a common exposure effect across moves, we find $\gamma = 0.030$ (s.e. = 0.004).

Together, these findings reinforce the robustness of our estimate on the exposure effect. Even when focusing on students who move multiple times, the results consistently show that moving to higher-STEM counties during one's school years has a meaningful and persistent impact on the likelihood of pursuing a STEM major. The estimated effects are similar in magnitude to those for students who moved only once, suggesting that the results are generalizable across different types of movers.

Table E.1: Exposure Effects Estimates Using Students who Move Multiple Times

	Parametric Estimation		Semi-Parametric Estimation	
	Separate Coefficients	Constrained Coefficient	Separate Coefficients	Constrained Coefficient
1st Destination Exposure Effect	0.018***		0.023**	
	(0.006)		(0.009)	
2nd Destination Exposure Effect	0.033***		0.035***	
	(0.005)		(0.006)	
3rd Destination Exposure Effect	0.032***		0.028***	
	(0.006)		(0.007)	
Average Exposure Effects		0.028***		0.030***
		(0.003)		(0.004)
Observations	196,731	196,731	172,829	172,829

Notes: The table reports estimates of the exposure effect, γ, using a sample of students who moved up to three times across counties during their school years. Column (1) presents estimates of separate exposure effects for the 1st, 2nd, and 3rd destination counties, allowing for heterogeneous impacts by move order, using the specification described in Appendix E. Column (2) imposes the restriction that the exposure effect is constant across all moves and reports a single coefficient estimate. Columns (3) and (4) replicate the specifications in Columns (1) and (2), respectively, but replace the origin STEM rate of non-movers—which was interacted with race-by-economic status-by-cohort-by-grade-of-move fixed effects—with origin county fixed effects. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

F Conceptual Framework for the Baseline Estimation Strategy

To clarify the intuition behind our empirical strategy, consider a hypothetical experiment in which students are randomly assigned to different locations p at a particular grade m and remain in that location for the rest of their school years, following Chetty and Hendren (2018). In such a setting, we can model a student's likelihood of majoring in STEM, denoted $STEM_i$, as a function of the average STEM rate among non-moving students in county p:

$$STEM_i = \alpha + \beta_m \overline{STEM}_p + \theta_i, \tag{F.1}$$

where $\overline{STEM}_p$ represents the average STEM major rate in county p, and θ_i captures individual-level unobservables such as family background, ability, or preferences that influence STEM choices. In this framework, $\overline{STEM}_p$ reflects the total influence of the local environment—encompassing schools, peers, community norms, and other contextual factors—on students' educational trajectories.

Under random assignment, $\overline{STEM}_p$ is uncorrelated with θ_i, and the coefficient β_m can be interpreted as the causal effect of spending grade m and all subsequent grades in a place with one percentage point higher STEM exposure. The marginal exposure effect of spending an additional grade in such a place—our parameter of interest—is given by the difference $\gamma_m = \beta_m - \beta_{m+1}$.

In practice, however, students are not randomly assigned to neighborhoods. Instead, we observe students who move across places at different points in their schooling, which introduces potential selection concerns. Specifically, in observational data, the error term θ_i may be correlated with the exposure variable $\overline{STEM}_p$, as families' decisions about where to move may be influenced by unobserved factors related to their children's educational trajectories.

To account for this, we define the observed regression coefficient from estimating equation (F.1) for students who moved at grade m as:

$$b_m = \beta_m + \delta_m,$$

where δ_m represents a selection bias term. This term can be expressed as:

$$\delta_m = \frac{\mathrm{cov}(\theta_i, \overline{STEM}_p)}{\mathrm{var}(\overline{STEM}_p)},$$

capturing the extent to which unobserved determinants of students' STEM choices are correlated with the average STEM rate of their destination place. Thus, b_m conflates the true exposure effect β_m with selection into destination locations.

Importantly, our identification strategy does not require that families' choices of destination places are random or uncorrelated with their children's future outcomes. Instead, the key identifying assumption is that the timing of the move—specifically, the grade at which a student relocates to a higher- or lower-STEM exposure area—is conditionally independent of the student's potential STEM major choice. In other words, while families may sort into different locations based on unobserved characteristics, we assume that, conditional on observable controls and fixed effects, the grade at which the move occurs is not systematically related to unobserved determinants of STEM outcomes. This allows us to isolate variation in exposure due to differences in the timing of the move, rather than differences in destination characteristics.

This framework motivates our empirical approach, which leverages variation in the timing of moves across grades to estimate grade-specific exposure effects. By holding origin and destination constant and focusing on when students move, we can net out the influence of selection into destination places. To further address potential confounding, we include rich fixed effects and covariates—such as race-by-economic-status-by-cohort-by-grade fixed effects and controls for origin and destination characteristics—to absorb baseline differences across students and families. This strategy enables us to credibly estimate the impact of neighborhood STEM exposure on students' long-term educational choices.

G Heterogeneous Exposure Effects by Grade Level

We proceed under the assumption that the exposure effects are consistent across grades, in line with the estimation results shown in Figure 4. While the linear trend provides a good overall fit, the figure also suggests that the effect of neighborhood STEM exposure is more pronounced for students who move during their later school years (grades 8–11). To explore this further, we replicate the main analysis from Figure 4 and Table 2, dividing the sample into two periods: grades 1–7 and grades 8–11.

Appendix Figure G.1 shows that this two-period division offers a better fit for the data. The estimated slope for students who moved during grades 1–7 is approximately 0.18, while for those who moved during grades 8–11, it increases sharply to 0.68. These results are consistent with the linearized estimates in Appendix Table G.1, further supporting the idea that exposure during later grades has a stronger impact on STEM major choice. This suggests that these later grades may be particularly formative for shaping students' academic trajectories and future career interests in STEM.

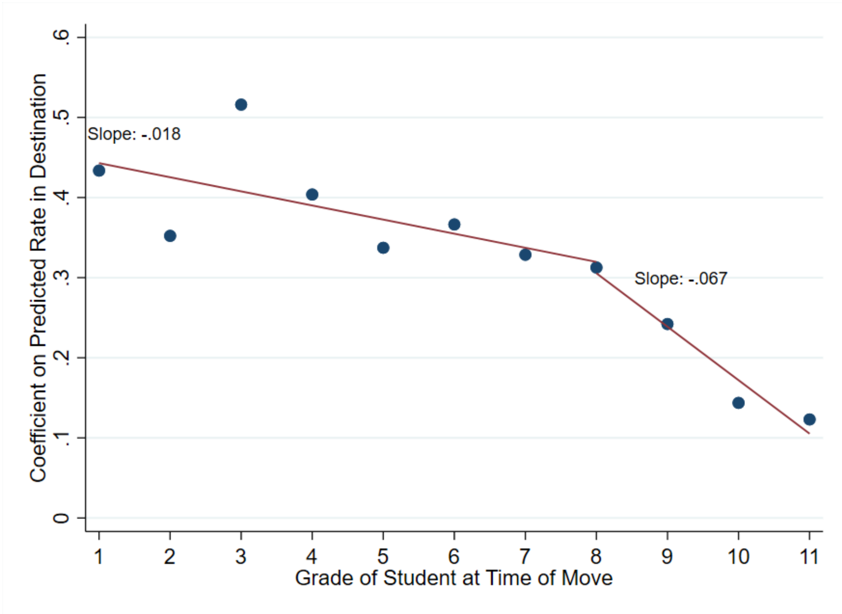
These findings also reinforce our identifying assumption that selection effects—captured by the correlation between unobserved determinants of STEM choice and the STEM rates of destination counties—are relatively constant across grades. As described in Section F, our identification strategy relies not on the assumption that students are randomly assigned to locations, but rather that the timing of moves is conditionally exogenous. If families systematically timed their moves based on factors related to students' future STEM choices—such as motivation, ability, or unobserved preferences—this would bias our estimates of exposure effects.

However, it is difficult to argue that such selection would differ meaningfully between those who move during early grades (1–7) and those who move during later grades (8–11). In fact, moving during students later grades often reflects exogenous factors such as job changes or housing needs, rather than long-term academic planning. Thus, the larger exposure effect estimated for grades 8–11 is more plausibly interpreted as a genuine increase in students' responsiveness to STEM environments during their later grades, rather than as the result of differential selection. This interpretation strengthens the causal interpretation of the exposure effects.

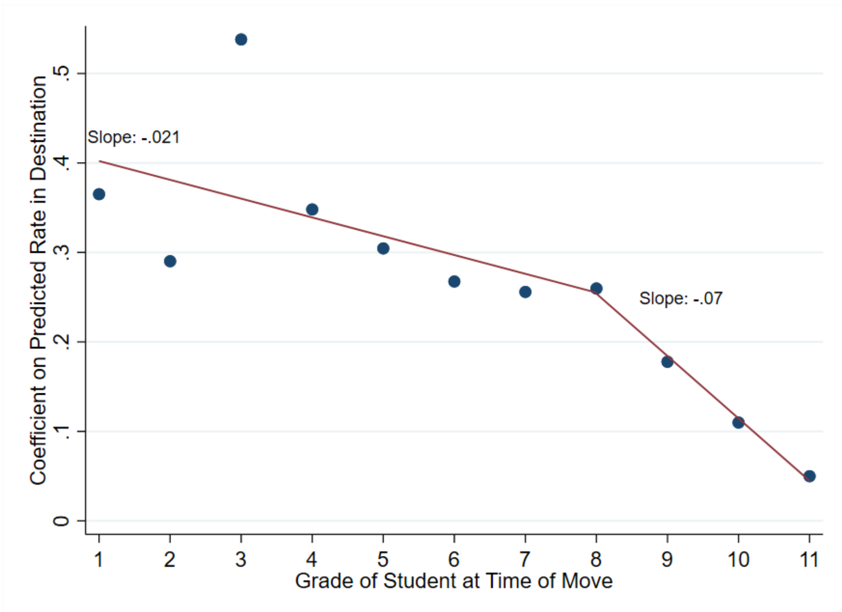
We do not use this grade-split specification in our main analysis due to sample size limitations. Several of our alternative specifications rely on subsamples that are too small to support precise estimates when further divided by grade ranges. Nonetheless, the patterns observed here provide evidence that the timing of exposure matters, with late school years being especially influential.

Fig. G.1. Exposure Effects on STEM Choice in College by Grade Range

(a) Semi-parametric Estimates



(b) Parametric Estimates



Notes: Figure (a) plots estimates of the coefficients b_m across the students' grade when they move (m) using the semi-parametric specification in equation (1), measuring students' STEM degree attainment. The sample includes all students in kindergarten in 1994-2003 who moved only once in their school years. Students' STEM choice $STEM_i$ is measured as obtaining an undergraduate degree in STEM majors by age 26. The b_m coefficients can be interpreted as the effect of moving to an area where non-moving students' STEM rate increases from 0 to 1 at age m. They are estimated by regressing the students' STEM degree attainment in college $STEM_i$ on $\Delta STEM_{odg} = \overline{STEM}_{dg}$ - $\overline{STEM}_{og}$, the difference in predicted STEM attainment between the destination and origin counties, based on non-moving students from the same racial and economic status group, interacted with each grade of the student at the time of the move m. We include origin county-by-race-by-economic disadvantage status-by-school cohort-by-grade at move fixed effects when estimating this specification. Figure (b) plots estimates from the parametric specification in equation (2). This specification replicates the specification used in Figure (a), replacing the origin fixed effects with the STEM rate for non-movers in the origin. Best-fit lines are estimated using unweighted OLS regressions of the b_m coefficients on m, separately for grades 1–7 and 8–11. The slopes of the regression line are reported. The magnitudes of the slope represent estimates of annual exposure effects.

Table G.1: Exposure Effects in STEM Major Choice by Grade Range

	Parametric				Semi-parametric			
	(1) Baseline	(2) + Destination info	(3) + Test scores & Labor char.	(4) Catchment FE	(5) Baseline	(6) + Destination info	(7) + Test scores & Orig.×dest. FE	(8) Catchment FE
Exposure effects	0.016*	0.016*	0.022**	0.021	0.019**	0.016*	0.020	0.023
Grade 1-7	(0.009)	(0.008)	(0.009)	(0.019)	(0.009)	(0.009)	(0.012)	(0.025)
Exposure effects	0.066***	0.064***	0.067***	0.070	0.069***	0.069***	0.095***	0.118
Grade 8-11	(0.021)	(0.022)	(0.021)	(0.072)	(0.021)	(0.022)	(0.025)	(0.100)
Observations	205,777	205,777	182835	87,000	199,502	199497	172,057	77,607

Notes: This table reports estimates of annual exposure effects on students' STEM degree attainment by age 26, separately for grades 1–7 and 8–11. The estimates can be interpreted as the impact of spending an additional school year in a county where non-moving students have a one percentage point higher STEM attainment rate. Each column reports estimates from a regression of a student's STEM degree attainment by 26 on the difference between non-moving students' STEM rates in the destination vs. the origin from the same race and economic disadvantage status, interacted with the grade of the student at the time of the move (m). Column (1) reports the estimate of γ and its standard error from equation (3), using the sample of one-time movers defined in the notes to Table 1. Column (2) adds to column (1) fixed effects for race-by-economic status interacted with the destination STEM rate of non-movers from the same race and economic status group. Column (3) further includes labor market characteristics (median income, employment rate, and poverty rate in origin and destination counties, constructed using the same method described in Section 5) and moving students' third-grade math and reading scores. Column (4) adds to column (1) finer fixed effects: race-by-economic status-by-origin school catchment area-by-destination catchment area-by-year of move. Columns (1) through (4) all include race-by-economic status-by-cohort-by-grade-of-move fixed effects, interacted with the origin STEM rate of non-movers from the same race and economic status group. Columns (5) through (8) replicate the specifications in Columns (1) through (4), respectively, but replace the origin STEM rate of non-movers—which was interacted with race-by-economic status-by-cohort-by-grade-of-move fixed effects—with origin county fixed effects. Additionally, Column (7) includes origin-by-destination county fixed effects by race and economic status instead of labor market controls. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

H Validation of Baseline Design

There is a long-standing economics literature of using mover designs to study the effects of neighborhoods (Aaronson 1998; Chetty, Friedman, and Saez 2013; Chetty, Hendren, and Katz 2016; Finkelstein, Gentzkow, and Williams 2016; Katz, Kling, and Liebman 2001; Ludwig et al. 2013; Oreopoulos 2003; Rosenbaum 1995). These mover designs were further developed and made more generalizable in Chetty and Hendren (2018), and have continued to be used to extract causal interpretations from neighborhoods. Building on this growing body of work, we conduct a series of validity tests for our baseline design. We begin by summarizing the validity tests used in recent studies in this literature and then discuss the results of our validity checks in detail.

Prior studies validate exposure effects using overidentification tests that examine heterogeneity across subgroups (Bell et al. 2019; Chetty and Hendren 2018; Chyn 2018; Derenoncourt 2022; Deutscher 2020; Finkelstein, Gentzkow, and Williams 2021; Laliberté 2021; Schwank 2024). The intuition behind these tests is that if the observed effects were driven by selection, we would be unlikely to see convergence in outcomes that is specific to particular groups. In Appendix H.1, we conduct similar analyses and find that students' outcomes converge to those of peers in the same gender and cohort group. Moreover, this convergence appears to be specific to the field of study.

Several papers also use family fixed effects to address concerns that families with more advantageous unobserved characteristics may choose to move to better neighborhoods at earlier ages (Alesina et al. 2021; Chetty and Hendren 2018; Deutscher 2020; Laliberté 2021; Maarseveen 2021; Schwank 2024). In general, including family fixed effects in these studies either confirms the original findings or results in larger estimated exposure effects. While we are unable to include family fixed effects directly due to data limitations, we proxy for family fixed effects by creating highly granular comparison groups, as detailed in Appendix H.2, and find that our results are consistent under this approach. We also account for changes in students' economic status at the time of the move, using a similar approach to that in Chetty and Hendren (2018), and find consistent results.

Previous research also restricts samples to moves likely driven by exogenous displacement to mitigate concerns that families systematically choose when to move (Alesina et al. 2021; Chetty and Hendren 2018; Maarseveen 2021). They further test the robustness of their estimates to the endogenous choice of destination. Alesina et al. (2021) and Derenoncourt (2022) use past migration patterns to predict destinations using a shift-share instrumental variables (IV) approach. Following these approaches, we restrict our sample to moves likely driven by exogenous displacement and use a shift-share IV strategy based on migration patterns of students in other cohorts. Both approaches yield consistent estimation results, as shown in Appendix H.3.

Finally, Schwank (2024) applies stability tests following Altonji, Elder, and Taber (2005) and Oster (2019) to assess sensitivity to unobserved selection. Our results suggest that selection on unobservables is unlikely to account for our main findings. We further test the relationship between baseline academic ability and age at move and do not find any systematic patterns. These results are more thoroughly discussed in H.4.

Table H.1: Validity Tests used in Recent Mover Design Studies

Study	**Validity Tests Used**
Chetty and Hendren (2018)	Overidentification, Family Fixed Effects, Displacement Shocks
Bell et al. (2019)	Overidentification
Deutscher (2020)	Overidentification, Family Fixed Effects
Alesina et al. (2021)	Family Fixed Effects, Displacement Shocks, Shift-Share IV
Finkelstein, Gentzkow, and Williams (2021)	Overidentification
Laliberté (2021)	Overidentification, Family Fixed Effects
Maarseveen (2021)	Family Fixed Effects, Displacement Shocks
Derenoncourt (2022)	Displacement Shocks, Shift-Share IV
Schwank (2024)	Overidentification, Family Fixed Effects, Stability Tests

H.1 Overidentification Tests: Group and College Major Specific Convergence

To test for bias from unobservable factors, we conduct placebo tests that exploit variation in average STEM rates across different subgroups, specifically, by cohort and gender. The causal exposure effect model predicts that a student's major choice should converge to the average STEM rate of non-moving students within their own subgroup. In contrast, models driven by omitted variables or selection would not generate such subgroup-specific convergence, under reasonable assumptions about parental information and preferences. This variation in STEM rates across subgroups of non-movers provides a set of overidentifying restrictions. Leveraging these restrictions allows us to more effectively assess whether neighborhoods have a causal impact on students' STEM choices, or whether the observed patterns might instead reflect unobserved confounding factors.

Intuitively, it is improbable that a correlated shock would coincide precisely with cohort (or gender)-level differences in place effects, as reflected in the STEM rates of non-moving students. Formally, this test is based on the assumption that if unobserved factors θ_i are correlated with exposure to the place effect for a given cohort $s(i)$ (proxied by the STEM rates of non-moving students), they should also be correlated with exposure to the place effects of adjacent cohorts s':

$$Cov(\theta_i, m\Delta STEM_{odg,s(i)}|X) > 0 \Rightarrow Cov(\theta_i, m\Delta STEM_{odg,s'}|X, m\Delta STEM_{odg,s(i)}) > 0$$

where X represents the vector of fixed effects and other controls in equation (3).

We begin with variation in place effects across cohorts. Although non-moving students' STEM choices are generally stable over time, those in some areas have improved over time, while others have gotten worse. Such changes could occur, for instance, because of changes in the local schools or other area-level characteristics that affect students' STEM choices. We exploit this heterogeneity across cohorts to test for confounds in the baseline research design. In the context of the causal exposure effect model, when a student moves to a new destination d, the change in STEM rates $\Delta STEM_{odg,s(i)}$ within that student's own cohort $s(i)$ should have a stronger influence on their STEM choice than the change in STEM rates $\Delta STEM_{odgs}$ for other cohorts $s \neq s(i)$. The rationale is that a student's decision to pursue STEM is more likely to be influenced by the STEM rates within their own cohort in the neighborhood, rather than the rates for older or younger cohorts. Conversely, unobserved factors θ_i are unlikely to vary significantly across different cohorts s in correlation with $\Delta STEM_{odgs}$, since these cohort-specific variations emerge only later in college and are not apparent at the time of the move. Thus, by comparing the predictive power of STEM rates from a student's own cohort to those from surrounding cohorts, we can evaluate the potential bias introduced by unobserved factors.

To conduct this analysis, we estimate the baseline specification in (3) by substituting the change in STEM rates for non-moving students within the student's own cohort, $\Delta STEM_{odg,s(i)}$, with corresponding STEM rates for neighboring school cohorts $s(i)+t$, $\Delta STEM_{odg,s(i)+t}$.[H.1] The results, presented in Table H.2 Panel A, show the exposure effect estimates (γ_t) from these regressions. The estimates for adjacent younger and older cohorts are similar to the baseline estimate of $\gamma = 0.032$, reflecting the strong correlation in STEM rates among non-moving students. In column (4), where predictions for both younger and older are included together, the

[H.1] We have 10 school cohorts from 1994-2003. Due to the size of the sample, we divide the cohorts into three groups: 1994-1997, 1998-2000, and 2001-2003 cohorts. If we cannot observe $\Delta STEM_{odgt}$, we set the younger or older exposure variables to zero and we include an indicator, $I_a = I\{\text{cohort s}(i) \text{ + a is missing}\}$ for each. For instance, students in the 1994-1997 cohort have their own cohort exposure effects and younger cohort exposure effects but older cohort exposure effects.

coefficients $\tilde{\gamma}_t$ for placebo exposure effects (γ_t for $t \neq 0$) are smaller and statistically insignificant. Notably, the estimate for younger cohorts becomes positive, indicating the unlikely impact on older cohorts' college major choices. However, the exposure effect estimate for the student's own cohort remains stable at 0.031.

Next, we perform another placebo test examining differences by student gender. We start by creating gender-specific predictions for the mean STEM choice rate of non-moving students, denoted as $\overline{STEM}^D_{pg}$. Here, the superscript D represents different groups: male (m) and female (f). Locations with high STEM rates for boys also tend to be high for girls, with a population-weighted correlation of 0.86 between $\overline{STEM}^m_{pg}$ and $\overline{STEM}^f_{pg}$ across counties. We use the residual variation between genders to conduct placebo tests, assuming that unobserved shocks are unlikely to have gender-specific effects.

In Table H.2 Panel B, we estimate exposure effect models similar to equation (3), but with predictions divided by gender. Column (1) replicates Table 2 column (1) using gender-specific predictions $\Delta STEM^D_{odg}$, rather than a combined prediction for both genders. This yields an exposure effect estimate of $\gamma = 0.032$ per year. In column (2), we use the prediction for the other gender, $\Delta STEM^{-D}_{odg}$, which results in an estimate of 0.028, consistent with the high correlation between outcomes for different genders. In column (3), we include predictions for both genders simultaneously, isolating the exposure effects based on differential variation across genders within counties. Here, the coefficient for the own-gender prediction is $\gamma = 0.027$, which is 3.4 times larger than the coefficient for the other-gender prediction.

To assess whether the estimated exposure effects are truly STEM-specific or simply reflect broader improvements in educational outcomes, we conduct an overidentification test using exposure measures based on changes in non-moving students' major choice rates across different fields of study. Specifically, we incorporate an alternative exposure measure for liberal arts majors, denoted as $\Delta LIBARTS_{odg}$, into our baseline model alongside the STEM exposure measure.

We define liberal arts majors broadly using the 2010 two-digit CIP codes, selecting fields with academic orientations of similar size to STEM majors. For example, among our baseline sample of students who moved once, the STEM major rate is 3.9 percent, while the liberal arts major rate is 4.2 percent. These include education (13), foreign languages, literatures, and linguistics (16), English language and literature/letters (23), liberal arts and sciences, general studies, and humanities (24), philosophy and religious studies (38), theology and religious vocations (39), social sciences (45), communication, journalism, and related programs (09), and history (54). This grouping is designed to capture a distinct but comparably sized set of majors to test whether neighborhood exposure effects are specific to STEM fields.

We estimate the baseline specification described in equation (3), now including both the STEM and liberal arts exposure measures as regressors. Panel A of Table H.3 reports the resulting estimates of the exposure effects (γ). In Column (1), the coefficient on the STEM exposure measure remains similar in magnitude to the baseline estimate, even after controlling for liberal arts exposure. In contrast, the coefficient on liberal arts exposure ($\gamma_{LIBARTS}$) is small and statistically insignificant, suggesting that STEM exposure predicts STEM major choice independently of liberal arts trends. In Column (2), we further explore the patterns by estimating the likelihood of choosing a liberal arts major as the outcome. Here, we again find evidence of major-specific convergence: students exposed to areas with higher liberal arts major rates are more likely to choose liberal arts majors, while exposure to higher STEM rates decreases the probability of selecting liberal arts.

Panel B extends this analysis by disaggregating STEM into two common subfields: engineering and biological sciences. Again, we observe strong major-specific convergence. Students exposed to areas with higher concentrations of engineering or biology majors are more likely to pursue those specific fields.

Together, these results support the interpretation that our estimates reflect genuine exposure to field-specific cultural or institutional environments, rather than a general improvement in educational attainment. This analysis helps to rule out the concern that students are simply responding to high-achieving environments regardless of major composition, strengthening the case for a STEM-specific exposure effect.

Table H.2: Exposure Effect Estimates: Group Specific Convergence

Panel A: Cohort

	(1)	(2)	(3)	(4)
Exposure effects-own cohort	0.032***			0.031***
	(0.005)			(0.007)
Exposure effects-older cohort		0.037***		0.007
		(0.007)		(0.009)
Exposure effects-younger cohort			0.016***	-0.011**
			(0.005)	(0.005)
Observations	205,777	205,777	205,777	205,777

Panel B: Gender

	(1)	(2)	(3)
Exposure effects-own gender	0.032***		0.027***
	(0.005)		(0.007)
Exposure effects-other gender		0.028***	0.008
		(0.004)	(0.007)
Observations	217,210	217,210	217,210

Notes: The table reports estimates of annual exposure effects γ using group-specific STEM attainment rates among non-movers. The sample consists of all one-time movers, defined in the notes to Table 1. *Panel A* shows cohort-specific convergence. Column (1) replicates column (1) of Table 2, replacing the average STEM rates in the origin and destination with the average STEM rates of students in the same cohort as the movers. See Appendix H.1 for more discussion on the test and construction of school cohorts. Columns (2) and (3) replicate column (1), with the STEM rates of older and younger cohorts, respectively. Column (4) combines the variables in columns (1), (2), and (3). *Panel B* shows gender-specific convergence. Column (1) replicates column (1) of Table 2, using STEM rates of non-moving students who have the same gender as the student who moves instead. Column (2) replicates column (1), replacing the own-gender STEM rate with the STEM rate of the opposite gender. Column (3) combines the variables in columns (1) and (2). Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

Table H.3: Exposure Effect Estimates: College Major Specific Convergence

Panel A: College Major		
	(1) STEM	(2) Liberal Arts
Exposure effects-STEM	0.027**	-0.028***
	(0.011)	(0.008)
Exposure effects-Liberal Arts	0.005	0.055***
	(0.011)	(0.009)
Observations	205,777	205,777

Panel B: Narrow College Major			
	(1) Engineering	(2) Biological Science	(3) Liberal Arts
Exposure effects-Engineering	0.030***	-0.012*	-0.052**
	(0.010)	(0.007)	(0.023)
Exposure effects-Biological Science	0.038	0.035***	0.005
	(0.026)	(0.009)	(0.044)
Exposure effects-Liberal Arts	-0.012*	0.003	0.051***
	(0.007)	(0.003)	(0.008)
Observations	205,777	205,777	205,777

Notes: The table reports estimates of annual exposure effects γ using major-specific STEM attainment rates among non-movers. The sample consists of all one-time movers, defined in the notes to Table 1. *Panel A* reports alternative exposure measures based on different college majors. Column (1) replicates column (1) of Table 2, adding another different exposure measures based on other college majors—liberal arts majors. The measure is constructed in the same way as STEM exposure measures ($\Delta libarts_{odg}$ = $\overline{libarts}_{dg}$ - $\overline{libarts}_{og}$). Column (2) replicates column (1), replacing the dependent variable with liberal arts major. *Panel B* Column (1) replicates column (1) of Table 2, and has exposure measures for engineering, biological science, and liberal arts. Columns (2) and (3) replicate column (1), replacing the dependent variable with biological science and liberal arts majors, respectively. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

H.2 Same School Catchment Area Comparison

To improve the comparability of students, we restrict the sample to a smaller group: students of the same race and economic status who move from the same school catchment area to the same destination catchment area in the same year. This is implemented by adding origin catchment area-by-destination catchment area-by-race and economic status-by-year of move fixed effects to the baseline fixed effects. Catchment areas are defined based on student flows: if more than 30% of students move from one school to another in the terminal grade, the two schools are considered part of the same catchment area. Based on this definition, we identify 1,180 catchment areas in Texas, comprising an average of 6 schools (with a median of 3 schools).

Columns (4) and (8) of Table 2 present results from this specification, applied to both the parametric and semi-parametric baseline models, respectively. The estimated exposure effects remain consistent and are slightly larger than the baseline estimates. However, this approach results in a roughly 60% reduction in sample size due to the catchment area construction process and the large number of fixed effects. As a result, approximately 81,000 singleton observations are dropped. These singletons are observations that are the only member of their fixed effect group and must be excluded because they offer no within-group variation. For our estimation, there are about 26,800 fixed effects, with an average of 3.2 students per comparison group. To address concerns about sample differences, Appendix Table H.4 presents results using the same sample across all specifications. We find that the estimates remain consistent after restricting to the common sample. Moreover, the estimation results remain consistent across different cutoff values used to define catchment areas, as shown in Appendix Table H.5.

This analysis helps alleviate concerns about comparing students from different family backgrounds by narrowing the sample to students likely to come from highly similar contexts—those of the same race and economic status, who move from the same place to the same place in the same year. On average, when we include the singletons, there are 1.6 students per group, and given that the average household with children includes about 2 children under age 18 (Census 2004), this effectively compares outcomes across one family. This strategy also partially addresses the lack of family identifiers in our data. Notably, prior studies using similar approaches report consistent results whether controlling for family fixed effects (Alesina et al. 2021; Chetty and Hendren 2018; Deutscher 2020; Laliberté 2021; Maarseveen 2021; Schwank 2024).

While comparisons at the catchment area level help address bias from fixed family characteristics—such as stable differences in parental background or household environment—they do not fully account for time-varying family inputs. For example, a family may move to a high-STEM area due to a parental job promotion, which itself could directly benefit the child's outcomes. Particularly if the move occurs earlier and the child is exposed longer to improved family circumstances at the destination, it biases our exposure effect estimates.

To help mitigate this concern, we restrict our comparisons to students whose families have the same economic status both in grade 0 (baseline) and at the time of the move. In other words, we modify the fixed effects term in equation (1) from α_{osgm} to $\alpha_{osg_1g_2m}$, where g_1 refers to economic status in grade 0 and g_2 refers to economic status at the time of the move along with race.

Appendix Table H.6 presents estimation results analogous to those in Table 2. The estimated effect is slightly larger, with $\gamma = 0.034$, compared to the baseline estimate of $\gamma = 0.030$. However, the results remain broadly consistent with the baseline findings, suggesting that time-varying family inputs are unlikely to explain the observed effects.

Table H.4: Exposure Effects in STEM Major Choice: Using the Same Sample Across Different Specifications

	Parametric				Semi-parametric			
	(1) Baseline	(2) + Destination info	(3) + Test scores & Labor char.	(4) Catchment FE	(5) Baseline	(6) + Destination info	(7) + Test scores & Orig.×dest. FE	(8) Catchment FE
Exposure effects	0.033***	0.032***	0.033***	0.041**	0.029***	0.030***	0.031***	0.035
	(0.009)	(0.009)	(0.008)	(0.018)	(0.010)	(0.010)	(0.011)	(0.024)
Observations	64,899	64,899	64,899	64,899	64,899	64,899	64,899	64,899

Notes: This table reports estimates of annual exposure effects on students' STEM degree attainment by age 26 with the common sample across different specifications. The estimates can be interpreted as the impact of spending an additional school year in a county where non-moving students have a 1% higher STEM attainment rate. Each column reports estimates from a regression of a student's STEM degree attainment by 26 on the difference between non-moving students' STEM rates in the destination vs. the origin, interacted with the grade of the student at the time of the move (m). Column (1) reports the estimate of γ and its standard error from equation (3), using the sample of one-time movers defined in the notes to Table 1. Column (2) adds to column (1) fixed effects for race-by-economic status interacted with the destination STEM rate of non-movers from the same race and economic status group. Column (3) further includes labor market characteristics (median income, employment rate, and poverty rate in origin and destination counties, constructed using the same method described in Section 5) and moving students' third-grade math and reading scores. Column (4) adds to column (1) finer fixed effects: race-by-economic status-by-origin school catchment area-by-destination catchment area-by-year of move. Columns (1) through (4) all include race-by-economic status-by-cohort-by-grade-of-move fixed effects, interacted with the origin STEM rate of non-movers from the same race and economic status group. Columns (5) through (8) replicate the specifications in Columns (1) through (4), respectively, but replace the origin STEM rate of non-movers—which was interacted with race-by-economic status-by-cohort-by-grade-of-move fixed effects—with origin county fixed effects. Additionally, Column (7) includes origin-by-destination county fixed effects by race and economic status instead of labor market controls. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

Table H.5: Exposure Effects in STEM Major Choice: Across Different Definitions of Catchment Areas

	(1) > 20%	(2) > 25%	(3) > 30%	(4) > 35%	(5) > 40%
Exposure effects	0.030***	0.036***	0.033**	0.038**	0.042**
	(0.008)	(0.011)	(0.014)	(0.016)	(0.018)
Observations	139,863	117,261	87,000	71,140	62,732

Notes: The table reports estimates of annual exposure effects on students' STEM degree attainment by age 26, using school catchment area fixed effects across different definitions of catchment areas. We define schools as belonging to the same catchment area if more than the percentage indicated in the column header of students move from one school to the other at the terminal grade. The estimates can be interpreted as the impact of spending an additional school year in a county where non-moving students have a one percentage point higher STEM attainment rate. Each column reports estimates from a regression of a student's STEM degree attainment by 26 on the difference between non-moving students' STEM rates in the destination vs. the origin, interacted with the grade of the student at the time of the move (m). Each regression also includes additional controls specified in equation (3). Observations are dropped due to fixed effects and cutoff definition. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

Table H.6: Exposure Effects in STEM Major Choice: Time-Varying Family Inputs

	Parametric				Semi-parametric			
	(1) Baseline	(2) + Destination info	(3) + Test scores & Labor char.	(4) Catchment FE	(5) Baseline	(6) + Destination info	(7) + Test scores & Orig.×dest. FE	(8) Catchment FE
Exposure effects	-0.034***	-0.035***	-0.037***	-0.038**	-0.033***	-0.024**	-0.024*	-0.071**
	(0.010)	(0.009)	(0.009)	(0.019)	(0.011)	(0.011)	(0.013)	(0.030)
Observations	188745	188745	169505	72998	174049	174011	147943	64200

Notes: This table reports estimates of annual exposure effects on students' STEM degree attainment by age 26. The coefficients can be interpreted as the impact of spending an additional school year in a county where non-moving students have a one percentage point higher STEM attainment rate. Each column presents results from a regression of students' STEM degree attainment on the difference in STEM rates between destination and origin counties for non-moving students of the same race and economic status group, interacted with the grade at which the student moved (m). Column (1) reports the estimate of γ and its standard error from equation (3), using the sample of one-time movers defined in the notes to Table 1. Column (2) adds to column (1) fixed effects for race-by-economic status interacted with the destination STEM rate of non-movers from the same race and economic status group. Column (3) further includes labor market characteristics (median income, employment rate, and poverty rate in origin and destination counties, constructed using the same method described in Section 5) and moving students' third-grade math and reading scores. Column (4) adds to column (1) finer fixed effects: race-by-economic status at grade 0-by-economic status in the year of move-by-origin school catchment area-by-destination catchment area-by-year of move. Columns (1) through (4) all include race-by-economic status at grade 0-by-economic status in the year of move-by-cohort-by-grade-of-move fixed effects, interacted with the origin STEM rate of non-movers from the same race and economic status group. Columns (5) through (8) replicate the specifications in Columns (1) through (4), respectively, but replace the origin STEM rate of non-movers—which was interacted with race-by-economic status at grade 0-by-economic status in the year of move-by-cohort-by-grade-of-move fixed effects—with origin county fixed effects. Additionally, Column (7) includes origin-by-destination county fixed effects by race and economic status instead of labor market controls. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

H.3 Displacement Shocks and Expected Destination of Movers

One approach to address concerns about unobserved confounding factors is to focus on a subset of moves where the underlying cause is plausibly exogenous. Specifically, we examine families who were plausibly forced to relocate from an origin county o to a nearby destination d due to displacement shocks such as natural disasters or plant closures. These shocks may result in varying changes in neighborhood STEM exposure due to relocation of families ($\Delta STEM_{od}$). In this setting, estimates of the exposure effect γ based on displaced students are less likely to be biased by selection into the timing of move, allowing for a more credible assessment of whether changes in students' STEM choices are attributable to their new environments.

To operationalize this approach, we identify displacement shocks using patterns of population outflows at the county level. Let K_{pt} denote the number of movers leaving county p in year t, and $\bar{K}_p$ represent the average annual outflow from county p over a multi-year period. We define the relative outflow as $k_{pt} = K_{pt}/\bar{K}_p$. This measure captures unusually large outflows, which often coincide with disruptive events. We validate this shock measure by examining historical natural disasters in Texas. For instance, the outflow from Orange County in 2006, following Hurricane Rita in 2005, ranks at the 95th percentile of k_{pt} values. Similarly, the 2008 outflow from Galveston County, after Hurricane Ike, ranks at the 93rd percentile. These examples confirm that our outflow-based measure effectively captures large-scale displacement events.

Figure H.1 presents the results from this analysis. To construct the figure, we sort observations into population-weighted bins based on the size of the relative outflow k_{dt}. The first point displays the estimated annual exposure effect γ using the full sample. Each subsequent point represents the estimate of γ for increasingly restrictive subsamples. The second point shows the estimate γ using observations where k_{dt} is at or above the 10th percentile. Subsequent points are calculated similarly, increasing the threshold by 10 percentiles each time, with the final point showing an estimate of γ based only on the county in the top two percentiles of outflow rates. The dotted lines represent the 95% confidence intervals for these estimates.

If the initial exposure effect estimates were primarily driven by selection bias—i.e., by families strategically moving to higher-STEM environments in earlier grades—then we would expect the estimated γ to disappear when the analysis is restricted to students whose moves were driven by potentially exogenous shocks. However, the results indicate that the estimated exposure effect remains stable, with $\gamma \approx 0.03$ even in the most restricted samples. In other words, students who moved due to significant displacement events and were exposed to higher-STEM neighborhoods in their formative years continue to show a higher likelihood of pursuing STEM majors as adults.

These findings support the causal interpretation of the baseline estimates. They suggest that the observed relationship between neighborhood STEM exposure and students' later outcomes is not solely the result of family selection into different environments in earlier grades. Instead, the exposure effect appears to reflect an influence of place on educational choices. Moreover, this analysis reinforces the generalizability of our findings, showing that even for families who did not voluntarily select into high-STEM areas, exposure to such environments had lasting impacts on students' academic trajectories.

While restricting the sample to displaced movers helps mitigate concerns about endogenous timing of the move, questions may remain about the endogenous choice of where families move. To address this, we follow Alesina et al. (2021) and implement a shift-share IV strategy to instrument for the actual destination. Because our data do not allow the construction of shift-share instruments from earlier cohorts, we use later cohorts (students starting kindergarten in 2007-2009) to predict destination patterns. Specifically, for each origin county o and group

g, we calculate the share of moves to each destination d based on these future cohorts. These shares reflect systematic migration patterns rather than individual sorting.

We use these shares as weights to construct the predicted change in STEM exposure after a move for each origin-by-group combination:

$$\Delta\widehat{STEM}^{w}_{og} = \sum_{d} \omega_{d(og)} \Delta STEM_{odg},$$

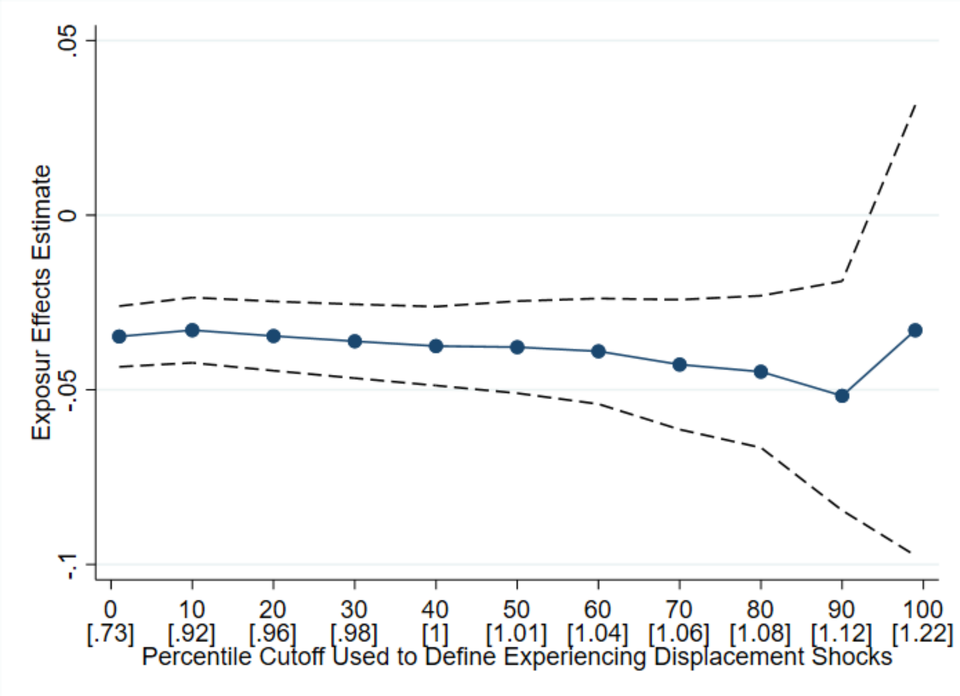
where $\omega_{d(og)}$ is the migration share to destination d for each origin o and group g in the future cohort.

Figure H.2 shows a binned scatterplot of the actual vs. predicted STEM exposure changes after move. Each dot represents the mean of 100 equally sized bins sorted by the predicted value. The slope is close to one, suggesting that our predicted STEM changes based on migration patterns from future cohorts closely approximate average actual exposure changes.

We then estimate the main regression using the predicted changes $\Delta\widehat{STEM}^{w}_{og}$ as an instrument for the actual exposure change $\Delta STEM_{odg}$. Table H.7 presents these IV estimation results alongside the reduced form estimation results. Columns (1) and (2) show reduced form estimates, while columns (3) and (4) show the IV results. The estimates are slightly smaller in magnitude than the baseline results reported in Table 2, but remain consistent in direction and significance, supporting the robustness of our identification strategy to the choice of destination.

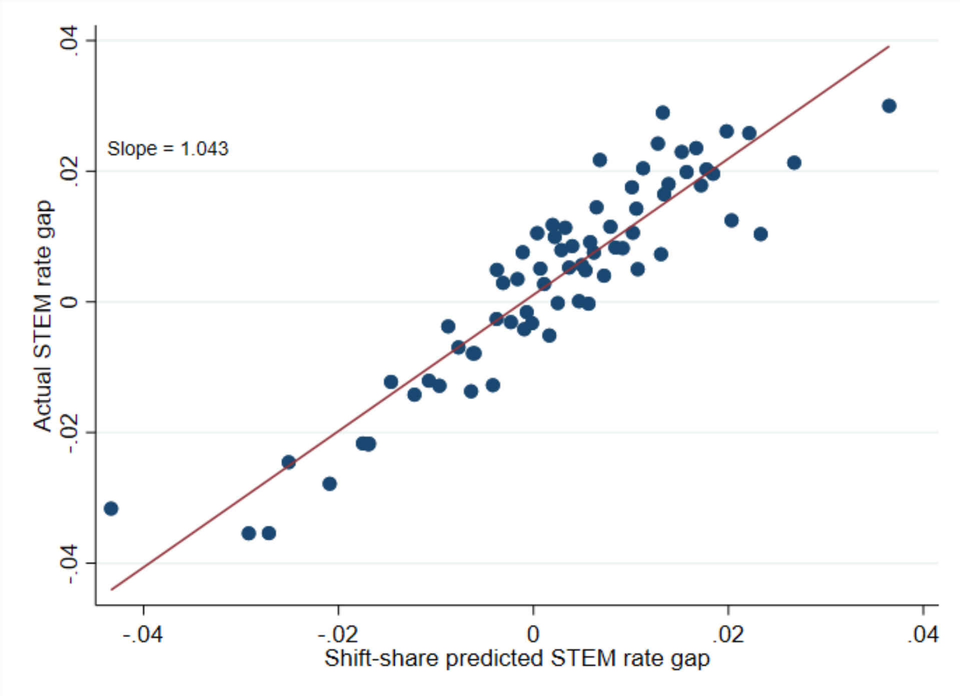
Finally, we examine whether these results hold even when we combine the two approaches, using the predicted exposure changes based on future cohort migration patterns while restricting the sample to students more likely to be displaced by shocks. Due to smaller sample sizes in these cases, we estimate reduced form regressions only. As shown in Appendix Figure H.3, the estimated effects remain consistent across subsamples (e.g., $\gamma_{k_{dt}>99} = 0.27$), despite some expected noise. Together, these findings alleviate concerns that endogenous destination choices, along with the timing of the move, are a major source of bias in our exposure effect estimates and lend further support to a causal interpretation.

Fig. H.1. Exposure Effects Estimates Using Displacement Shocks



Notes: The figure presents estimates of annual exposure effects for areas experiencing displacement shocks, defined as county-by-year cells with large outflows of students. We measure outflows by defining K_{pt} as the number of students who leave county p in year t in our one-time movers sample, and $\bar{K}_p$ as the average outflow over the years. The outflow in year t for county p is then defined as $k_{pt} = K_{pt}/\bar{K}_p$. The sample is divided into 10 population-weighted bins based on the size of the relative outflow k_{pt}. For each subset of observations where k_{pt} exceeds the percentile threshold listed on the x-axis, we estimate γ using equation (3). The dashed lines represent 95% confidence intervals for the estimates. The mean value of the relative outflow k_{pt} used in each subsample is shown in brackets below the percentile thresholds.

Fig. H.2. Predicted and Actual STEM Rate Gap



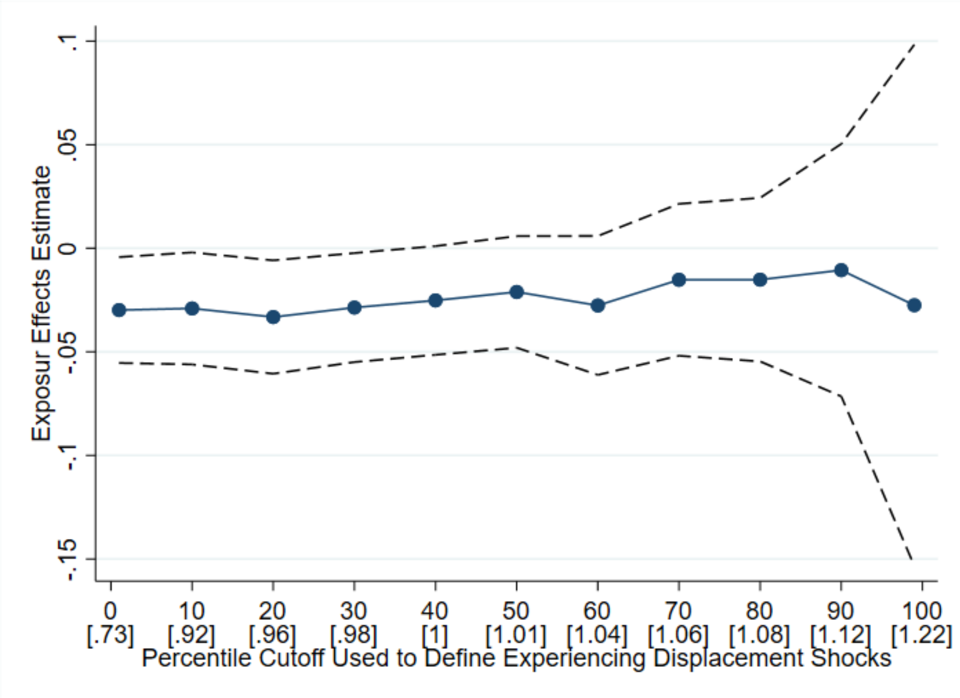
Notes: This figure displays a binned scatterplot of the actual origin-destination STEM rate gap, $\Delta STEM_{odg}$, on the vertical axis against the predicted STEM rate gap, $\Delta\widehat{STEM}^{w}_{og}$ on the horizontal axis. The binned scatterplot is created by dividing the individual-level data into 100 equally sized bins of the predicted STEM rate gap. Each dot in the figure represents the within-bin means of the actual and predicted STEM rate gap. The line in the figure is the linear fit through the binned means. The slope quantifies the strength of the shift-share prediction in explaining observed origin-destination STEM rate gaps.

Table H.7: Exposure Effects in STEM Major Choice: Shift-Share Instrument

	(1) Baseline (RF)	(2) + Test scores & Labor char. (RF)	(3) Baseline (SS)	(4) + Test scores & Labor char. (SS)
Exposure effects	-0.025*	-0.024*	-0.022	-0.024*
	(0.014)	(0.013)	(0.027)	(0.014)
Observations	205,777	182,830	205,777	182,830

Notes: This table reports estimates of annual exposure effects on students' STEM degree attainment by age 26. The coefficients can be interpreted as the impact of spending an additional school year in a county where non-moving students have a one percentage point higher STEM attainment rate. Each column presents results from a regression of students' STEM degree attainment on the difference in STEM rates between destination and origin counties for non-moving students of the same race and economic status group, interacted with the grade at which the student moved (*m*). Column (1) reports the estimate of γ and its standard error from equation (3), while we replace $\Delta STEM_{odg}$ to $\Delta\widehat{STEM}^{w}_{og} = \sum_d \omega_{odg}\Delta STEM_{odg}$ where ω_{odg} is the migration share from origin o to destination d for each group g in the future cohort. The table uses the sample of one-time movers defined in the notes to Table 1. Column (2) further includes labor market characteristics (median income, employment rate, and poverty rate in origin and destination counties, constructed using the same method described in Section 5) and moving students' third-grade math and reading scores. Columns (1) and (2) present the reduced form estimates. Columns (3) and (4) present the 2SLS estimates for the same specifications as Columns (1) and (2) where the actual $\Delta STEM_{odg}$ is instrumented with the predicted $\Delta\widehat{STEM}^{w}_{og}$. Standard errors are shown in parentheses and clustered at the origin county level. *** p<0.01, ** p<0.05, * p<0.10

Fig. H.3. Exposure Effects Estimates Using Displacement Shocks with Predicted STEM Rate Changes



Notes: The figure presents estimates of annual exposure effects for areas experiencing displacement shocks, defined as county-by-year cells with large outflows of students. We measure outflows by defining K_{pt} as the number of students who leave county p in year t in our one-time movers sample, and $\bar{K}_p$ as the average outflow over the years. The outflow in year t for county p is then defined as $k_{pt} = K_{pt}/\bar{K}_p$. The sample is divided into 10 population-weighted bins based on the size of the relative outflow k_{pt}. For each subset of observations where k_{pt} exceeds the percentile threshold listed on the x-axis, we estimate γ using equation (3) while we replace $\Delta STEM_{odg}$ to $\Delta\widehat{STEM}^{w}_{og} = \sum_d \omega_{odg}\Delta STEM_{odg}$ where ω_{odg} is the migration share from origin o to destination d for each group g in the future cohort. The dashed lines represent 95% confidence intervals for the estimates. The mean value of the relative outflow k_{pt} used in each subsample is shown in brackets below the percentile thresholds.

H.4 Stability Test of Coefficients with Respect to Unobservables

A key identifying assumption of our research design is that the timing of a family's move is not systematically correlated with unobserved characteristics that also affect students' decisions to pursue STEM majors. This assumption would be violated if families' unobserved traits were related to the grade at which they choose to move, introducing selection bias. For example, if more affluent families or those who are particularly invested in their children's education tend to move during earlier grades, these early movers could exhibit stronger selection effects, leading to an overestimation of the exposure effects.

To evaluate the robustness of our results to concerns about omitted variable bias, we follow the methodology proposed by Altonji, Elder, and Taber (2005) and formalized by Oster (2019), which assesses the stability of coefficient estimates when observed controls are added to the model. The underlying intuition is that if including observed controls substantially increases the explained variance (R^2) while causing only a modest change in the coefficient estimate, then it would require an implausibly high degree of selection on unobservables to fully account for the estimated effect.

We use our baseline model (equation 3) with controls for students' third-grade math and reading scores, origin and destination undergraduate degree attainment rates, and fixed effects for race-by-economic status-by-school cohort-by-grade at the time of the move, interacted with predicted neighborhood STEM rates. In contrast, our restricted model drops characteristics linked to family and student factors that could reflect unobserved characteristics influencing families' moving decisions. Specifically, we exclude fixed effects for race and economic status as well as test scores, replacing them with fixed effects only for school cohort-by-grade at the time of the move, interacted with predicted STEM rates. Since race, economic status, and early test scores are variables likely correlated with unobservable family and student characteristics, omitting them allows us to estimate the extent of potential bias from selection on unobservables.

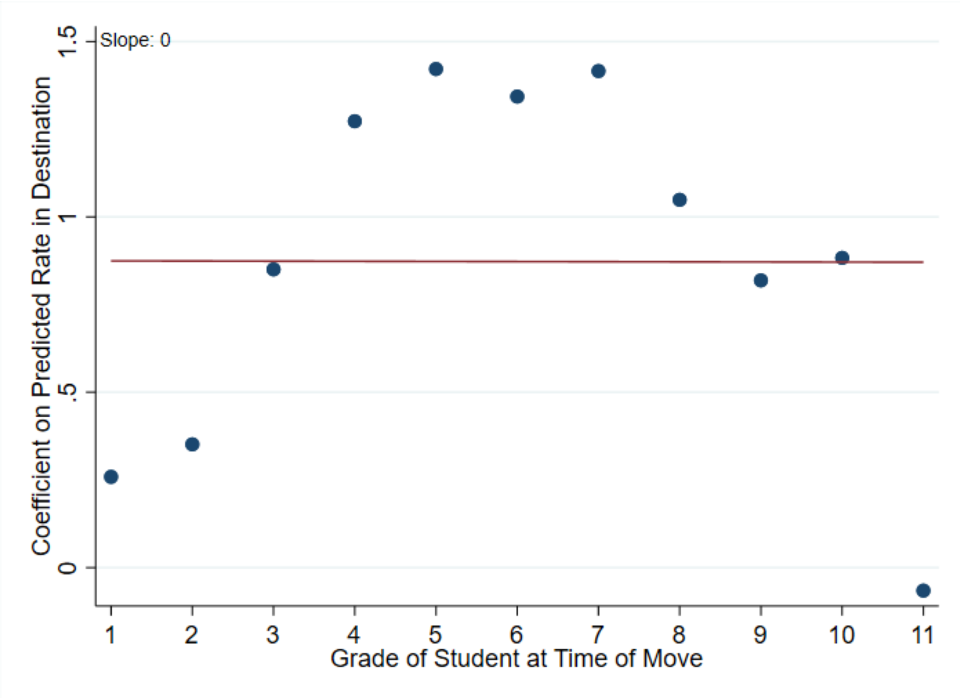
Following Oster (2019), we calculate the robustness parameter δ^{Oster}, which captures the relative degree of selection on unobservables needed to drive our estimated coefficient, γ, to zero. Oster (2019) suggests a value of $\delta^{Oster} \geq 1$ as the cutoff for robust results. Anything greater than 1 suggests that an exceptionally high degree of selection on unobservables is required for the true effect to be zero, indicating the results are unlikely to be driven by omitted variable bias. For our calculation, we assume the maximum possible R^2 (R^{max}) is 1.3 times the R^2 from our full model, consistent with the recommendation in Oster (2019).

In the restricted model, the coefficient estimate is -0.033 with an R^2 of 0.017, whereas in the full model, the coefficient shifts slightly to -0.031 with an R^2 of 0.029. This small change in the coefficient, combined with the substantial increase in R^2, indicates that our estimates are relatively stable. Calculating the robustness parameter yields $\delta^{Oster} = 6.05$, suggesting that selection on unobservables would need to be roughly six times stronger than selection on observables to drive our estimated effect to zero, which we believe is implausible. This provides strong evidence that our main findings are unlikely to be driven by omitted variable bias.

Furthermore, we directly assess whether students' baseline academic ability is systematically related to the grade at which they move, using students' standardized math scores in grade 3 as the outcome variable. These scores serve as a proxy for students' academic ability and are assumed to be unaffected by future neighborhood exposure, as most moves occur after this baseline assessment and test scores do not usually change significantly in a few years. If families with higher- or lower-ability children tend to move at specific points in time, this could bias our estimates of neighborhood exposure effects. We estimate our baseline specification, as described in equation (3), but use grade 3 math scores as the dependent variable instead of STEM major choice.

Appendix Figure H.4 displays these estimated coefficients across grades. If families strategically timed their moves based on their children's academic potential—for example, moving earlier for higher-ability students—we would expect to observe the same downward trends in these coefficients. In contrast, the estimates do not show clear patterns. This result provides reassuring evidence in support of our identifying assumption: the timing of moves does not appear to be systematically related to students' unobserved ability. In other words, unobservable traits linked to early academic performance do not predict the grade at which families choose to relocate. This finding strengthens the credibility of our empirical strategy, suggesting that variation in exposure driven by the grade of the move can be interpreted as plausibly exogenous with respect to baseline student ability.

Fig. H.4. Placebo Test using Math score in Grade 3



Notes: This figure plots estimates of the coefficients b_m across the students' grade when they move (m) using the same specification of Figure 4 (b), replacing the dependent variable with third-grade math score. The sample includes all students in kindergarten in 1994-2003 who moved only once in their school years. The b_m coefficients can be interpreted as the effect of moving to an area where non-moving students' STEM rate increases from 0 to 1 at age m. They are estimated by regressing the students' third-grade math score on $\Delta_{odg} = \overline{STEM}_{dg}$ - $\overline{STEM}_{og}$, the difference in predicted STEM attainment between the destination and origin counties, based on non-moving students from the same racial and economic status group, interacted with each grade of the student at the time of the move m. We include race-by-economic status-by-school cohort-by-grade at move fixed effects, interacted with predicted STEM rate for non-moving students from the same racial and economic status group in the origin.

I Details for Counterfactual Analysis

I.1 Using STEM Occupation Exposure Measure

In this section, we outline the methodology for our back-of-the-envelope counterfactual analysis. To quantify exposure, we calculate an annual exposure rate based on the proportion of STEM occupations in students' neighborhoods. Specifically, for all students (i) who began kindergarten between 1994 and 2003, we compute each student's average STEM occupation exposure as:

$$AvgSTEMOcc_i = \frac{1}{S_i} \sum_{s=1}^{S_i} STEMOcc_{p(s)g},$$

where S_i is the number of school years observed for student i, and $STEMOcc_{p(s)g}$ represents the STEM occupation rate for the neighborhood (county p) and group g in year s. We then calculate the group-level exposure by averaging $AvgSTEMOcc_i$ across all students in each group.

In our sample, 4.6 percent of White students major in STEM, compared to 1.4 percent of Black students and 2.4 percent of Hispanic students. To assess how differences in neighborhood exposure contribute to these gaps, we simulate a scenario in which underrepresented students are exposed to the same neighborhood STEM rates as White students. On average, White students are exposed to neighborhoods where 8.8 percent of neighbors are in STEM occupations, compared to 6.8 percent for Black students and 5.3 percent for Hispanic students. This corresponds to a 29 percent higher exposure rate for White students than Black students, and a 66 percent higher rate than for Hispanic students. If Black students were exposed to the same neighborhood STEM rates as White students for 13 years, their probability of majoring in STEM would increase by $(0.088-0.068) \times 13 \times 2.7 = 0.7$ percentage points, closing 22 percent of the Black–White gap. For Hispanic students, the increase would be $(0.088-0.053) \times 13 \times 2.0 = 0.91$ percentage points, explaining 41 percent of the Hispanic–White gap (see Appendix Table 3 Panel B for group-specific neighborhood exposure effects).

A similar pattern holds for economic status. Among economically non-disadvantaged students, 5.5 percent major in STEM, compared to 1.7 percent of economically disadvantaged students. The average neighborhood STEM exposure rate is 8.8 for non-disadvantaged students and 5.7 for disadvantaged students, which is a 54 percent difference. If disadvantaged students experienced the same exposure rate as their non-disadvantaged peers for 13 years, their STEM major rate would rise by $(0.088-0.057) \times 13 \times 1.7 = 0.69$ percentage points, closing 18 percent of the gap.

Although male and female students live in the same neighborhoods, they are influenced more by workers of the same gender, which results in different effective exposure levels. To explore this, we simulate a scenario in which female students are exposed to the same gender-specific exposure level as male students. In our data, 4.2 percent of male students major in STEM, compared to 2.6 percent of female students. Male students are exposed to neighborhoods where 8.0 percent of males are in STEM occupations, while the corresponding figure for females is 6.3 percent—a 28 percent difference. If female students experienced the same exposure level as male students for 13 years, their probability of majoring in STEM would increase by $(0.080-0.063) \times 13 \times 1.3 = 0.29$ percentage points, closing 11 percent of the gender gap.

I.2 Using Baseline STEM Exposure Measure

Similarly, we conduct a counterfactual analysis using our baseline STEM exposure measure—the proportion of STEM choices among non-moving students. In this case, rather than constructing the exposure at the county-by-group level, we use all students within each school to calculate the exposure rate. This approach better reflects the actual environment students experience during their school years.

In our sample, 4.6 percent of White students major in STEM, compared to 1.4 percent of Black students and 2.4 percent of Hispanic students. To assess how differences in neighborhood exposure contribute to these gaps, we simulate a scenario in which underrepresented students are exposed to the same neighborhood STEM rates as White students. On average, White students are exposed to neighborhoods where 5.2 percent of non-movers choose STEM majors, compared to 3.0 percent for Black students and 3.3 percent for Hispanic students. This corresponds to a 73 percent higher exposure rate for White students than Black students, and a 58 percent higher rate than for Hispanic students. If Black students were exposed to the same neighborhood STEM rates as White students for 13 years, their probability of majoring in STEM would increase by $(0.052-0.030) \times 13 \times 5.1 = 1.46$ percentage points, closing 46 percent of the Black–White gap. For Hispanic students, the increase would be $(0.052-0.033) \times 13 \times 2.4 = 0.59$ percentage points, explaining 27 percent of the Hispanic–White gap (see Appendix Table 3 Panel A for group-specific neighborhood exposure effects).

A similar pattern holds for economic status. Among economically non-disadvantaged students, 5.5 percent major in STEM, compared to 1.7 percent of economically disadvantaged students. The average neighborhood STEM exposure rate is 0.055 for non-disadvantaged students and 0.031 for disadvantaged students—a 77 percent difference. If disadvantaged students experienced the same exposure rate as their non-disadvantaged peers for 13 years, their STEM major rate would rise by $(0.055-0.031) \times 13 \times 1.7 = 0.53$ percentage points, closing 14 percent of the gap.

Although male and female students live in the same neighborhoods, they are influenced more by peers of the same gender, which results in different effective exposure levels. To explore this, we simulate a scenario in which female students are exposed to the same gender-specific exposure level as male students. In our data, 4.2 percent of male students major in STEM, compared to 2.6 percent of female students. Male students are exposed to neighborhoods where 5.3 percent of male students choose STEM majors, while the corresponding figure for female students is 3.2 percent—a 66 percent difference. If female students experienced the same exposure level as male students for 13 years, their probability of majoring in STEM would increase by $(0.053-0.032) \times 13 \times 1.8 = 0.49$ percentage points, closing 31 percent of the gender gap.

J List of STEM majors (DHS 2016)

CIP	Title	CIP	Title
1.1101	Plant Sciences, General.	15.1304	Civil Drafting and Civil Engineering CAD/CADD.
1.1102	Agronomy and Crop Science.	15.1305	Electrical/Electronics Drafting and Electrical/Electronics CAD/CADD.
1.1103	Horticultural Science.	15.1306	Mechanical Drafting and Mechanical Drafting CAD/CADD.
1.1104	Agricultural and Horticultural Plant Breeding.	15.1399	Drafting/Design Engineering Technologies/Technicians, Other.
1.1105	Plant Protection and Integrated Pest Management.	15.1401	Nuclear Engineering Technology/Technician.
1.1106	Range Science and Management.	15.1501	Engineering/Industrial Management.
1.1199	Plant Sciences, Other.	15.1502	Engineering Design.
1.1201	Soil Science and Agronomy, General.	15.1503	Packaging Science.
1.1202	Soil Chemistry and Physics.	15.1599	Engineering-Related Fields, Other.
1.1203	Soil Microbiology.	15.1601	Nanotechnology.
1.1299	Soil Sciences, Other.	15.9999	Engineering Technologies and Engineering-Related Fields, Other.
3.0101	Natural Resources/Conservation, General.	26	Biological and Biomedical Sciences.
3.0103	Environmental Studies.	27	Mathematics and Statistics.
3.0104	Environmental Science.	28.0501	Air Science/Airpower Studies.
3.0199	Natural Resources Conservation and Research, Other.	28.0502	Air and Space Operational Art and Science.
3.0205	Water, Wetlands, and Marine Resources Management.	28.0505	Naval Science and Operational Studies.
3.0502	Forest Sciences and Biology.	29.0201	Intelligence, General.
3.0508	Urban Forestry.	29.0202	Strategic Intelligence.
3.0509	Wood Science and Wood Products/Pulp and Paper Technology.	29.0203	Signal/Geospatial Intelligence.
3.0601	Wildlife, Fish and Wildlands Science and Management.	29.0204	Command & Control (C3, C4I) Systems and Operations.
4.0902	Architectural and Building Sciences/Technology.	29.0205	Information Operations/Joint Information Operations.
9.0702	Digital Communication and Media/Multimedia.	29.0206	Information/Psychological Warfare and Military Media Relations.
10.0304	Animation, Interactive Technology, Video Graphics and Special Effects.	29.0207	Cyber/Electronic Operations and Warfare.
11.0101	Computer and Information Sciences, General.	29.0299	Intelligence, Command Control and Information Operations, Other.
11.0102	Artificial Intelligence.	29.0301	Combat Systems Engineering.
11.0103	Information Technology.	29.0302	Directed Energy Systems.
11.0104	Informatics.	29.0303	Engineering Acoustics.
11.0199	Computer and Information Sciences, Other.	29.0304	Low-Observables and Stealth Technology.
11.0201	Computer Programming/Programmer, General.	29.0305	Space Systems Operations.
11.0202	Computer Programming, Specific Applications.	29.0306	Operational Oceanography.
11.0203	Computer Programming, Vendor/Product Certification.	29.0307	Undersea Warfare.
11.0299	Computer Programming, Other.	29.0399	Military Applied Sciences, Other.
11.0301	Data Processing and Data Processing Technology/Technician.	29.0401	Aerospace Ground Equipment Technology.
11.0401	Information Science/Studies.	29.0402	Air and Space Operations Technology.
11.0501	Computer Systems Analysis/Analyst.	29.0403	Aircraft Armament Systems Technology.
11.0701	Computer Science.	29.0404	Explosive Ordinance/Bomb Disposal.
11.0801	Web Page, Digital/Multimedia and Information Resources Design.	29.0405	Joint Command/Task Force (C3, C4I) Systems.
11.0802	Data Modeling/Warehousing and Database Administration.	29.0406	Military Information Systems Technology.
11.0803	Computer Graphics.	29.0407	Missile and Space Systems Technology.
11.0804	Modeling, Virtual Environments and Simulation.	29.0408	Munitions Systems/Ordinance Technology.
11.0899	Computer Software and Media Applications, Other.	29.0409	Radar Communications and Systems Technology.
11.0901	Computer Systems Networking and Telecommunications.	29.0499	Military Systems and Maintenance Technology, Other.
11.1001	Network and System Administration/Administrator.	29.9999	Military Technologies and Applied Sciences, Other.
11.1002	System, Networking, and LAN/WAN Management/Manager.	30.0101	Biological and Physical Sciences.
11.1003	Computer and Information Systems Security/Information Assurance.	30.0601	Systems Science and Theory.
11.1004	Web/Multimedia Management and Webmaster.	30.0801	Mathematics and Computer Science.
11.1005	Information Technology Project Management.	30.1001	Biopsychology.
11.1006	Computer Support Specialist.	30.1701	Behavioral Sciences.

Continued on next page

CIP	Title	CIP	Title
11.1099	Computer/Information Technology Services Administration and Management, Other.	30.1801	Natural Sciences.
13.0501	Educational/Instructional Technology.	30.1901	Nutrition Sciences.
13.0601	Educational Evaluation and Research.	30.2501	Cognitive Science.
13.0603	Educational Statistics and Research Methods.	30.2701	Human Biology.
14	Engineering.	30.3001	Computational Science.
15	Engineering Technology, General.	30.3101	Human Computer Interaction.
15.0101	Architectural Engineering Technology/Technician.	30.3201	Marine Sciences.
15.0201	Civil Engineering Technology/Technician.	30.3301	Sustainability Studies.
15.0303	Electrical, Electronic and Communications Engineering Technology/Technician.	40	Physical Sciences.
15.0304	Laser and Optical Technology/Technician.	41	Science Technologies/Technicians, General.
15.0305	Telecommunications Technology/Technician.	41.0101	Biology Technician/Biotechnology Laboratory Technician.
15.0306	Integrated Circuit Design.	41.0204	Industrial Radiologic Technology/Technician.
15.0399	Electrical and Electronic Engineering Technologies/Technicians, Other.	41.0205	Nuclear/Nuclear Power Technology/Technician.
15.0401	Biomedical Technology/Technician.	41.0299	Nuclear and Industrial Radiologic Technologies/Technicians, Other.
15.0403	Electromechanical Technology/Electromechanical Engineering Technology.	41.0301	Chemical Technology/Technician.
15.0404	Instrumentation Technology/Technician.	41.0303	Chemical Process Technology.
15.0405	Robotics Technology/Technician.	41.0399	Physical Science Technologies/Technicians, Other.
15.0406	Automation Engineer Technology/Technician.	41.9999	Science Technologies/Technicians, Other.
15.0499	Electromechanical and Instrumentation and Maintenance Technologies/Technicians, Other.	42.2701	Cognitive Psychology and Psycholinguistics.
15.0501	Heating, Ventilation, Air Conditioning and Refrigeration Engineering Technology/Technician.	42.2702	Comparative Psychology.
15.0503	Energy Management and Systems Technology/Technician.	42.2703	Developmental and Child Psychology.
15.0505	Solar Energy Technology/Technician.	42.2704	Experimental Psychology.
15.0506	Water Quality and Wastewater Treatment Management and Recycling Technology/Technician.	42.2705	Personality Psychology.
15.0507	Environmental Engineering Technology/Environmental Technology.	42.2706	Physiological Psychology/Psychobiology.
15.0508	Hazardous Materials Management and Waste Technology/Technician.	42.2707	Social Psychology.
15.0599	Environmental Control Technologies/Technicians, Other.	42.2708	Psychometrics and Quantitative Psychology.
15.0607	Plastics and Polymer Engineering Technology/Technician.	42.2709	Psychopharmacology.
15.0611	Metallurgical Technology/Technician.	42.2799	Research and Experimental Psychology, Other.
15.0612	Industrial Technology/Technician.	43.0106	Forensic Science and Technology.
15.0613	Manufacturing Engineering Technology/Technician	43.0116	Cyber/Computer Forensics and Counterterrorism.
15.0614	Welding Engineering Technology/Technician.	45.0301	Archeology.
15.0615	Chemical Engineering Technology/Technician.	45.0603	Econometrics and Quantitative Economics.
15.0616	Semiconductor Manufacturing Technology.	45.0702	Geographic Information Science and Cartography.
15.0699	Industrial Production Technologies/Technicians, Other.	49.0101	Aeronautics/Aviation/Aerospace Science and Technology, General.
15.0701	Occupational Safety and Health Technology/Technician.	51.1002	Cytotechnology/Cytotechnologist.
15.0702	Quality Control Technology/Technician.	51.1005	Clinical Laboratory Science/Medical Technology/Technologist.
15.0703	Industrial Safety Technology/Technician.	51.1401	Medical Scientist.
15.0704	Hazardous Materials Information Systems Technology/Technician.	51.2003	Pharmaceutics and Drug Design.
15.0799	Quality Control and Safety Technologies/Technicians, Other.	51.2004	Medicinal and Pharmaceutical Chemistry.
15.0801	Aeronautical/Aerospace Engineering Technology/Technician.	51.2005	Natural Products Chemistry and Pharmacognosy.
15.0803	Automotive Engineering Technology/Technician.	51.2006	Clinical and Industrial Drug Development.
15.0805	Mechanical Engineering/Mechanical Technology/Technician.	51.2007	Pharmacoeconomics/Pharmaceutical Economics.
15.0899	Mechanical Engineering Related Technologies/Technicians, Other.	51.2009	Industrial and Physical Pharmacy and Cosmetic Sciences.
15.0901	Mining Technology/Technician.	51.201	Pharmaceutical Sciences.
15.0903	Petroleum Technology/Technician.	51.2202	Environmental Health.
15.0999	Mining and Petroleum Technologies/Technicians, Other.	51.2205	Health/Medical Physics.

Continued on next page

CIP	Title	CIP	Title
15.1001	Construction Engineering Technology/Technician.	51.2502	Veterinary Anatomy.
15.1102	Surveying Technology/Surveying.	51.2503	Veterinary Physiology.
15.1103	Hydraulics and Fluid Power Technology/Technician.	51.2504	Veterinary Microbiology and Immunobiology.
15.1199	Engineering-Related Technologies, Other.	51.2505	Veterinary Pathology and Pathobiology.
15.1201	Computer Engineering Technology/Technician.	51.2506	Veterinary Toxicology and Pharmacology.
15.1202	Computer Technology/Computer Systems Technology.	51.251	Veterinary Preventive Medicine, Epidemiology, and Public Health.
15.1203	Computer Hardware Technology/Technician.	51.2511	Veterinary Infectious Diseases.
15.1204	Computer Software Technology/Technician.	51.2706	Medical Informatics.
15.1299	Computer Engineering Technologies/Technicians, Other.	52.1301	Management Science.
15.1301	Drafting and Design Technology/Technician, General.	52.1302	Business Statistics.
15.1302	CAD/CADD Drafting and/or Design Technology/Technician.	52.1304	Actuarial Science.
15.1303	Architectural Drafting and Architectural CAD/CADD.	52.1399	Management Science and Quantitative Methods, Other.

References for Online Appendix

Aaronson, Daniel. 1998. “Using Sibling Data to Estimate the Impact of Neighborhoods on Children’s Educational Outcomes.” *Journal of Human Resources,* 915–946.

Aina, Carmen, and Cheti Nicoletti. 2018. “The Intergenerational Transmission of Liberal Professions.” *Labour Economics* 51:108–120.

Ajzenman, Nicolás, Gregory Elacqua, Diana Hincapié, Analía Jaimovich, Florencia López Boo, Diana Paredes, and Alonso Román. 2021. “Career Choice Motivation Using Behavioral Strategies.” *Economics of Education Review* 84 (October): 102173.

Alesina, Alberto, Sebastian Hohmann, Stelios Michalopoulos, and Elias Papaioannou. 2021. “Intergenerational Mobility in Africa.” *Econometrica* 89 (1): 1–35.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. “Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools.” *Journal of Political Economy* 113 (1): 151–184.

Anelli, Massimo, Kevin Shih, and Kevin Williams. 2023. “Foreign Students in College and the Supply of STEM Graduates.” *Journal of Labor Economics* 41, no. 2 (April): 511–563.

Arcidiacono, Peter, Esteban M Aucejo, and V Joseph Hotz. 2016. “University Differences in the Graduation of Minorities in Stem Fields: Evidence from California.” *American Economic Review* 106 (3): 525–562.

Arcidiacono, Peter, V Joseph Hotz, Arnaud Maurel, and Teresa Romano. 2020. “Ex Ante Returns and Occupational Choice.” *Journal of Political Economy* 128 (12): 4475–4522.

Arold, Benjamin W. 2024. “Evolution Vs. Creationism in the Classroom: The Lasting Effects of Science Education*.” *The Quarterly Journal of Economics* (June).

Astorne-Figari, Carmen, and Jamin D. Speer. 2019. “Are Changes of Major Major Changes? The Roles of Grades, Gender, and Preferences in College Major Switching.” *Economics of Education Review* 70 (June): 75–93.

Baker, Rachel, Eric Bettinger, Brian Jacob, and Ioana Marinescu. 2018. “The Effect of Labor Market Information on Community College Students’ Major Choice.” *Economics of Education Review* 65 (August): 18–30.

Baran, Cavit, Eric Chyn, and Bryan A. Stuart. 2024. “The Great Migration and Educational Opportunity.” *American Economic Journal: Applied Economics* 16, no. 3 (July): 354–398.

Barrios-Fernández, Andrés. 2022. “Neighbors’ Effects on University Enrollment.” *American Economic Journal: Applied Economics* 14 (3): 30–60.

Beffy, Magali, Denis Fougère, and Arnaud Maurel. 2012. “Choosing the Field of Study in Postsecondary Education: Do Expected Earnings Matter?” *The Review of Economics and Statistics* 94, no. 1 (February): 334–347.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen. 2019. “Who Becomes an Inventor in America? the Importance of Exposure to Innovation.” *The Quarterly Journal of Economics* 134 (2): 647–713.

Bleemer, Zachary, and Aashish Mehta. 2021. “College Major Restrictions and Student Stratification.” *Available at SSRN 3981921.*

Bostwick, Valerie K., and Bruce A. Weinberg. 2022. “Nevertheless She Persisted? Gender Peer Effects in Doctoral STEM Programs.” *Journal of Labor Economics* 40, no. 2 (April): 397–436.

Breda, Thomas, Julien Grenet, Marion Monnet, and Clémentine Van Effenterre. 2023. "How Effective Are Female Role Models in Steering Girls Towards STEM? Evidence from French High Schools." *The Economic Journal* 133, no. 653 (July): 1773–1809.

Brooks-Gunn, Jeanne, Greg J Duncan, Pamela Kato Klebanov, and Naomi Sealand. 1993. "Do Neighborhoods Influence Child and Adolescent Development?" *American journal of sociology* 99 (2): 353–395.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2014. "Gender, Competitiveness, and Career Choices." *The quarterly journal of economics* 129 (3): 1409–1447.

Canaan, Serena, and Pierre Mouganie. 2021. "Does Advisor Gender Affect Women's Persistence in Economics?" *AEA Papers and Proceedings* 111 (May): 112–116.

Card, David, and A Abigail Payne. 2017. *High School Choices and the Gender Gap in Stem.* Technical report. National Bureau of Economic Research.

Carrell, Scott E, Marianne E Page, and James E West. 2010. "Sex and Science: How Professor Gender Perpetuates the Gender Gap." *The Quarterly journal of economics* 125 (3): 1101–1144.

Castleman, Benjamin L., Bridget Terry Long, and Zachary Mabel. 2018. "Can Financial Aid Help to Address the Growing Need for STEM Education? The Effects of Need-Based Grants on the Completion of Science, Technology, Engineering, and Math Courses and Degrees." *Journal of Policy Analysis and Management* 37 (1): 136–166.

Ceci, Stephen J, Donna K Ginther, Shulamit Kahn, and Wendy M Williams. 2014. "Women in Academic Science: A Changing Landscape." *Psychological science in the public interest* 15 (3): 75–141.

Census, U.S. Bureau of the. 2004. "Average Number of Children Per Family and Per Family with Children, by State: 1990 Census." https://www2.census.gov/programs-surveys/demo/tables/families/states/tabst-f1-2000.xls.

Chetty, Raj, John N. Friedman, and Emmanuel Saez. 2013. "Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings." *American Economic Review* 103, no. 7 (December): 2683–2721.

Chetty, Raj, and Nathaniel Hendren. 2018. "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *The Quarterly Journal of Economics* 133 (3): 1107–1162.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review* 106 (4): 855–902.

Chise, Diana, Margherita Fort, and Chiara Monfardini. 2021. "On the Intergenerational Transmission of Stem Education Among Graduate Students." *The BE Journal of Economic Analysis & Policy* 21 (1): 115–145.

Christian, Alvin, Matthew Ronfeldt, and Basit Zafar. 2024. *College Students and Career Aspirations: Nudging Student Interest in Teaching.* Technical report. National Bureau of Economic Research.

Chyn, Eric. 2018. "Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children." *American Economic Review* 108 (10): 3028–3056.

Chyn, Eric, Robert Collinson, and Danielle H Sandler. 2025. "The Long-Run Effects of America's Largest Residential Racial Desegregation Program: Gautreaux." *The Quarterly Journal of Economics* (February): qjaf011.

Cohodes, Sarah R, Helen Ho, and Silvia C Robles. 2022. *Stem Summer Programs for Underrepresented Youth Increase Stem Degrees.* Technical report. National Bureau of Economic Research.

Conlon, John J. 2021. “Major Malfunction: A Field Experiment Correcting Undergraduates’ Beliefs About Salaries.” *Journal of Human Resources* 56, no. 3 (July): 922–939.

Darolia, Rajeev, Cory Koedel, Joyce B Main, J Felix Ndashimye, and Junpeng Yan. 2020. “High School Course Access and Postsecondary Stem Enrollment and Attainment.” *Educational Evaluation and Policy Analysis* 42 (1): 22–45.

De Giorgi, Giacomo, Michele Pellizzari, and Silvia Redaelli. 2010. “Identification of Social Interactions Through Partially Overlapping Peer Groups.” *American Economic Journal: Applied Economics* 2 (2): 241–275.

De Philippis, Marta. 2023. “Stem Graduates and Secondary School Curriculum: Does Early Exposure to Science Matter?” *Journal of Human Resources* 58 (6): 1914–1947.

Delaney, Judith M., and Paul J. Devereux. 2021. “High School Rank in Math and English and the Gender Gap in STEM.” *Labour Economics* 69 (C).

Derenoncourt, Ellora. 2022. “Can You Move to Opportunity? Evidence from the Great Migration.” *American Economic Review* 112, no. 2 (February): 369–408.

Deutscher, Nathan. 2020. “Place, Peers, and the Teenage Years: Long-Run Neighborhood Effects in Australia.” *American Economic Journal: Applied Economics* 12, no. 2 (April): 220–249.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. “Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach.” *Econometrica* 64 (5): 1001–1044.

Ding, Yanqing, Wei Li, Xin Li, Yinduo Wu, Jin Yang, and Xiaoyang Ye. 2021. “Heterogeneous Major Preferences for Extrinsic Incentives: The Effects of Wage Information on the Gender Gap in Stem Major Choice.” *Research in Higher Education* 62 (8): 1113–1145.

Ellison, Glenn, and Ashley Swanson. 2010. “The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions.” *Journal of Economic Perspectives* 24 (2): 109–128.

Elsner, Benjamin, Ingo E Isphording, and Ulf Zölitz. 2021. “Achievement Rank Affects Performance and Major Choices in College.” *The Economic Journal* 131 (640): 3182–3206.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2016. “Sources of Geographic Variation in Health Care: Evidence from Patient Migration.” *The quarterly journal of economics* 131 (4): 1681–1726.

———. 2021. “Place-based Drivers of Mortality: Evidence from Migration.” *American Economic Review* 111, no. 8 (August): 2697–2735.

Fischer, Stefanie. 2017. “The Downside of Good Peers: How Classroom Composition Differentially Affects Men’s and Women’s Stem Persistence.” *Labour Economics* 46:211–226.

Gottfried, Michael A, and Robert Bozick. 2016. “Supporting the Stem Pipeline: Linking Applied Stem Course-Taking in High School to Declaring a Stem Major in College.” *Education Finance and Policy* 11 (2): 177–202.

Goulas, Sofoklis, Silvia Griselda, and Rigissa Megalokonomou. 2024. “Comparative Advantage and Gender Gap in STEM.” *Journal of Human Resources* 59, no. 6 (November): 1937–1980.

Griffith, Amanda L., and Joyce B. Main. 2019. "First Impressions in the Classroom: How Do Class Characteristics Affect Student Grades and Majors?" *Economics of Education Review* 69 (April): 125–137.

———. 2021. "The Role of the Teaching Assistant: Female Role Models in the Classroom." *Economics of Education Review* 85 (December): 102179.

Hanushek, Eric A, Babs Jacobs, Guido Schwerdt, Rolf Van der Velden, Stan Vermeulen, and Simon Wiederhold. 2023. *Where Do Stem Graduates Stem from? the Intergenerational Transmission of Comparative Skill Advantages.* Technical report. National Bureau of Economic Research.

Hoffmann, Florian, and Philip Oreopoulos. 2009. "A Professor Like Me: The Influence of Instructor Gender on College Achievement." *Journal of Human Resources* 44, no. 2 (March): 479–494.

Jencks, Christopher, Susan E Mayer, et al. 1990. "The Social Consequences of Growing Up in a Poor Neighborhood." *Inner-city poverty in the United States* 111:186.

Jia, Ning. 2021. "Do Stricter High School Math Requirements Raise College Stem Attainment?" *Economics of Education Review* 83:102140.

Joensen, Juanna Schrøter, and Helena Skyt Nielsen. 2016. "Mathematics and Gender: Heterogeneity in Causes and Consequences." *The Economic Journal* 126 (593): 1129–1163.

Justman, Moshe, and Susan J Méndez. 2018. "Gendered Choices of Stem Subjects for Matriculation Are Not Driven by Prior Differences in Mathematical Achievement." *Economics of education review* 64:282–297.

Katz, Lawrence F, Jeffrey R Kling, and Jeffrey B Liebman. 2001. "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment." *The quarterly journal of economics* 116 (2): 607–654.

Kugler, Adriana D., Catherine H. Tinsley, and Olga Ukhaneva. 2021. "Choice of Majors: Are Women Really Different from Men?" *Economics of Education Review* 81 (April): 102079.

Laliberté, Jean-William. 2021. "Long-Term Contextual Effects in Education: Schools and Neighborhoods." *American Economic Journal: Economic Policy* 13 (2): 336–377.

Landaud, Fanny, Son Thierry Ly, and Éric Maurin. 2020. "Competitive Schools and the Gender Gap in the Choice of Field of Study." *Journal of Human Resources* 55, no. 1 (January): 278–308.

Lavy, Victor, and Edith Sand. 2015. *On the Origins of Gender Human Capital Gaps: Short and Long Term Consequences of Teachers' Stereotypical Biases.* Technical report. National Bureau of Economic Research.

Liu, Jing, Cameron Conrad, and David Blazar. 2024. *Computer Science for All? the Impact of High School Computer Science Courses on College Majors and Earnings.* Technical report. IZA Discussion Papers.

Long, Mark C., Dan Goldhaber, and Nick Huntington-Klein. 2015. "Do Completed College Majors Respond to Changes in Wages?" *Economics of Education Review* 49 (December): 1–14.

Ludwig, Jens, Greg J Duncan, Lisa A Gennetian, Lawrence F Katz, Ronald C Kessler, Jeffrey R Kling, and Lisa Sanbonmatsu. 2013. "Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity." *American economic review* 103 (3): 226–231.

Maarseveen, Raoul van. 2021. "Urbanization and Educational Attainment: Evidence from Africa." *Available at SSRN 3836097.*

Montmarquette, Claude, Kathy Cannings, and Sophie Mahseredjian. 2002. "How Do Young People Choose College Majors?" *Economics of Education Review* 21, no. 6 (December): 543–556.

Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson. 2022. "The Gift of Moving: Intergenerational Consequences of a Mobility Shock." *The Review of Economic Studies* 89 (3): 1557–1592.

Ngo, Diana, and Andrew Dustan. 2024. "Preferences, Access, and the Stem Gender Gap in Centralized High School Assignment." *American Economic Journal: Applied Economics* 16 (4): 257–287.

Oguzoglu, Umut, and Serkan Ozbeklik. 2016. *Like Father, Like Daughter (unless There is a Son): Sibling Sex Composition and Women's Stem Major Choice in College.* Technical report. IZA Discussion Papers.

Oliver, Daniel, Robert Fairlie, Glenn Millhauser, and Randa Roland. 2021. "Minority Student and Teaching Assistant Interactions in Stem." *Economics of education review* 83:102125.

Oreopoulos, Philip. 2003. "The Long-Run Consequences of Living in a Poor Neighborhood." *The quarterly journal of economics* 118 (4): 1533–1575.

Ost, Ben. 2010. "The Role of Peers and Grades in Determining Major Persistence in the Sciences." *Economics of education review* 29 (6): 923–934.

Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics* 37, no. 2 (April): 187–204.

Patnaik, Arpita, Joanna Venator, Matthew Wiswall, and Basit Zafar. 2022. "The Role of Heterogeneous Risk Preferences, Discount Rates, and Earnings Expectations in College Major Choice." *Journal of Econometrics,* Annals Issue: Subjective Expectations & Probabilities in Economics, 231, no. 1 (November): 98–122.

Porter, Catherine, and Danila Serra. 2020. "Gender Differences in the Choice of Major: The Importance of Female Role Models." *American Economic Journal: Applied Economics* 12 (3): 226–254.

Price, Joshua. 2010. "The Effect of Instructor Race and Gender on Student Persistence in Stem Fields." *Economics of Education Review* 29 (6): 901–910.

Rask, Kevin. 2010. "Attrition in Stem Fields at a Liberal Arts College: The Importance of Grades and Pre-Collegiate Preferences." *Economics of education review* 29 (6): 892–900.

Reuben, Ernesto, Matthew Wiswall, and Basit Zafar. 2017. "Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender." *The Economic Journal* 127, no. 604 (September): 2153–2186.

Rosenbaum, James E. 1995. "Changing the Geography of Opportunity by Expanding Residential Choice: Lessons from the Gautreaux Program." *Housing Policy Debate* 6 (1): 231–269.

Saltiel, Fernando. 2023. "Multi-Dimensional Skills and Gender Differences in Stem Majors." *The Economic Journal* 133 (651): 1217–1247.

Sampson, Robert J, Jeffrey D Morenoff, and Thomas Gannon-Rowley. 2002. "Assessing "neighborhood Effects": Social Processes and New Directions in Research." *Annual review of sociology* 28 (1): 443–478.

Sampson, Robert J, Stephen W Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *science* 277 (5328): 918–924.

Schwank, Hanna. 2024. "Childhood Migration and Educational Attainment: Evidence from Indonesia." *Journal of Development Economics* 171 (October): 103338.

Sharkey, Patrick, and Jacob W Faber. 2014. “Where, When, Why, and for Whom Do Residential Contexts Matter? Moving Away from the Dichotomous Understanding of Neighborhood Effects.” *Annual review of sociology* 40:559–579.

Sovero, Veronica, Moshe Buchinsky, and Matthew D Baird. 2021. “Playing Catch Up: A Term-Level Investigation of the Racial Gap in Stem Retention.” *Economics of Education Review* 83:102146.

Speer, Jamin D. 2023. “Bye Bye Ms. American Sci: Women and the Leaky Stem Pipeline.” *Economics of Education Review* 93:102371.

Stinebrickner, Ralph, and Todd R. Stinebrickner. 2014. “A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout.” *The Review of Economic Studies* 81, no. 1 (January): 426–472.

Tai, Robert H, Christine Qi Liu, Adam V Maltese, and Xitao Fan. 2006. “Planning Early for Careers in Science.” *Science* 312 (5777): 1143–1144.

THECB. 2017. “Texas Public High School Graduates’ Out-of-State College Matriculation Class of 2011 – 2015.” https://reportcenter.highered.texas.gov/reports/data/special-interest-brief-texas-public-high-school-graduates-out-of-state-college-matriculation/.

Wang, Xueli. 2013. “Why Students Choose Stem Majors: Motivation, High School Learning, and Postsecondary Context of Support.” *American Educational Research Journal* 50 (5): 1081–1121.

Weinstein, Russell. 2022. “Local Labor Markets and Human Capital Investments.” *Journal of Human Resources* 57 (5): 1498–1525.

Wilson, William Julius. 2012. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* University of Chicago Press.

Wiswall, Matthew, and Basit Zafar. 2015. “Determinants of College Major Choice: Identification Using an Information Experiment.” *The Review of Economic Studies* 82 (2): 791–824.

———. 2018. “Preference for the Workplace, Investment in Human Capital, and Gender*.” *The Quarterly Journal of Economics* 133, no. 1 (February): 457–507.

———. 2021. “Human Capital Investments and Expectations About Career and Family.” *Journal of Political Economy* 129, no. 5 (May): 1361–1424.

Wodtke, Geoffrey T, David J Harding, and Felix Elwert. 2011. “Neighborhood Effects in Temporal Perspective: The Impact of Long-Term Exposure to Concentrated Disadvantage on High School Graduation.” *American sociological review* 76 (5): 713–736.

Xie, Yu, Michael Fang, and Kimberlee Shauman. 2015. “Stem Education.” *Annual review of sociology* 41:331–357.

Zafar, Basit. 2013. “College Major Choice and the Gender Gap.” *Journal of Human Resources* 48 (3): 545–595.

Zölitz, Ulf, and Jan Feld. 2021. “The Effect of Peer Gender on Major Choice in Business School.” *Management Science* 67, no. 11 (November): 6963–6979.