Linked in Economic Graph

WORKING PAPER

Decomposing Gender Gaps in US STEM Transition from Education to Employment

Matthew Baird LinkedIn Economic Graph Research & Insights

Nikhil Gahlawat LinkedIn Economic Graph Research & Insights

Rosie Hood LinkedIn Economic Graph Research & Insights

Paul Ko LinkedIn Economic Graph Research & Insights

Silvia Lara LinkedIn Economic Graph Research & Insights

APRIL 2023

EG WP No. 2

Any opinions expressed in this paper are those of the author(s) and not those of LinkedIn. This working paper represents work in progress distributed to encourage discussion and feedback. Citation should account for the provisional character.

Decomposing Gender Gaps in US STEM Transition from Education to Employment

Matthew Baird¹, Nikhil Gahlawat², Rosie Hood³, Paul Ko⁴, and Silvia Lara⁵

April 2023

ABSTRACT⁶

Women are underrepresented in STEM, both in the United States and internationally. Prior literature has primarily focused on disproportionate outflows from STEM that occur during the educational process, from K-12 through college. However, the eventual goal is equitable representation within STEM jobs. We investigate millions of LinkedIn profiles in the United States to understand the transition from STEM degrees to STEM employment. We examine the large drop-off occurring between graduation and the first year of employment, which is especially severe for women, with an approximately ten percentage point widening of the gender STEM gap (the proportion of workers in STEM that are men minus the proportion that are women) during this period. Using a Blinder-Oaxaca decomposition, we demonstrate that several factors impact this gap: major choice, STEM job posting views and applications, and the proportion of the local STEM workforce that are the same gender as the graduate. We estimate that if all of the measured factors were equal between men and women, women would be slightly more likely than men to persist in STEM employment.

Keywords: STEM, gender, human capital, labor JEL classifications: J16, J24, I29

¹ Corresponding author. LinkedIn Corporation. <u>mdbaird@linkedin.com</u>

² LinkedIn Corporation. <u>ngahlawat@linkedin.com</u>

³ LinkedIn Corporation. <u>rhood@linkedin.com</u>

⁴ LinkedIn Corporation. <u>pko@linkedin.com</u>

⁵ LinkedIn Corporation. <u>slara@linkedin.com</u>

⁶ Acknowledgements: we would like to thank Karin Kimbrough, Anne Trapasso, Peter McCrory, Sue Marquez, and Virginia Roode for feedback on this paper, as well as members of the Economic Graph Research & Insights team.

1. Introduction

Science, Technology, Engineering, and Mathematics (STEM) training and subsequent employment are important parts of a healthy economy and in many cases offer higher paying, stable pathways into career progression (Katz & Margo, 2014; Manski, 1992). In recognition of this, many countries, including the United States, have invested in initiatives aimed at encouraging STEM education and increasing the STEM workforce through immigration (National Academies of Sciences & Medicine, 2016). However, despite the potential benefits of careers in STEM, certain populations, particularly women and underrepresented racial and ethnic groups in the United States, remain underrepresented along the entire training to work pipeline, from elementary and secondary school through post-secondary education and employment.

Previous research has mainly focused on underrepresentation and attrition in the educational STEM pipeline (Rodriguez-Solorio, 2022). However, while the STEM gender and minority disparities do grow during education, there is also known and documented widening of the disparities that happens between post-secondary education and STEM employment (Baird et al., 2017). Insufficient attention has examined the timing and reasons for the divergence at this juncture, despite the fact that it represents in many ways the most costly leakage in the pipeline. After years of STEM training, including post-secondary education, a non-negligible fraction of graduates (both men and women) end up not working in STEM, leaving behind at least a portion of their gained human capital (Baird, Ko, et al., 2023). As underrepresentation affects women and underrepresented racial and ethnic groups more often, this creates entrenched inefficiencies for the economy and exacerbates existing equity shortcomings, as these populations are already earning less on average and end up in what may be lower paying career trajectories.

In this paper, we use a proprietary data set of LinkedIn members from graduating cohorts 2019 through 2021 to examine the differences in outcomes between men and women transitioning from STEM post-secondary education to early career employment. We focus our attention on the drop-off that happens between graduation with a STEM degree and one year later, which Baird et al. (2023) show is the leakiest point of the

1

pipeline and the largest contributor to the expansion of the gender gap in STEM fields within the first five years. Specifically, the share of women in STEM fields decreases from 40.3 percent for bachelor's degree STEM graduates to 32.4 percent for bachelor's degree graduates one year later.

In this paper, we perform regression analysis and Oaxaca-Blinder decompositions and find that several factors impact the gender gap, including college major choice within STEM, STEM job search and application behavior, and the proportion of local jobs that are STEM and the proportion of STEM workers locally that are the same gender as the graduate. These results suggest that there may be early differences in desire or perceived fit for entry into STEM careers, or differences in women's perceived access to opportunities in STEM fields. This has important implications for policy options aimed at narrowing the gender gap, such as targeted mentoring programs and recruiting initiatives for companies, educators, and platforms such as LinkedIn.

1.1. Related Literature

The discourse around the underrepresentation of women and minorities in STEM fields often frames the discussion around a leaky pipeline model (Baird et al., 2017; Rodriguez-Solorio, 2022). The gender STEM gap begins as early as middle school (Choney, 2018; Rabenberg, 2013). Gender gaps continue to widen through numerous junctures along the educational and professional path, including initial major declaration in college (Cimpian et al., 2020; Gottfried & Bozick, 2016; Rainey et al., 2018), retention in STEM majors (Sovero et al., 2021), graduation in STEM majors (Arcidiacono et al., 2016), and eventual employment in STEM fields (Baird et al., 2017).

Rodriguez-Solorio (2022) provides an insightful overview of the existing literature on this issue and offers a comprehensive framework for understanding the attrition of women and minorities in STEM. In addition to tracking this attrition through various stages of formal education, Rodriguez-Solorio also highlights the importance of considering the influence of earlier stages in the educational pipeline, such as STEM preparation in high school, on later outcomes, including college persistence in STEM. For example, Card and Payne (2021) show that the gender gap in college persistence in STEM is largely determined by differences in STEM preparation in high school (see also Sovero et al., 2021).

The root causes of differential gender attrition in STEM education have been the focus of numerous studies, with the goal of finding solutions to increase representation. Ireland et al., (2018) summarize the findings of 60 studies and offer recommendations for improvement, and Eddy and Brownell (2016) provide a similar synthesis of research in this field. Additionally, Rodriguez-Solorio (2022) surveys the various programs aimed at increasing retention in STEM majors in college, offering insights into the types of initiatives that are most effective in addressing this issue.

However, despite these efforts, the gender STEM gap continues to persist. While the leaky pipeline model provides a useful framework for understanding why this is the case, deeper analyses are needed to address this issue and ensure that women and minorities are properly represented in STEM fields.

2. Data and Measurements

We use for primary data source the education and employment records of U.S. members of LinkedIn. Additionally, we recognize that gender is not binary, but is a spectrum. Nevertheless, given data limitations, in this paper we use the binary conceptualization of gender, men and women. Furthermore, we rely on developed methodology by LinkedIn which infers gender based on pronouns and names, where feasible (and excludes them when we cannot with confidence infer gender.¹

A full description of the methodology for defining STEM occupations can be found in (Baird, Gahlawat, et al., 2023). We summarize the methods here, which rely on (1) classifying STEM degrees, (2) classifying STEM skills based on STEM degree holders, and (3) classifying STEM occupations based on STEM skills used in those occupations.

¹ The official LinkedIn gender methodology statement: "Gender identity isn't binary and we recognize that some LinkedIn members identify beyond the traditional gender constructs of "man" and "woman." If not explicitly self-identified, we have inferred the gender of members included in this analysis either by the pronouns used on their LinkedIn profiles, or inferred on the basis of first name. Members whose gender could not be inferred as either man or woman were excluded from this analysis."

2.1. Classifying STEM Degrees

For the purposes of this paper, which focuses on retention in STEM fields, we focus our attention on individuals who have earned a STEM college degree (associate's, bachelor's or graduate). However, it is important to note that our data is limited to those who have recorded their education degrees and fields of study on the LinkedIn platform. Therefore, our findings may not be generalizable to all individuals with STEM degrees. In order to define STEM fields, we use the U.S. Department of Homeland Security's (DHS) STEM Designated Degree Program list of majors, which utilizes CIP (Classification of Instructional Programs) codes. The DHS based their list on the U.S. Department of STEM fields. However, there may be limitations in the use of a single governmental source for defining STEM fields.

2.2. Classifying STEM Skills

Using this set of individuals, we create a novel classification of STEM occupations driven by STEM skills. This decision is not innocuous and has important implications for such outcomes as the estimated gender gap (Anderson et al., 2021). In our process, we first develop a list of STEM skills (most closely related to the classification approach of Rothwell (2013)). We define a STEM skill as one for which at least 100 members have added the skill, and in which the probability that a STEM graduate adds it is at least five times as likely as the probability that a non-STEM graduate adds the skill. We chose a threshold of five for several reasons. First, it separates two modes in the density of skill add ratios relatively cleanly (see Baird et al. 2023b), with the larger mass below the threshold representing non-STEM skills and the smaller density mass above the threshold representing STEM skills. Second, the threshold of five aligned the U.S. estimates of STEM in the workforce and gender ratios relatively closely to external benchmarks, such as Rothwell (2013) and Baird et al. (2017). Third, the threshold offers a clean, intuitive separation point.

The most commonly added STEM skills in the United States are Python (Programming Language), SQL, Engineering, Java, JavaScript, C++, MATLAB, Software

Development, Linux, and C (Programming Language). Thus, the list is dominated by computer programming and engineering skills. The skill add ratio can be used to rank how STEM-focused a skill is, with the highest-rated STEM skills among common skills (at least 500,000 members globally have added it) being Core Java (34.5 times more likely for STEM degree holders), Spring Framework (32.2), Algorithms (27.9), Spring Boot (27.9), and C (25.1).

2.3. Classifying STEM Occupations

With the derived list of STEM skills, we next classify which occupations are STEM. We do so using the LinkedIn Skills Genome², which calculates TF-IDF scores across members within occupation to determine the most important and unique skills to each occupation. We classify a STEM occupation as one which has at least one STEM skill in its top ten skills. Using the LinkedIn occupational taxonomy of occupation representative IDs (groupings of occupations yielding 3,194 occupations), 825 are classified as STEM in the United States, or 25.8 percent of the occupations.

The most common STEM occupations in the United States include Software Engineer, Professor, Manufacturing Engineer, System Engineer, Data Analyst, Quality Assurance Manager, Engineering Manager, Mechanical Engineer, Director of Information Technology, and Design Engineer. The most core-STEM occupations (as defined in Baird et al. 2023b) are Geology Specialist, Exploration Manager, Java Consultant, Thermal Engineer, Geophysicist, Computational Biologist, Analytical Chemist, Professor of Chemistry, Research Instructor, and Bioinformatician.

3. Results

3.1. Descriptive Statistics

We used this skills-based approach to estimate the proportion of LinkedIn members who work in STEM occupations in the United States, and found that 19.3% of the membership falls under this category. This is similar to the estimates made by other researchers such

² https://engineering.linkedin.com/blog/2019/how-we-mapped-the-skills-genome-of-emerging-jobs

as Rothwell 2013 at 20% and Anderson et al. 2022 at 19.8%. Our estimate is higher than the traditional estimates that are skills-based, such as the Bureau of Labor Statistics' estimate of 6.2%.³ We acknowledge that our choice of a skills-based approach may have contributed to the higher estimate. Additionally, it is possible that the LinkedIn membership in the United States is more likely to participate in STEM occupations than the overall U.S. population, which may have also contributed to the higher estimate. Therefore, we caution that these findings should be interpreted with this caveat in mind.

Baird, Ko, et al., (2023) have additional details on STEM statistics and trends in the United States using the same methodology as in this paper. We reproduce a selection of their estimates here, which serve as the basis for our regression analysis sample. For instance, we found that among sub-baccalaureate workers with STEM degrees, 40.6% of men and 19.0% of women are currently employed in STEM, resulting in a gap of 21.6%. Although the gap is somewhat smaller for bachelor's degree holders with a STEM degree at 14.6% (47.7% of men and 33.2% of women in STEM occupations), it remains substantial. Baird, Ko, et al., (2023) also show that the fraction of STEM graduates working in STEM increases each year for both men and women, regardless of educational attainment, but the gender gap in STEM employment grows slightly each year. Figure A.1 further disaggregates this finding by field of study, where we see persistent gender gaps even within STEM field of study groups.

Figure 1 presents an exhibit from Baird, Ko, et al., (2023) paper that is of particular interest to our analysis. It shows the proportion of STEM graduates and workers who are women both at graduation from college (year 0) and in employment for years 1 through 5 after graduation. Although we observe a very slight decrease in women's representation after year 1, this decrease is relatively small compared to the substantial increase in the gender gap between graduation and one year after graduation. For example, for bachelor's degree or higher STEM graduates, 40.3% of the graduating class in our data are women, while one year later only 32.4 percent of STEM workers are women. Thereafter the decline is shallow, ending in 31.7 percent of STEM workers five years after

³ <u>https://www.bls.gov/emp/tables/stem-employment.htm</u>

graduating a STEM degree being women. This suggests that the majority of the attrition resulting in under-representation of women in STEM occurs within the first year after graduation, which holds true for both associate degree holders and bachelor's or higher workers.



Figure 1: Proportion of STEM that are women among STEM degree holders, by time since graduation

Note: Repeated from Baird, Ko, et al., (2023). Each group is limited only to those with STEM degrees from the 2016 cohort.

Additionally, Figures A.2 and A.3 in the appendix provide Sankey charts for employment status in each of the first five years after graduation for STEM bachelor's degree holders. We compare the trends for men and women for the 2016 graduating cohort. Three attributes stand out: first, the transitions between states are small compared to the proportion of people who remain stable within their respective groups, indicating that most STEM workers tend to stay in STEM jobs over time after their initial placement one year after graduation, and the same for non-STEM workers in non-STEM. This is not necessarily implied from Figure 1, which could have occurred with high levels of turnover between STEM and non-STEM. This is not the case though, as shown, with workers staying within STEM or non-STEM. Secondly, the small but source of entry of workers into STEM after year 1 comes primarily from workers in non-STEM, as opposed to those not working. Third, the gender gap in STEM is reinforced and evident throughout as well, but the turnover rates are relatively similar between women and men.

Given the substantial decrease in representation for women between graduation and one year later, and the greater stability in STEM employment for the following five years, we focus on the initial drop-off period. Specifically, we use working in STEM one year after graduation as the outcome variable and limit our analysis to the graduation cohorts of 2019, 2020, and 2021 to include key predictors that are only available for recent years. Table 1 provides the summary statistics for the analytical sample used in our regression analysis. The mean values for women and men differ significantly for all variables at the 1 percent level with the exception of the proportion not working one year after graduation (4.7 percent for both men and women).

In our sample of recent graduates, we find that women are 14.1 percentage points less likely to work in STEM one year after graduation than men, and 14.6 percentage points more likely to work in non-STEM fields. The gender differences in other employment outcomes, such as being in an unknown occupation or not working, are small.

We next examine the averages for several potential mechanisms and control variables in our regressions. We observe large gender gaps in job search behavior, with women viewing and applying to fewer job postings overall, and a smaller proportion of STEM job postings compared to men. Specifically, female STEM graduates in our sample have 39.2% of viewed jobs and 42.0% of job applications in STEM occupations, compared to 58.6% and 63.8% for men, respectively. For regression analysis, we set the fraction

equal to zero when they did not view or apply to any job postings. We control for total views and applications, which thus controls for this imputation. Additionally, setting equal to zero is consistent with the notion that they did not view or apply for any STEM job postings

	Women		Men				
	Obs.	Mean	SD	Obs.	Mean	SD	Diff.
Employment and education outcom	пе						
STEM job	673,823	0.337	0.473	942,182	0.478	0.5	-0.141
Non-STEM job	673,823	0.456	0.498	942,182	0.31	0.462	0.146
Unknown occupation	673,823	0.159	0.365	942,182	0.164	0.37	-0.005
No job	673,823	0.047	0.213	942,182	0.047	0.212	0.000
Individual job search activity							
Sub-baccalaureate worker	673,823	0.034	0.182	942,182	0.061	0.238	-0.027
IHS number of job posting views	673,823	3.247	2.489	942,182	3.46	2.54	-0.213
% job posting views that are STEM	520,245	0.392	0.382	744,951	0.586	0.394	-0.194
% job posting views that are STEM (zero-added)	673,823	0.302	0.374	942,182	0.463	0.424	-0.161
IHS number of job applications	673,823	0.797	1.518	942,182	0.998	1.693	-0.201
% job applications that are STEM	198,017	0.42	0.438	323,614	0.638	0.427	-0.218
% job applications that are STEM (zero-added)	673,823	0.124	0.305	942,182	0.219	0.393	-0.095
Local market conditions							
IHS number of local job posts	673,823	9.938	5.558	942,182	9.866	5.499	0.072
% local posts that are STEM	522,149	0.262	0.075	731,945	0.259	0.075	0.003
% local posts that are STEM (zero-added)	673,823	0.203	0.128	942,182	0.202	0.127	0.001
% local STEM seniority that are women	528,321	0.33	0.029	741,655	0.327	0.03	0.003
% local Non-STEM seniority that are women	528,327	0.535	0.019	741,666	0.534	0.022	0.001

Table 1: Summary statistics from analytic sample

Note: Obs.: number of observations. SD: standard deviation. Diff: difference between men and women mean values. The difference between men and women's mean values is statistically different for each variable at p<0.01 except for "no job".

When looking at local STEM conditions, men and women do not live in substantially different local labor markets (e.g., metropolitan areas). There is only approximately three-tenths of a percentage point difference in local postings that are for STEM positions between men and women, both at around 26.2 percent of posts. Additionally, both men and women live in markets where approximately 33 percent of workers in STEM are women, and 53% of non-STEM workers are women.

3.2. Regression Analysis

We next turn to regression analysis to understand the specific transition from STEM education to their employment situation a year after graduation, where we see the most dramatic drop-off in the representation of women in STEM as shown in Figure 1.

To investigate the relationship between various predictors and the likelihood of being employed in a STEM job one year after graduation, we estimate a series of linear regressions with different sets of covariates as well as an indicator for gender. Table 2 summarizes the results of these models. Contrasting these regressions allows us to examine the extent to which the gender gap changes when controlling for other factors, as well as to investigate which factors impact STEM participation. The first column presents the raw difference without controlling for any covariates. In the subsequent columns, we add additional controls, including college major (column 2), personal activity and attributes (column 3), local labor conditions (column 4), and all covariates combined (column 5).

First, we observe that controlling for major choice reduces the gender gap in STEM employment by nearly half, from 17.1 percentage points to 10.2 percentage points. This suggests that women tend to select STEM majors (such as psychology and majors within social sciences) that are less likely to lead to STEM employment compared to men, who are more likely to major in fields such as engineering and computer science that have higher transition rates into STEM jobs. Including job search covariates such as job views and search behavior (column 3) leads to a larger reduction in the gender gap. Although not causally identified, this reduction from 17.1 to 5.5 percentage points suggests that these individual behaviors are significant in determining STEM employment outcomes.

In other words, if a man and woman have similar STEM job search behavior, the difference in their likelihood of being employed in STEM one year after graduation is relatively small, although it exists.

	No controls	Field of study controls	Job search	Market covariates	All covariates
Male	0.171***	0.102***	0.055***	0.088***	-0.018*
	(0.001)	(0.026)	(0.001)	(0.013)	(0.008)
Sub-baccalaureate			-0.098***		-0.102***
			(0.002)		(0.002)
IHS number of job posting			-0.043***		-0.037***
views			(0.000)		(0.001)
IHS number of job			0.015***		0.013***
applications			(0.000)		(0.000)
% job posting views that			0.667***		0.600***
are STEM			(0.001)		(0.004)
% job applications that are			0.104***		0.090***
STEM			(0.001)		(0.002)
IHS number of local job				0.000	0.001
posts				(0.002)	(0.001)
% local posts that are				0.170**	0.111**
STEM				(0.054)	(0.037)
% STEM workers who are				0.211***	0.151***
the same gender				(0.031)	(0.021)
% Non-STEM workers who				-0.136*	-0.071*
are the same gender				(0.054)	(0.028)
FE: Major		Х			Х
FE: Market area				Х	Х
FE: Year					Х
Number of observations	1354832	1354832	1354832	1354832	1354832
R ² adjusted	0.029	0.134	0.311	0.036	0.332
R ² within adjusted		0.010		0.029	0.234
AIC	1927305	1772230	1461627	1917452	1420232
Std. Errors	IID	By major	IID	Market area	Market area
				urcu	uicu

Table 2: Linear Regressions of STEM employment one year after graduation

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. IHS: inverse hyperbolic sine.

When accounting for market conditions in column 4, which includes market fixed effects, the gender gap in STEM participation decreases from 17.1 percentage points to 8.8 percentage points. The likelihood of working in STEM is higher when the market has a greater proportion of STEM-related job openings and when workers in STEM are predominantly of the same gender as the graduating student. Additionally, after controlling for all covariates in column 5, the gender gap actually reverses, and women are 1.8 percentage points more likely than men to work in STEM one year after graduation. This reversal of a gap has been observed before, as shown for example in Sovero et al. (2021) in Table 7, where the racial gaps in STEM graduation in college are eliminated and flipped when controlling for all covariates.

Several coefficients in model 5 are of interest. Viewing more job postings overall is negatively related to the likelihood of ending up in STEM (while holding constant the number of jobs applied to), while applying for more jobs (holding constant job viewing behavior) is positively related to ending up in STEM. This suggests that having a high applications-to-views ratio is positively related to working in STEM. On average, women in the sample view and apply to fewer jobs than men, with a larger disparity in application behavior, which may contribute to the gender gap in STEM. The likelihood of ending up in STEM is highly related to viewing and applying to a greater proportion of STEM jobs, which as discussed earlier may serve as an indicator of interest in STEM work. As shown in Table 1, women tend to view and apply to a smaller proportion of STEM jobs compared to men, indicating possible gender-based differences in preferences and beliefs regarding barriers, access, and perceived fit into STEM jobs (Cheryan et al., 2011; Moè et al., 2021; Rittmayer & Beier, 2008). These beliefs may include concerns about gender discrimination within STEM jobs and whether women feel confident in their ability to be considered strong candidates for these positions.

Our analysis shows that local labor conditions have explanatory power, even after controlling for other factors. Specifically, we find that a ten percentage point increase in the proportion of job postings that are STEM-related is associated with a one percentage point increase in the likelihood of working in STEM, highlighting the significance of local demand for STEM jobs. Interestingly, this effect is of a similar magnitude as the impact of increasing the proportion of job applications that are STEM-related, suggesting an unexpected equality in impact between these labor demand and supply effects. Additionally, there is evidence for homophily in the data. A ten percentage point increase in the share of local STEM workers who share the same gender as the graduate increases the likelihood of working in STEM by 1.5 percentage points. Conversely, a 10 percentage point increase in the share of non-STEM workers who share the same gender as the graduate decreases the likelihood of working in STEM by 0.7 percentage points. These findings suggest that market conditions, including local demand for STEM and homophily, play a role in shaping transition into STEM careers.

	Men	Women	Difference
Sub-baccalaureate	-0.098***	-0.113***	0.016***
	(0.003)	(0.003)	(0.004)
IHS number of job posting	-0.041***	-0.032***	-0.009***
views	(0.001)	(0.001)	(0.000)
IHS number of job applications	0.014***	0.011***	0.003***
	(0.000)	(0.001)	(0.000)
% job posting views that are	0.572***	0.649***	-0.077***
STEM	(0.004)	(0.004)	(0.002)
% job applications that are	0.089***	0.107***	-0.018***
STEM	(0.002)	(0.003)	(0.002)
IHS number of local job posts	0.001	0.001	0.000
	(0.001)	(0.001)	(0.000)
% local posts that are STEM	0.117**	0.095**	0.022
	(0.042)	(0.032)	(0.018)
% STEM workers who are the	0.140**	0.105**	0.035
same gender	(0.046)	(0.036)	(0.074)
% Non-STEM workers who are	-0.070	0.005	-0.075
the same gender	(0.068)	(0.060)	(0.121)

Table 3: Linear Regressions of STEM employment one year after graduation, by gender

Fixed effects: major, market area, year

Number of observations: 1,354,832

R² adjusted: 0.334, R² within adjusted: 0.236, AIC: 1416175

Std. errors clustering: market area

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. IHS: inverse hyperbolic sine. Regressions additionally control for gender. In Table 3, we re-estimate the full model from column 5 in Table 2 but interact each coefficient with gender to estimate separate coefficients for men (column 1 in Table 3) and women (column 2). Column 3 presents the difference in coefficients. Men have large responsiveness to number of postings, but smaller to the types of job posts (the fraction that are STEM). Meanwhile, there are no significant differences in the responsiveness to local market conditions.

3.3. Oaxaca-Blinder Decomposition

We next take the full model presented in Table 2 column 5 and implement a Oaxaca-Blinder decomposition (Blinder, 1973; Oaxaca, 1973) which allows us to investigate the extent to which gaps may be explained by differences in the variables (or, "endowments" as the literature calls them, resulting in the "explained" portion of the gap) compared to differences in the coefficients (the "unexplained" portion of the gap). There are many versions of the model based on whether the decomposition is a two-fold or three-fold, and if a two-fold, what weights are used to determine the common coefficients. Here, we use a two-fold model with the pooled regression weights with group indicators (Jann, 2008). We use the "oaxaca" package in R to estimate the model (Hlavac, 2022). Additionally, given package and data limitations, for this analysis we dropped markets which had 100 or fewer observations across the three years, which accounted for around one half of one percent of the sample. Table A.1 in the appendix contrasts the regression coefficients for the full sample and this 99.5% subsample. The results are relatively similar here.

Table 4 presents the break-down of the explained versus unexplained portions if just the college major, the job search covariates, or the local market conditions are included, as well as if all are included. These results align with the findings from Table 2. When college major fixed effects are used, approximately 41% (0.0696/(0.0696+0.1022)) is explained by differences in the endowments (differences in which majors they are graduating from), whereas just over half remain unexplained and is unexplained. When just the job search measures are included, 68% of the difference is explained by differences in the endowments, and market conditions explain 45%. When all covariates are included, all of the gap is explained and more, given the earlier-reported finding where all covariates flip the sign of the gap, giving women a slight edge.

		_		
	Covariates included			
	College major	Job search	Market	All
Explained portion	0.0696	0.1165	0.0770	0.1875
	(0.0004)	(0.0005)	(0.0073)	(0.0064)
Unexplained portion (using	0.1022	0.0553	0.0948	-0.0156
men as reference group)	(0.0009)	(0.0007)	(0.0072)	(0.0063)
_				

Table 4: Blinder-Oaxaca Decomposition Results

Note: Standard errors in parentheses

Figure 2 presents a graphical representation of the Oaxaca-Blinder decomposition of the gap by endowment and unexplained portions by covariates. The explained portion of the gap driven by differences in the covariates are driven by two primary factors: the proportion of job views that are STEM positions, and the proportion of the local STEM workforce who are the same gender. The remaining covariates are substantially smaller in their explained impact on the gap. As for the unexplained portion, the overall impact of differences in coefficients is small in explaining the gap (see Table 4), and is driven primarily by differences in the impact of job views—both total count and fraction STEM—on STEM employment. And in fact, the differences in the coefficients help narrow the gap—that is, the gender gap in STEM would be even wider if men and women shared similar coefficients.

Take for example the proportion of job posting views that are STEM. From Table 3, we can see that for a ten percentage point increase in the share of job posting views that are STEM for women (say, from viewing 40% STEM postings to 50%), the likelihood of ending up in a STEM position is predicted to increase by 6.49 percentage points. For the same ten percentage point increase in the share of job posting views that are STEM for men, we predict an increase in the likelihood of working in STEM of only 5.72 percentage points, slightly lower. Thus, the higher responsiveness of women than men to these factors actually keep the gender gap narrower than it would be otherwise. Nevertheless,

this advantage is not large enough to compensate for how different the endowments are. Taking this example again, while women may be more responsive in the data to viewing STEM job postings and the translation into working in a STEM job, they are much less likely to view STEM job postings than men (39.2 versus 58.6), overshadowing any advantages they would have had from that stronger relationship between the two.



Figure 2: Oaxaca-Blinder decomposition by covariate

4. Discussion

Increasing gender representation in STEM employment may have several important benefits, both for productivity and scientific advancement as well as equity and narrowing of income gaps between men and women. Most research in the past has been focused on documenting and exploring gender STEM gaps in education, especially in post-secondary education when individuals most concretely commit to a career path. This paper contributes to the literature by focusing on what happens after graduation from college with a STEM degree, and an exploration into potential reasons for the gap. In some ways, separation from STEM at this point in the pipeline is the most costly, as it may involve skills and proficiencies that took years to develop that the worker does not put into use in their employment at all. Insofar as structural barriers reinforce gender gaps, this would lead to widened inequities.

We find that the most important point of gap-widening is within the first year after graduation from STEM. Thereafter, while there is still an increase in the share of STEM graduates working in STEM over time for both men and women, the gap between them stays relatively stable after the first year of employment. For example, while women representation in STEM majors for bachelor's degree holders is 40.3 percent, one year later in the labor market it drops to 32.4 percent. Thereafter, there is a shallow continual decrease in female representation over the next few years, to 31.7 percent five years after graduation.

In exploring that first-year outcome, we find that may factors are important predictors of the gap – from college major, to job search behavior, to local STEM market conditions. Of that observable gap, the largest difference driving the observed gap in outcomes is due to differences in the fraction of job postings that they view that are STEM, with women coming from STEM degrees having around 39.2% of all job postings they view being for STEM positions, while men view around 58.6%. This difference is suggestive of differences in preferences and beliefs about potential success in the STEM labor market. Reinforcing this, the second most impactful difference between men and women in explaining the STEM gender gap is the proportion of the local STEM workforce that are the same gender as the graduate – which is approximately 1/3rd of the workforce

for women, and 2/3rds for men. This may help explain why women are less likely to view and subsequently apply for STEM positions. In addition, college major (among the STEM majors) is a large predictor of whether the graduate ends up in a STEM job or not, explaining along 41% of the overall gap. Some STEM majors have strong pipelines into STEM occupations, such as engineering, while other majors have weaker pipelines to STEM occupations, such as some within social science.

While most of the gap is explained by difference in endowments, such as those explained in the prior paragraph, differences in how men and women respond also contribute to the gap. However, most of these actually help narrow the gap. That is, the gap would be even wider if men and women had the same relationships between the predictors and working in STEM. For example, for women, each increase in the proportion of job posting views that are STEM leads to a bigger increase in the probability of ending up in STEM than the same increase for men. However, these stronger relationships are not enough to offset the larger differences in the proportion of job views that are STEM compared to men, they have a much smaller level currently. Women views contain a smaller fraction of STEM postings than men.

These findings suggest that the key to decreasing gender gaps in STEM lies both earlier in the educational pipeline (to increase representation of women in strong STEM feeder majors, such as engineering), as well as to find ways to encourage new women graduates to consider – and apply for – STEM jobs. This could mean more outreach from companies towards women just graduating, as well as communication from educators about women having a place in STEM. There also may be a role for platforms such as LinkedIn to provide insights encouraging women into STEM and to highlight specific STEM jobs that a graduate would be a good fit for. This would then serve a reinforcing mechanism, as the local market would contain more gender representation which in turn would increase the retention of women in STEM as well.

There are many paths for future research. We hope to investigate discouragement more concretely after applying for STEM positions, and the extent to which nudges towards STEM jobs can lead to higher view and application rates. We further plan to investigate international trends to compare to the United States. This will also help in contrasting majors and industries with higher or lower fractions of women to separate out preferences for other industry characteristics that may be correlated with STEM representation of women. We will also use their peer network to understand the impact of changes into and out of STEM for peers on their own decisions, recruiter outreach differences, and the role of the rise of remote work, to understand if this leads to an increase in women representation in STEM and what the long term implications are, especially if career advancement differs depending on remote work status.

REFERENCES

- Anderson, D., Baird, M., & Bozick, R. (2021). Who Gets Counted as STEM? A New Approach for Measuring the STEM Workforce and its Implications for Identifying Gender Disparities in the Labor Market. *International Journal of Gender, Science and Technology*, 13(3), 254–279.
- Arcidiacono, P., Aucejo, E. M., & Hotz, V. J. (2016). University differences in the graduation of minorities in STEM fields: Evidence from California. *American Economic Review*, 106(3), 525–562.
- Baird, M., Bozick, R., & Harris, M. (2017). Postsecondary Education and STEM Employment in the United States: An Analysis of National Trends with a Focus on the Natural Gas and Oil Industry. RAND.
- Baird, M., Gahlawat, Ni., Hood, R., Ko, P., & Lara, S. (2023). LinkedIn STEM Classification Methodology. *LinkedIn Economic Graph Technical Note*, 1. https://economicgraph.linkedin.com/content/dam/me/economicgraph/enus/PDF/linkedin-stem-classification-methodology-egtn01.pdf
- Baird, M., Ko, P., Gahlawat, N., Lara, S., & Hood, R. (2023). Measuring gender gaps in the U.S. STEM workforce. *Economic Graph White Paper*.
- Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 436–455.
- Card, D., & Payne, A. A. (2021). High school choices and the gender gap in STEM. *Economic Inquiry*, *59*(1), 9–28.
- Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B. J., & Kim, S. (2011). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM? *Social Psychological and Personality Science*, 2(6), 656–664.
- Choney, S. (2018). Why do girls lose interest in STEM? New research has some answers and what we can do about it. *Microsoft Stories*.
- Cimpian, J. R., Kim, T. H., & McDermott, Z. T. (2020). Understanding persistent gender gaps in STEM. *Science*, *368*(6497), 1317–1319.
- Eddy, S. L., & Brownell, S. E. (2016). Beneath the numbers: A review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines. *Physical Review Physics Education Research*, 12(2), 020106.
- Gottfried, M. A., & Bozick, R. (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.
- Hlavac, M. (2022). *oaxaca: Blinder-Oaxaca Decomposition in R*. [R package version 0.1.5.]. https://CRAN.R-project.org/package=oaxaca
- Ireland, D. T., Freeman, K. E., Winston-Proctor, C. E., DeLaine, K. D., McDonald Lowe, S., & Woodson, K. M. (2018). (Un) hidden figures: A synthesis of research

examining the intersectional experiences of Black women and girls in STEM education. *Review of Research in Education*, 42(1), 226–254.

- Jann, B. (2008). The Blinder–Oaxaca decomposition for linear regression models. *The Stata Journal*, *8*(4), 453–479.
- Katz, L. F., & Margo, R. A. (2014). Technical change and the relative demand for skilled labor: The united states in historical perspective. In *Human capital in history: The American record* (pp. 15–57). University of Chicago Press.
- Manski, C. (1992). Income and higher education. *Focus*, 14(3). https://scholar.google.com/scholar_lookup?title=Income%20and%20higher%20 education&author=C.F.%20Manski&publication_year=1992
- Moè, A., Hausmann, M., & Hirnstein, M. (2021). Gender stereotypes and incremental beliefs in STEM and non-STEM students in three countries: Relationships with performance in cognitive tasks. *Psychological Research*, *85*(2), 554–567.
- National Academies of Sciences & Medicine. (2016). *Promising practices for strengthening the regional STEM workforce development ecosystem*. National Academies Press.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 693–709.
- Rabenberg, T. A. (2013). *Middle school girls' STEM education: Using teacher influences, parent encouragement, peer influences, and self efficacy to predict confidence and interest in math and science* [PhD Thesis]. Drake University.
- Rainey, K., Dancy, M., Mickelson, R., Stearns, E., & Moller, S. (2018). Race and gender differences in how sense of belonging influences decisions to major in STEM. *International Journal of STEM Education*, 5(1), 1–14.
- Rittmayer, A. D., & Beier, M. E. (2008). Overview: Self-efficacy in STEM. *SWE-AWE CASEE Overviews*, 1(3), 12.
- Rodriguez-Solorio, C. (2022). The Impact of a Summer Program on the Development of Historically Underrepresented Minorities in STEM.
- Rothwell, J. (2013). *The hidden STEM economy*. Metropolitan Policy Program at Brookings.
- Sovero, V., Buchinsky, M., & Baird, M. D. (2021). Playing catch up: A term-level investigation of the racial gap in STEM retention. *Economics of Education Review*, *83*, 102146.

APPENDIX

Defining STEM skills

Our first step is to define STEM skills. Here, by STEM skills we mean skills that are specific to STEM, and not universal skills that would be part of any proficiency or competency, such as dependability and basic literacy and numeracy. To create this list, we calculated the probability that a STEM graduate had added each given skill and contrasted it to the probability that a non-STEM graduate had added the skill. We took the ratio of the two probabilities and classified a skill as STEM if this odds ratio exceeded 5. See Baird, Gahlawat, et al., (2023) for details into this distribution and the methodology.

Defining STEM occupations

We define STEM occupations using the list of STEM skills. To do so, we merged STEM skills onto the list of skills unique to different occupations using LinkedIn's Skills Genome.¹⁰ Prior work at LinkedIn created a list of skills related with each occupation using the skills genome. This list is ranked by TF-IDF scores within each occupation. For each occupation, we retained the top ten skills for each occupation. We merged this onto the list of STEM skills developed above and classified the occupation as STEM if at least one skill was in the top ten for that occupation is a STEM skill.

¹⁰ https://engineering.linkedin.com/blog/2019/how-we-mapped-the-skills-genome-of-emerging-jobs

	Full	Subsample
Male	-0.018*	-0.016+
	(0.008)	(0.009)
Sub-baccalaureate	-0.102***	-0.102***
	(0.002)	(0.002)
IHS number of job posting views	-0.037***	-0.037***
	(0.001)	(0.001)
IHS number of job applications	0.013***	0.013***
	(0.000)	(0.000)
% job posting views that are STEM	0.600***	0.600***
	(0.004)	(0.004)
% job applications that are STEM	0.090***	0.090***
	(0.002)	(0.002)
IHS number of local job posts	0.001	0.001
	(0.001)	(0.001)
% local posts that are STEM	0.111**	0.111**
	(0.037)	(0.039)
% STEM workers who are the same gender	0.151***	0.145***
	(0.021)	(0.022)
% Non-STEM workers who are the same gender	-0.071*	-0.067*
	(0.028)	(0.030)
Number of observations	1,354,832	1,347,369
R2 Adj.	0.332	0.333
R2 Within Adj.	0.234	0.234
AIC	1420232	1411380
FE: Market	Х	Х
FE: Major	Х	Х
FE: Year	Х	Х

Table A.1: Comparison of main results under sample restriction dropping markets with

 100 or fewer observations

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors clustered at the market level in parentheses

Figure A.1: Retention in STEM one year after graduation by field of study and gender





Figure A.2: Sankey flow diagram for women with STEM bachelor's degrees, transitions between stages in first five years after graduation from STEM degree



Figure A.3: Sankey flow diagram for men with STEM bachelor's degrees, transitions between stages in first five years after graduation from STEM degree

