Friederike Hertweck

Judith Lehner

The Gender Gap in STEM: (Female) Teenagers' ICT Skills and Subsequent Career Paths

Imprint

Friederike Hertweck and Judith Lehner

The Gender Gap in STEM: (Female) Teenagers' ICT Skills and Subsequent Career Paths

Bibliografische Informationen
der Deutschen Nationalbibliothek

Friederike Hertweck and Judith Lehner[1]

The Gender Gap in STEM: (Female) Teenagers' ICT Skills and Subsequent Career Paths

Abstract

Skills shortage in the fields of Sciences, Technology, Engineering, and Mathematics (STEM) poses a significant challenge for industries globally. This study examines the interrelationship between high school students' gender, their proficiency in Information and Communication Technology (ICT), and their career preferences in the STEM domain. Based on representative data for German teenagers, our study shows that female adolescents are less likely to choose a career in STEM unless they have strong ICT skills in secondary school. The relationship does not hold for male students. Our findings can be explained with evidence that teenagers sort into occupations they believe to be good at and that female teenagers rather underestimate their true potential. Using different empirical approaches, we also show that ICT skills act as a moderator and not as a mediator in the gender-specific choice of training upon graduating from secondary school.

JEL-Codes: I20, J24, J08

Keywords: Career choice; ICT skills; digital literacy; gender gap; STEM

June 2023

1 Introduction

Many companies and industries are struggling to find qualified candidates to fill open positions in sciences, technology, engineering, and mathematics (STEM). This shortage of STEM talent is due to the persistent gender gap in STEM, the lack of investment in STEM education, inadequate training, and the high level of competition for top talent (e.g., Card & Payne, 2021; Cedefop, 2018; Calvani, Fini, Ranieri, & Picci, 2012; Steedman, Wagner, & Foreman, 2006). To overcome skills shortage in STEM, more adolescents in general and female youth in particular must choose a STEM occupation after completing high school.

This paper analyses the interplay between (female) teenagers' digital skills and their enrollment in STEM fields upon graduating from high school. The study is based on the German National Educational Panel Study (NEPS) that combines data from testing teenagers' skills in Information and Communications Technology (ICT skills) with their subsequent educational trajectories.

We use a panel of 9,315 teenagers who were first tested in 2010 during ninth grade and were repeatedly asked about their educational and occupational choices. In a set of linear probability models, we regress educational and occupational choices during the five years after graduating from high school on students' ICT skills and a large set of socio-economic characteristics. Moreover, we perform moderation and mediation analyses to understand the interplay between gender, ICT skills, and choice of STEM fields.

Our results show that girls are much more likely to choose a STEM field upon graduating from high school only if they have strong ICT skills in ninth grade. This relationship is not present for boys. More precisely, an increase in girls' ICT skills by 10 percentage points in ninth grade is associated with an increase in the probability to choose a STEM field upon graduating from high school by 2.95 percentage points or around 25 %. ICT skills in grade 9 moderate the gender-specific sorting into STEM fields. We explain this finding with existing evidence from psychological research that shows that teenagers rather choose occupations that span tasks they believe to be good at (Tracey & Hopkins, 2001) and that female teenagers rather underestimate their true potential (Correll, 2001; Lubienski, Robinson, Crane, & Ganley, 2013; Jouini, Karehnke, & Napp, 2018).

Moreover, we find that higher ICT skills close to graduating from high school (i.e., in grade 12) are positively associated with choosing a STEM field – irrespective of gender. A potential underlying reason for the positive association is that students in upper secondary school have to specialize in some fields and, as Card and Payne (2021) show, the selection into or out of STEM already partially takes place in high school. Thereby, our results complement recent evidence provided by Card and Payne (2021). We show that the selection out of STEM fields is at least partly driven by teenagers' ICT skills in ninth grade and that girls sort into STEM fields upon graduating from high school only if they have above average ICT skills.

By testing whether and how teenagers' ICT skills affect their subsequent fields of training, this study helps identify potential intervention points for addressing gender disparities in STEM fields. We show that ICT skills in ninth grade affect long-run educational decisions. Promoting ICT skills and confidence from secondary school onwards especially among female youth can boost enrollment in STEM and may combat the persistent gender gap in STEM. This finding is of particular importance because recent research on the gender gap in ICT skills shows that gender-specific difference in digital skills is negligibly small among teenagers and emerges at age 18 in favor of boys (e.g., Hatlevik & Christophersen, 2013; van Deursen & van Diepen, 2013; Bachmann, Hertweck, Kamb, Lehner, & Niederstadt, 2022). Despite these small differences, girls often select themselves out of STEM-related fields in upper secondary school (Card & Payne, 2021).

Yet, existing studies on the gender gap in STEM or in ICT skills do not address the potential consequences of female teenagers' ICT skills on labor market decisions. By addressing the interplay between

gender, ICT skills, and educational choices, the present study uncovers an additional lever of how to mitigate skills shortage in STEM. The findings of this study thereby also add to the ongoing policy debate on increasing the intensity of ICT training in secondary school (e.g., Vegas, Hansen, & Fowler, 2021; KMK, 2016). By fostering ICT skills and strengthening confidence in ICT, especially among adolescent girls in secondary education, the likelihood of more teenagers pursuing STEM careers can be enhanced. This approach may effectively address the enduring gender disparity in STEM fields.

The remainder of the paper is organised as follows. We describe the data in section 2 and our empirical strategy in section 3. The results are presented in section 4 and discussed in section 5. Section 6 concludes.

2 Data and Methods

2.1 Definition and Testing of ICT skills

ICT skills are often labelled as *ICT literacy* or *digital literacy* and have been initially described as "the ability to understand and use information in multiple formats from a wide range of sources when it is presented via computers" (Gilster, 1997, p. 1). However, since the required digital competencies are subject to continuous technological change, it is difficult to find a universally valid definition of the term over a long period of time (Leu, Kinzer, Coiro, & Cammack, 2004; Sparks, Katz, & Beile, 2016).

In 2002, the Educational Testing Service (ETS) established the ICT Literacy Panel and defined ICT Literacy as "using digital technology, communications tools, and/or networks to access, manage, integrate, evaluate, and create information in order to function in a knowledge society" (International ICT Literacy Panel, 2002, p. 2). These five components – access, manage, integrate, evaluate, and create – have subsequently been used in this (or slightly modified) form in numerous studies to describe ICT skills.

For instance, Sparks et al. (2016) find most of these five components in all twelve peer-reviewed definitions from national and international organizations, including OECD and UNESCO. In their studies on ICT skills, Binkley et al. (2012), Chetty et al. (2018), Leu et al. (2004), Eshet-Alkalai (2004), and van Deursen, Helsper, and Eynon (2016) also base their measurements on these components. Some studies also address ethical and legal issues relating to the use of information on the Internet (Chetty et al., 2018; Sparks et al., 2016; Van Laar, van Deursen, van Dijk, & de Haan, 2017) or add the time needed to solve an ICT problem as an additional dimension (Goldhammer, Naumann, & Keßel, 2013).

2.2 Description of the data

The present study is based on the German National Educational Panel Study (NEPS).[1] The NEPS is a representative multi-sequence study designed to research educational trajectories over the life course. It provides the possibility to link data from detailed competence tests to teenagers' educational trajectories.

ICT skills are tested among kids, teenagers, college students, and adults. The NEPS takes up four of the aforementioned process components – access, manage, evaluate, and create – in the design of the ICT skill tests. Each component covers both technological and cognitive aspects (Senkbeil, Ihme, & Wittwer, 2013, 2021).

[1]Researchers who work with data from the National Educational Panel Study are obliged to indicate the use of the data by including the following statement and reference: *This paper uses data from the National Educational Panel Study (NEPS; see NEPS-Netzwerk (2021)). The NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi, Germany) in cooperation with a nationwide network.*

Because we are interested in establishing the (gender-specific) link between teenagers' ICT skills and their subsequent career paths, we concentrate our analysis on a panel of teenagers who were first surveyed in the year 2010 when they attended ninth grade of high school (NEPS-Netzwerk, 2021). Since then, the respondents have been repeatedly asked about their lives. Hence, the panel dataset comprises data on each student's educational and occupational choices including the type and field of training. We can only distinguish between female and male teenagers in the data.

In 2010, the teenagers participated in ICT skill tests. The majority of these teenagers had an additional test in the school year 2013/14 during 12th grade shortly before graduating from high school. During the ICT skill test in ninth grade (12th grade), respondents had to answer 39 (32) multiple choice items of different complexity that measured the extent to which they were able to access, manage, evaluate, and create digital content using various software application groups including word processing, spreadsheet, presentations, e-mail, Internet, and search engines (Senkbeil et al., 2021). The tests were paper-based and contained pictures of screenshots of standard software applications.

The NEPS test design comprises both simple and complex multiple-choice questions. It may be argued that the tests may rather capture the stock of knowledge and not necessarily the practical skills. However, a comparison between NEPS and the simulation-based items of the International Computer and Information Literacy Study (ICILS) shows that the NEPS in fact provides a good mapping of digital skills (Sparks et al., 2016; Senkbeil & Ihme, 2020).

Figure 1: Timeline of ICT skills tests and training choices
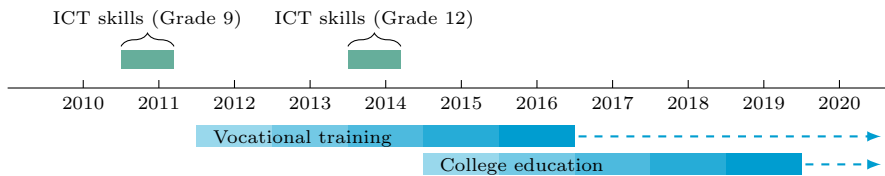
Note: The figure illustrates the timeline of events extracted from NEPS data. Secondary school students are tested in grade 9 (school year 2010/11) and some also again in grade 12 (school year 2013/14). Tests are indicated by the green boxes. Teenagers who are interested in vocational training can start the training after grade 9 at the earliest, though most students start after grade 10 or even later. Teenagers who intend to go to college must receive their university entrance qualification before enrolling. Hence, starting college is possible in 2014 at the very earliest. For each individual we consider a five-year time window after leaving secondary school as indicated by the blue bars. Potential differences in the exact timing of the five-year window are indicated by the arrows.

2.3 Development of a score of ICT skills

To summarize the large set of up to 32 items of the competence tests, we construct a weighted score of each respondent's digital skills. In doing so, we scale all items k_f between 0 and 1 and add the number of correctly answered items relative to the total number of items queried:

$$ICTskills_i = \frac{1}{F} \sum_{f=1}^{F} k_f \tag{1}$$

Items that remained unanswered or received an incorrect response receive a value of 0.[2] The resulting score, $ICTskills_i$, lies between 0 to 1. A respondent's score of, for instance, 0.5 indicates that half of the queried items had been answered correctly. One correctly answered item increases the ICT skills by 0.026 in ninth grade and by 0.032 in 12th grade.

---

[2]The majority of respondents answered all questions. Please refer to Bachmann et al. (2021) for details on the items and the success rate.

Table 1: Descriptive statistics of the sample

Statistic	Overall		Male		Female	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
ICT skills in 9th grade	0.569	0.174	0.565	0.179	0.573	0.169
Female	0.501	0.500	–	–	–	–
Migration Background	0.219	0.414	0.210	0.407	0.229	0.420
Parent(s) in STEM Occupation	0.337	0.473	0.347	0.476	0.326	0.469
Cognitive Skills	0.499	0.500	0.564	0.496	0.435	0.496
Choose STEM: longest training	0.326	0.469	0.513	0.500	0.140	0.347
Choose STEM: first training	0.337	0.473	0.526	0.499	0.149	0.356
Choose STEM: last training	0.317	0.465	0.501	0.500	0.134	0.341

The table provides the mean and standard deviation of variables of interest. The overall sample consists of 9,315 teenagers for whom all variables are available. The subsamples contain 4,747 female and 4,668 male teenagers. The overall sample is slightly positive selected - please see Table 4 in Appendix A for a comparison of the (representative) full data and the sample.

2.4 Sample selection

We restrict our analysis to teenagers who were tested in ninth grade and participated in subsequent waves of the panel survey in which educational and occupational choices were surveyed. Moreover, we exclud all students from our analysis of which data on parents or the student's gender or migration background were missing. The relevant information is available for 57 % of the data.

Our sample spans data on 9,315 students whose ICT skills were tested in ninth grade in 2010. Table 1 provides summary statistics of the sample. Among the 9,315 students in the sample, 50.1 % are female and 21.9 % have a migration background. Around one third of the adolescents (33.7 %) have at least one parent who works in a STEM occupation.[3]

The average *ICT skills* is 0.569, meaning that the average respondent correctly answered 22 out of 39 questions (56.9 %). Female teenagers have on average slightly higher ICT skills than their male fellows. Table 1 already reveals some stereotypical career choices: More than half of the male teenagers choose a career in STEM while less than 15 % of the female teenagers go into STEM after completing secondary school.

Adolescents can start and drop out of their educational choices. We therefore include three different possibilities to measure whether a student chooses a STEM occupation after graduating from high school. Within five years after completing high school, we can distinguish between the *longest* training period, the *first* training period, and the *last* training period.[4] Table 1 shows that one third of the adolescents (33.7 %) choose a STEM occupation directly after completing high school (*Choose STEM: first training*). The table also shows that slightly fewer students have STEM as their longest (32.6 %) or last training period (31.7 %).

---

[3]NEPS SC4 is representative of German high school students. Table 4 in Appendix A shows that our sample is slightly positive selected in terms of ICT skills and overall cognitive abilities. The sample comprises fewer students with a migration background but more students whose parents work in STEM - compared to the full NEPS data. We will discuss in the results section how the slightly positive selection may affect the results.

[4]The first training period is the training period that starts directly upon graduating from secondary school. The last training period is the training period that is closest to the end of the five-year time window of interest. The longest training period is the training period with the longest duration, i.e., the training period that usually has been completed successfully.

3 Empirical analysis

We are interested in understanding the interplay between teenagers' digital skills and gender-specific selection into STEM fields upon graduation from secondary school. The study's objective is to address the following research questions:

1. Do teenagers' ICT skills correlate with their subsequent career choices and, if yes, to what extent? (RQ1)

2. Do differences exist in ICT skills and career choices of female and male teenagers? (RQ2)

3. Is the gender-specific selection into STEM fields mitigated by ICT skills in secondary school? (RQ3)

As we do not know a priori whether a teenager's ICT skills affect subsequent career choices at all or act as a moderator or effect modifier, we employ three different statistical models. These are illustrated in Figure 2 and are described in more detail below.

Figure 2: Illustration of statistical models

A) Direct (gender independent) relationship between ICT skills and STEM career
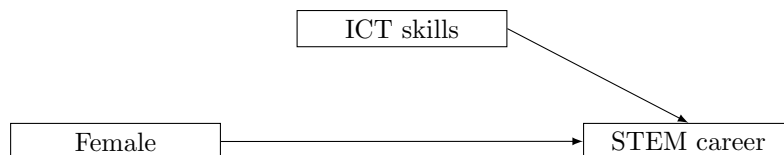
B) ICT skills as a moderator
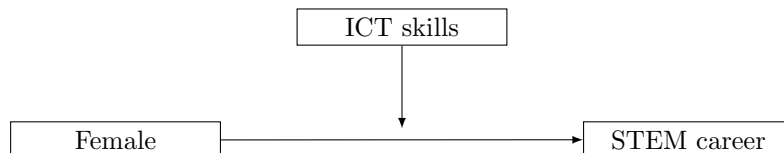
C) ICT skills as a mediator
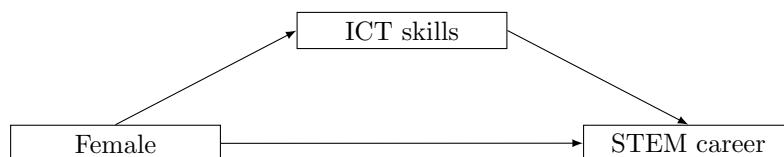
Note: Panels A, B, and C illustrate stylized versions of the three possible scenarios of how gender, ICT skills, and the decision towards a career in STEM can relate to each other. The models in panels A and B are estimated by utilizing linear probability models. The model in panel C is estimated by running a mediation analysis.

Direct relationship between ICT skills and STEM career

To answer research question RQ1, we estimate a set of linear probability models with the method of ordinary least squares that postulate independent links between gender and ICT skills (Panel A in Figure 2). Our outcome variable y_i is a binary variable that equals one if the respondent chooses a STEM major

at college or a vocational training in a STEM field. More specifically, we estimate:

$$P(y_i = 1 | X_{ij}) = \alpha + \beta ICTskills_i + \sum_{j=1}^{J} \gamma_j X_{ij} + \epsilon_i \qquad (2)$$

where $ICTskills_i$ captures the digital literacy of a respondent i as described by equation 1. X_{ij} comprises a set of j individual control variables including a respondent's gender, migration background, and whether at least one of the parents worked in a STEM occupation. Moreover, X_{ij} includes a binary proxy for an individual's overall cognitive skills. It is 0 if students have below-average test results in mathematics and 1 otherwise.

Our main specification utilizes the longest training period within five years after having completed secondary education. In a set of robustness checks, we change this definition from the longest to the first and to the last training period within five years after secondary schooling.

ICT skills as a moderator

To answer RQ2, we add interactions between the ICT skills and gender to Equation 2. The interaction term helps us understand whether the gender-specific sorting into STEM occupations is moderated by ICT skills (Panel B in Figure 2). Formally, we estimate:

$$P(y_i = 1 | X_{ij}) = \alpha + \beta_1 ICTskills_i + \beta_2 ICTskills_i \cdot gender_i + \sum_{j=1}^{J} \gamma_j X_{ij} + \epsilon_i \qquad (3)$$

By including the interaction term between ICT skills and gender as formalized in Equation 3, we can differentiate whether ICT skills directly affect the probability to choose a STEM career (provided by β_1) or whether gender-specific differences in these skills affect the relationship (provided by β_2).

Mediation analysis: Understanding a potential indirect effect

In a third set of regressions, we address RQ3 and aim at understanding whether the gender-specific selection into STEM fields is mitigated by ICT skills. We further explore the triangle between gender, ICT skills, and career choices by employing mediation analysis (Panel C in Figure 2). Mediation analysis can help unravel the underlying causal mechanism through which an explanatory variable affects an outcome into an effect running through changes of an intermediate variable and through other pathways. In our setting, as illustrated in Panel C in Figure 2, we are interested in the mediating effect of ICT skills on the gender-specific selection of STEM fields. The goal of the analysis in Panel C is to disentangle the total effect of gender on subsequent career paths into a direct effect and an indirect effect via a student's ICT skills. Thereby, the potential mediator – ICT skills – can help us clarify the mechanisms underlying the relationship between teenagers' ICT skills, gender, and career choices.

4 Results

4.1 Direct relationships and ICT skills as a moderator

Table 2 provides the regression results from running linear probability models as formalized in Equations 2 and 3. The dependent variable is always a binary variable that equals 1 if a teenager chooses college or vocational training in a STEM field after secondary schooling, and 0 otherwise. Columns 1a and 1b of Table 2 show that those teenagers rather choose a career in STEM that have high ICT skills (column 1a) or are male (column 1b). The coefficient for ICT skills in column 1a can be interpreted as follows. If a teenager's ICT skills increase by 10 percentage points (i.e., correctly answering an additional of three to four questions in the ICT skills test), the probability to choose a STEM field increases by 1.49 percentage points or 4.57 %.

In an additional set of regressions, we run the set of linear probability models outlined in Equations 2 and 3 on the subsample of teenagers who stayed in high school and participated in another test on ICT skills in 12th grade. The subsample comprises 2,789 students. Compared to our baseline sample that has been surveyed in ninth grade we now have more female students and fewer students with a migrant background. The complete descriptive statistics are in Table 5 in Appendix A.

Direct relationships

In columns 1c and 2a, we provide the results of jointly including ICT skills and gender in the regression as provided by Equation 2. A higher digital literacy as well as being male is associated with an increase in the likelihood of starting a career in STEM. Both coefficients are highly statistically significant. Interestingly, the magnitude of these coefficients is similar to the raw correlation between the variables and the outcome as provided by columns 1a and 1b.

In columns 1c and 2a, two further relationships emerge: First, we find no differences in choosing a STEM career among students of different migration backgrounds. Second, teenagers who have at least one parent who works in a STEM field are more likely to choose a STEM career themselves.

ICT skills as a moderator

In columns 1d and 2b we add the interaction of a teenager's gender and ICT skills to the regression (as provided by Equation 3). The results are striking for teenagers in 9th grade: While females rather do not choose a career in STEM, those girls with higher digital competencies are much more likely to choose either a vocational training or an undergraduate college course in STEM. The coefficient of the interaction in 9th grade is 0.258. In combination with the coefficients for ICT skills (0.037 but lacking statistical significance) and female (-0.519, highly statistically significant), the interpretation is that an increase in girls' ICT skills by 10 percentage points is associated with an increase in the probability to choose a STEM career by 24.99 %. The direction of the relationship of interest holds even if we control for overall cognitive skills in column 1e and when interacting cognitive skills with the variable for gender in column 1f. Correctly solving an additional three to four questions is still associated with an increase in the likelihood of choosing a STEM career by 21.71 %.

For teenagers in grade 12 the relationship is different: While the direct relationship of ICT skills and gender remains high in magnitude and strong in statistical significance, the interaction of the two variables lacks economic and statistical significance. That means that the positive link between ICT skills and the choice of a career in STEM holds for female and male teenagers and no longer for female teenagers only. We explore this finding further in the next section.

Table 2: ICT skills in 9th and 12th grade and career paths: Direct relationship and moderating effect

Dependent variable:
Respondent chooses STEM occupation after high school

	ICT skills in 9th grade						**ICT skills in 12th grade**			
	(1a)	(1b)	(1c)	(1d)	(1e)	(1f)	(2a)	(2b)	(2c)	(2d)
ICT skills	0.149***	-	0.158***	0.037	−0.046	−0.016	0.305***	0.308***	0.205*	0.154
	(0.028)	-	(0.026)	(0.035)	(0.038)	(0.041)	(0.067)	(0.094)	(0.094)	(0.096)
Female	-	−0.373***	−0.372***	−0.519***	−0.509***	−0.496***	−0.264***	−0.259**	−0.191*	−0.184*
	-	(0.009)	(0.009)	(0.030)	(0.030)	(0.031)	(0.017)	(0.079)	(0.079)	(0.079)
ICT skills * Female	-	-	-	0.258***	0.256***	0.194***	-	−0.008	−0.077	0.024
	-	-	-	(0.051)	(0.051)	(0.059)	-	(0.132)	(0.131)	(0.137)
Migration Background	-	-	−0.014	−0.012	−0.008	−0.008	−0.022	−0.022	−0.010	−0.010
	-	-	(0.011)	(0.011)	(0.011)	(0.011)	(0.023)	(0.023)	(0.023)	(0.023)
Parent(s) in STEM	-	-	0.076***	0.076***	0.073***	0.074***	0.075***	0.075***	0.070***	0.069***
	-	-	(0.009)	(0.009)	(0.009)	(0.009)	(0.018)	(0.018)	(0.018)	(0.018)
Cognitive Skills	-	-	-	-	0.060***	0.038***	-	-	0.131***	0.193***
	-	-	-	-	(0.010)	(0.015)	-	-	(0.019)	(0.031)
Cognitive Skills * Female	-	-	-	-	-	0.043**	-	-	-	−0.098*
	-	-	-	-	-	(0.021)	-	-	-	(0.039)
Constant	0.241***	0.513***	0.400***	0.468***	0.481***	0.476***	0.282***	0.280***	0.242***	0.225***
	(0.017)	(0.006)	(0.017)	(0.021)	(0.022)	(0.022)	(0.043)	(0.058)	(0.058)	(0.058)
Observations	9,315	9,315	9,315	9,315	9,315	9,315	2,789	2,789	2,789	2,789
R²	0.003	0.158	0.168	0.171	0.174	0.174	0.099	0.099	0.114	0.116
Adjusted R²	0.003	0.158	0.168	0.170	0.173	0.173	0.098	0.097	0.112	0.114

Note: The dependent variable is a binary variable that equals 1 if the respondents' *longest* training period within five years after completing secondary schooling was within a STEM field. The ICT skills relate to the tests conducted in 9th grade (columns 1a to 1f) and 12th grade (columns 2a to 2d). Robust standard errors are clustered on the levels of schools. Significance: *p<0.1; **p<0.05; ***p<0.01.

8

When accounting for overall cognitive skills, the isolated effect of the digital skills slightly decreases (column 2c) and lacks statistical significance (column 2d), while at the same time the cognitive skills seem to be more relevant for subsequent educational choices than in ninth grade. Selection into more complex courses in high school as determined by Card and Payne (2021) may be a driver of the link between skills and the choice of a STEM career.

Overall, the results in Table 2 show that ICT skills rather act as a moderator in the career choice of girls in ninth grade while ICT skills and a student's gender rather independently affects career paths after 12th grade.

Type of STEM training

Given these results, we dig deeper into the types of training in a second set of regressions on the subsample of students in grade 9.[5] After having completed secondary schooling, students most commonly either start with an apprenticeship (vocational training) or go to college. We run our regressions from Equations 2 and 3 separately on different subsamples depending on gender and the training track. Table 3 presents the results.

Column 1a provides the baseline results (as in column 1f in Table 2). Columns 1b and 1c provide the results after running the pooled regressions separately for female (1b) and male students (1c). As expected, the association between female teenagers' ICT skills and choosing a STEM career is large in magnitude and highly statistically significant – while there is no clear pattern for male students.

Columns 1a, 2a, and 3a of Table 3 provide that female students are less likely to choose a STEM career. Furthermore, it appears that all students are affected by their parents' occupations. The relationship is stronger for male than for female students though: Male teenagers are more likely to choose any STEM career if at least one parent was working in STEM (columns 1c, 2c, and 3c).

Our main variable of interest is a respondent's ICT skills. Looking at the relationship between the ICT skills and the probability to choose a STEM career, Table 2 shows that the strong association from our baseline regression disappears when conditioning on college education. All our results from column 1f in Table 2 seem to be driven by teenagers who start vocational training after completing high school. It is worth noting that the coefficient is large in magnitude and strong in statistical significance despite the small subset of only 2,802 female students who decide for a vocational training.

Summing this up, the preceding analysis suggests that teenagers' ICT skills correlate with their subsequent career choices (RQ 1): Adolescents with higher ICT skills during secondary education are more likely to start a career in STEM. We also show that female teenagers select out of STEM careers. Regarding the interrelationship between a teenager's gender, ICT skills, and choice of STEM education, we can conclude that ICT skills act as a moderator for those who start a vocational training: Female teenagers are more likely to start an apprenticeship in a STEM field if they already have high ICT skills in 9th grade. Such a relationship neither exists for studying a STEM field at college nor for teenagers' ICT skills in 12th grade (RQ2).

4.2 Mediation Analysis

In the preceding results, we have already identified a joint action of gender and ICT skills. In this section, we now turn to the results of a mediation analysis (RQ 3). We used the *mediation* package provided by Tingley, Yamamoto, Hirose, Keele, and Imai (2014) for R to conduct the mediation analysis. The results summarised in Figure 3 reveal that the gender-specific selection into STEM fields is not mediated via the teenagers' ICT skills.

---

[5]Running the same analysis for students in grade 12 lacks statistical power due to the small sample sizes.
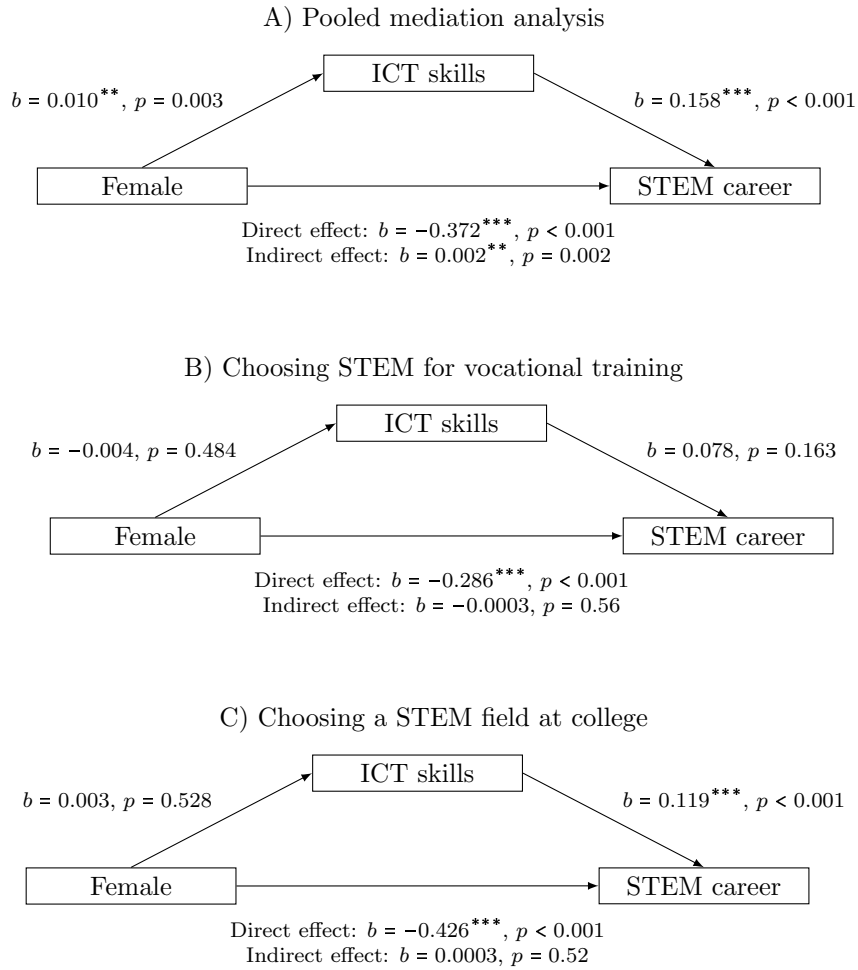
Table 3: Heterogeneity across type of STEM training

Dependent variable:
Respondent chooses STEM occupation after high school

	Pooled			Vocational training			College		
	Pooled (1a)	Female (1b)	Male (1c)	Pooled (2a)	Female (2b)	Male (2c)	Pooled (3a)	Female (3b)	Male (3c)
ICT skills in 9th grade	−0.016 (0.041)	0.187*** (0.034)	−0.020 (0.048)	−0.015 (0.048)	0.180*** (0.037)	−0.020 (0.059)	0.055 (0.084)	0.002 (0.075)	0.056 (0.093)
ICT skills * Female	0.194*** (0.059)	-	-	0.181** (0.071)	-	-	−0.050 (0.117)	-	-
Female	−0.496*** (0.031)	-	-	−0.512*** (0.036)	-	-	−0.247*** (0.075)	-	-
Migration background	−0.008 (0.011)	0.004 (0.012)	−0.022 (0.018)	−0.024* (0.013)	−0.009 (0.013)	−0.039* (0.022)	0.019 (0.020)	0.014 (0.025)	0.025 (0.033)
Parent(s) in STEM occupation	0.074*** (0.009)	0.051*** (0.011)	0.096*** (0.015)	0.071*** (0.011)	0.053*** (0.011)	0.086*** (0.019)	0.070*** (0.017)	0.051** (0.021)	0.093*** (0.028)
Cognitive skills	0.038*** (0.015)	0.083*** (0.012)	0.036** (0.017)	0.047*** (0.017)	0.052*** (0.013)	0.045** (0.020)	0.054 (0.036)	0.057*** (0.024)	0.052 (0.040)
Cognitive skills * Female	0.043** (0.021)	-	-	0.004 (0.026)	-	-	0.002 (0.044)	-	-
Constant	0.476*** (0.022)	−0.021 (0.020)	0.475*** (0.026)	0.486*** (0.025)	−0.031 (0.020)	0.488*** (0.031)	0.400*** (0.056)	0.161*** (0.050)	0.391*** (0.064)
Observations	9,315	4,668	4,647	5,895	2,802	3,093	3,197	1,759	1,438
R^2	0.174	0.039	0.010	0.225	0.033	0.011	0.097	0.007	0.011
Adjusted R^2	0.173	0.038	0.010	0.224	0.032	0.009	0.095	0.005	0.008

Note: The dependent variable is a binary variable that equals 1 if the respondents' longest training period within five years after completing secondary schooling was within a STEM field. Robust standard errors are clustered on the levels of schools. The data of 107 female and 216 male respondents has been dropped in columns (2) and (3) because we could not clearly define the training as vocational or college studies. Significance: * p<0.1; ** p<0.05; *** p<0.01

The results in Figure 3 imply that the gender-specific sorting into STEM fields is not caused by gender-specific development of ICT skills. We find a strong negative direct effect of being female on the choice of a STEM field ($b = -0.286$ for pooled analysis in panel A, $b = -0.286$ for vocational training in panel B, $b = -0.426$ for college in panel C) and a positive though not always statistically significant effect of having larger ICT skills on choosing a STEM field ($b = 0.158$ for pooled, $b = 0.078$ for vocational training, $b = 0.119$ for college). The indirect effect is negligibly small and even lacks statistical significance for the subsamples that include only those teenagers who start vocational training (Panel B) or who go to college (Panel C). Overall, the mediation analysis shows that there is no causal pathway along which a teenager's gender causes their ICT skills (the mediator) and the ICT skills cause the choice of a career in STEM.

Figure 3: Results from mediation analysis

A) Pooled mediation analysis

$b = 0.010^{**}$, $p = 0.003$ ICT skills $b = 0.158^{***}$, $p < 0.001$

Female STEM career

Direct effect: $b = -0.372^{***}$, $p < 0.001$
Indirect effect: $b = 0.002^{**}$, $p = 0.002$

B) Choosing STEM for vocational training

$b = -0.004$, $p = 0.484$ ICT skills $b = 0.078$, $p = 0.163$

Female STEM career

Direct effect: $b = -0.286^{***}$, $p < 0.001$
Indirect effect: $b = -0.0003$, $p = 0.56$

C) Choosing a STEM field at college

$b = 0.003$, $p = 0.528$ ICT skills $b = 0.119^{***}$, $p < 0.001$

Female STEM career

Direct effect: $b = -0.426^{***}$, $p < 0.001$
Indirect effect: $b = 0.0003$, $p = 0.52$

The figure illustrates the results of running the mediation analysis. Additional covariates are teenagers' migrant background and whether at least one of their parents works in a STEM occupation (as in column 1d of Table 2). Standard errors of the indirect effect are bootstrapped with 1,000 replications. Sample sizes are 9,315 for the pooled mediation (Panel A), 3,197 for the subsample of teenagers who start a vocational training (Panel B), and 5,895 for the subsample of college students (Panel C). Detailed regression tables are available on request.

5   Discussion

The results provide insights into the relationship between teenagers' gender, ICT skills, and career choices. First, higher ICT skills do not per se translate into a career in STEM. We find, however, very strong

positive associations between girls' ICT skills in ninth grade and their decisions to start a vocational training in a STEM field.

Interestingly, ICT skills act as a moderator and not as a mediator in the triangle of interest. One explanation for this finding comes from psychological research: Tracey and Hopkins (2001) have shown that young people are interested in occupations that imply tasks they believe to be good at while Correll (2001), Lubienski et al. (2013), and Jouini et al. (2018) show that female teenagers tend to underestimate their true potential. Combining these two strands may imply that female teenagers' possible lack of confidence in being able to develop strong ICT skills may hinder them from starting a career in STEM.

Another relevant finding is that higher ICT skills close to receiving a university entrance qualification come with a stronger selection into studying STEM. This relationship as presented in Table 2 holds for female and male teenagers. One of the underlying reasons may be that students in high schools have to specialize in some fields and the overall selection into or out of STEM at least partly already takes place in upper secondary schooling as shown by Card and Payne (2021). When students choose more STEM subjects they are likely to increase their ICT skills at school by attending more ICT-related courses (e.g., in Physics) and also outside school because they are more interested in improving their skills.

Changes in the definition of the training period

So far, we have concentrated on the longest training period within five years after completing secondary schooling. However, students may change their desired career within the first years after graduating from high school.

The aforementioned results as presented in Table 2 use the *longest* training period as the outcome variable. We believe that the longest training period best captures the long-term career choice. To understand whether different training periods affect the results, we run our baseline regression from equation 2 but use the *first* training period after completing high school as well as the *last* training period within the five years after high school instead of the longest training period.

Table 6 provides the results of the alternative specification of the outcome variable. Column (1) provides the baseline specification (as in column 1f in Table 2). Overall, the results are very similar, indicating – as already suggested by Table 1 – that students are very consistent in choosing either STEM or non-STEM even if they change the exact subfield or switch their training provider.

An interesting difference between the three specifications in Table 6 is, however, the decrease in importance of female teenagers' ICT skills when considering the last training period (column 3) while at the same time, the cognitive skills increase in their importance. The ad-hoc selection into STEM appears to be (partly) driven by teenagers' ICT skills.

Overall, it appears that the results are very robust to considering the different timing of the training. The ICT skills remain highly correlated with the decision towards a career in STEM – but for female students only.

Selectivity of sample

NEPS data is representative of German high school students. However, not all students participated in the ICT skills test. Table 4 in Appendix A compares our sample to the full data and shows that our sample is slightly positive selected in terms of ICT skills, overall cognitive abilities, and parents working in STEM fields. In turn, the sample comprise fewer students from migrant families than the full data.

Overall, the sample is slightly positively selected and is thereby no longer fully representative of German teenagers: Adolescents in the sample live in households with above-average parental job engagement in STEM and below-average migrant backgrounds. We believe that the students in the sample have a

higher than population-average probability to choose a career in STEM because their parents can serve as role models. Thereby, we may underestimate the true (population-wide) relationship between teenagers' digital skills and their subsequent choice of a STEM occupation.

6   Conclusion

The present study aims at understanding whether the level of teenagers' ICT skills affects their career choices and whether the gender-specific sorting into STEM fields can be explained by differences in ICT skills. The study is based on data from the National Educational Panel Study (NEPS) that comprises detailed data about individual educational trajectories and also provides test-based results on teenagers' ICT skills. We use a sample of teenagers who were in ninth grade in the school year 2010/11 and were repeatedly asked about their educational choices and pathways afterwards. Tests on students' ICT skills were conducted in 2010/2011 and again in 2013/14. We look at the transition into vocational training or higher education after high school.

While we find – in line with recent evidence by Gnambs (2021) and Bachmann et al. (2022) – no significant differences between girls' and boys' ICT skills in ninth grade, the subsequent trajectories of female and male teenagers are different: Only those girls with strong ICT skills select into an apprenticeship in STEM. A student's ICT skills act as a moderator in the the gender-specific selection into STEM. The relationship between an indvidual's ICT skills and subsequent career choices is, however, different among students in their final years of high school: For students in 12th grade, higher ICT skills (and higher cognitive skills) directly affect the choice of a STEM field while we could neither find a moderating nor a mediating effect of ICT skills on the choice of a STEM field.

Our results complement recent evidence by Card and Payne (2021) who show that the persistent gender gap in STEM is partly caused by some students' lack of engagement in complex STEM fields in high school. The present study indicates that this behaviour is at least partly driven by ICT skills and that girls need above average ICT skills to sort into STEM fields upon graduating from high school. We explain our findings with evidence from research in psychology that shows that teenagers sort into occupations they believe to be good at (Tracey & Hopkins, 2001) and that female teenagers rather underestimate their true potential (Correll, 2001; Lubienski et al., 2013; Jouini et al., 2018).

The shortage in the supply of STEM workers is a major concern in many countries and industries (Cedefop, 2018). Given the persistent gender-specific selection into STEM fields, our study contributes to a set of studies that aim at understanding the lack of female participation in STEM occupations. While research has shown that, for instance, teachers and role models can reduce stereotypical sorting into jobs (Kahn & Ginther, 2017), the present study shows that a (perceived) lack of ICT skills may hinder girls from choosing a STEM field after high school. The findings of this study add to the ongoing policy debate on increasing the intensity of ICT training in secondary school (Vegas et al., 2021; KMK, 2016). Thereby, the findings of the present study have strong policy implications: Promoting ICT skills and strengthening confidence in ICT, especially among female teenagers during their secondary schooling, can increase the number of teenagers starting a career in STEM and thereby combat the persistent gender gap in STEM.

References

Bachmann, R., Hertweck, F., Kamb, R., Lehner, J., & Niederstadt, M. (2022). Digitale Kompetenzen in Deutschland. *Zeitschrift für Wirtschaftspolitik*, *71*(3), 266-286. doi: 10.1515/zfwp-2022-2082

Bachmann, R., Hertweck, F., Kamb, R., Lehner, J., Niederstadt, M., & Rulff, C. (2021). Digitale Kompetenzen in Deutschland – eine Bestandsaufnahme. *RWI Materialien No. 150*. doi: 10419/249684

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining Twenty-First Century Skills. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and Teaching of 21st Century Skills* (pp. 17–66). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-007-2324-5_2

Calvani, A., Fini, A., Ranieri, M., & Picci, P. (2012). Are young generations in secondary school digitally competent? A study on Italian teenagers. *Computers & Education*, *58*(2), 797-807. doi: 10.1016/j.compedu.2011.10.004

Card, D., & Payne, A. (2021). High school choices and the gender gap in STEM. *Economic Inquiry*, *59*(1), 9-28. doi: 10.1111/ecin.12934

Cedefop. (2018). Insights into skill shortages and skill mismatch: Learning from Cedefop's European skills and jobs survey. *Cedefop reference series*, *106*.

Chetty, K., Qigui, L., Gcora, N., Josie, J., Wenwei, L., & Fang, C. (2018). Bridging the digital divide: Measuring digital literacy. *Economics*, *12*(1), 20180023. doi: 10.5018/economics-ejournal.ja.2018-23

Correll, S. J. (2001). Gender and the Career Choice Process: The Role of Biased Self-assessments. *International Reading Association*, *106*(6), 1691-1730. doi: 10.1086/321299

Eshet-Alkalai, Y. (2004). Digital Literacy: A Conceptual Framework for Survival Skills in the Digital era. *Journal of Educational Multimedia and Hypermedia*, *13*(1), 93–106.

Gilster, P. (1997). *Digital literacy.* John Wiley & Sons.

Gnambs, T. (2021). The development of gender differences in information and communication technology (ICT) literacy in middle adolescence. *Computers in Human Behavior*, *114*, 106533. doi: 10.1016/j.chb.2020.106533

Goldhammer, F., Naumann, J., & Keßel, Y. (2013). Assessing Individual Differences in Basic Computer Skills. *European Journal of Psychological Assessment*, *29*(4), 263-275. doi: 10.1027/1015-5759/a000153

Hatlevik, O. E., & Christophersen, K.-A. (2013). Digital competence at the beginning of upper secondary school: Identifying factors explaining digital inclusion. *Computers & Education*, *63*, 240–247. doi: 10.1016/j.compedu.2012.11.015

International ICT Literacy Panel. (2002). Digital transformation. A framework for ICT literacy. *Educational Testing Service*.

Jouini, E., Karehnke, P., & Napp, C. (2018). Stereotypes, underconfidence and decision-making with an application to gender and math. *Journal of Economic Behavior & Organization*, *148*, 34–45. doi: 10.1016/j.jebo.2018.02.002

Kahn, S., & Ginther, D. (2017). Women and STEM. *NBER Working Paper 23525*. doi: 10.3386/w23525

KMK. (2016). Strategie „Bildung in der digitalen Welt" (Beschluss der Kultusministerkonferenz vom 08.12.2016 in der Fassung vom 07.12.2017). *Kultusministerkonferenz (KMK) Berlin/Bonn*.

Leu, D. J. J., Kinzer, C. K., Coiro, J., & Cammack, D. W. (2004). Toward a Theory of New Literacies Emerging From the Internet and Other Information and Communication Technologies. *Theoretical models and processes of reading*, *5*(1), 1570–1613.

Lubienski, S. T., Robinson, J. P., Crane, C. C., & Ganley, C. M. (2013). Girls' and Boys' Mathematics Achievement,Affect,and Experiences: Findings From ECLS-K. *Journal for Research in Mathematics Education*, *44*(4), 634–645. doi: 10.5951/jresematheduc.44.4.0634

NEPS-Netzwerk. (2021). Nationales Bildungspanel, Scientific Use File of Starting Cohort Grade Nine. doi: 10.5157/NEPS:SC4:12.0.0

Senkbeil, M., & Ihme, J. M. (2020). Diagnostik von ICT Literacy: Messen Multiple-Choice-Aufgaben und simulationsbasierte Aufgaben vergleichbare Konstrukte? *Diagnostica*, *66*, 147-157. doi: 10.1026/0012-1924/a000243

Senkbeil, M., Ihme, J. M., & Wittwer, J. (2013). Entwicklung und erste Validierung eines Tests zur Erfassung technologischer und informationsbezogener Literacy (TILT) für Jugendliche am Ende der Sekundarstufe I. *Zeitschrift für Erziehungswissenschaft*, *16*, 671–691.

Senkbeil, M., Ihme, J. M., & Wittwer, J. (2021). The NEPS ICT Literacy Framework and Item Examples. *NEPS Research Data*.

Sparks, J. R., Katz, I. R., & Beile, P. M. (2016). Assessing Digital Information Literacy in Higher Education: A Review of Existing Frameworks and Assessments With Recommendations for Next-Generation Assessment. *ETS Research Report Series*.

Steedman, H., Wagner, K., & Foreman, J. (2006). ICT skill supply in the UK and Germany. *European journal of vocational training*, *37*(1), 9–21.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software*, *59*.

Tracey, T. J. G., & Hopkins, N. (2001). Correspondence of interests and abilities with occupational choice. *Journal of counseling psychology*, *48*(2), 178–189.

van Deursen, A. J. A. M., Helsper, E. J., & Eynon, R. (2016). Development and validation of the Internet Skills Scale (ISS). *Information, Communication & Society*, *19*(6), 804-823. doi: 10.1080/1369118X.2015.1078834

van Deursen, A. J. A. M., & van Diepen, S. (2013). Information and strategic internet skills of secondary students: A performance test. *Computers & Education*, *63*, 218-226. doi: 10.1016/j.compedu.2012.12.007

Van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, *72*, 577-588. doi: 10.1016/j.chb.2017.03.010

Vegas, E., Hansen, M., & Fowler, B. (2021). Building skills for life: How to expand and improve computer science education around the world. *Report by Center for Universal Education at Brookings*.

A Additional tables

Table 4: Comparison of sample with full data

	Sample		Full data		Comparison	
	N	Mean	N	Mean	Diff.	p-value
ICT skills in 9th grade	9,315	0.569	14,486	0.552	0.017	0.000
Female	9,315	0.501	16,319	0.495	0.006	0.318
Migration Background	9,315	0.219	16,159	0.263	-0.044	0.000
Parent in STEM Occupation	9,315	0.337	14,912	0.305	0.032	0.000
Cognitive Skills	9,315	0.499	14,523	0.459	0.040	0.000
Choose STEM: longest training	9,315	0.326	11,016	0.320	0.006	0.381
Choose STEM: first training	9,315	0.337	11,001	0.331	0.006	0.355
Choose STEM: last training	9,315	0.317	10,966	0.313	0.005	0.488

The table summarizes the differences between the means of the variables of interest in the full data and the sample we selected. We restricted the sample to students for which data on competence tests, on socio-demografic characteristics, and their educational decisions is available.

Table 5: Descriptive statistics of students tested in grade 12

Statistic	N	Mean
ICT skills in 12th grade	2,789	0.299
Female	2,789	0.564
Migration Background	2,789	0.170
Parent in STEM Occupation	2,789	0.343
Cognitive Skills	2,789	0.619
Choose STEM: longest training	2,789	0.332
Choose STEM: first training	2,789	0.345
Choose STEM: last training	2,789	0.320

The table provides the mean of variables of interest. These are based on the sample of 2,789 students who participated in a second test on digital skills when attending 12th grade of high school.

Table 6: Alternative definitions of outcome in main regression

| | Dependent variable: | | |
| | Respondent chooses STEM occupation after high school | | |
	Longest training (1)	First training (2)	Last training (3)
ICT skills	−0.016	0.009	−0.011
	(0.041)	(0.041)	(0.041)
ICT skills * Female	0.194***	0.176***	0.165***
	(0.059)	(0.059)	(0.059)
Female	−0.496***	−0.483***	−0.481***
	(0.031)	(0.031)	(0.031)
Migration Background	−0.008	−0.002	−0.005
	(0.011)	(0.011)	(0.011)
Parent(s) in STEM occupation	0.074***	0.074***	0.071***
	(0.009)	(0.009)	(0.009)
Cognitive Skills	0.038***	0.042***	0.029*
	(0.015)	(0.015)	(0.015)
Cognitive Skills * Female	0.043**	0.027	0.056***
	(0.021)	(0.021)	(0.021)
Observations	9,315	9,315	9,315
R^2	0.174	0.174	0.171
Adjusted R^2	0.173	0.173	0.170

Note: The dependent variable is a binary variable that equals 1 if the respondents' *longest* training period within five years after completing secondary schooling was within a STEM field (column 1), if the respondents' *first* training period after completing secondary schooling was within a STEM field (column 2) or if the respondents' *last* training period within five years after completing secondary schooling was within a STEM field (column 3). All regressions also include a constant. Robust standard errors are clustered on the levels of schools. Significance: *p<0.1; **p<0.05; ***p<0.01