Gender differences in the labor market entry of STEM graduates: Does fertility play a role?

Work in Progress

Jakob Schwerter*, Lena Ilg

June 13, 2019

Abstract

Females do not work in science, technology, engineering, and mathematics (STEM) occupation, even though they have a degree in those subjects of study. Unlike most other studies on the topic, we include both genders as well as STEM and non-STEM graduates in order to compare, again, unlike most other studies, the non-entry behavior of male and female STEM graduates relative to all other degree fields. Random effects regressions show that female STEM graduates, in general, have a lower field of degree transition advantage to STEM occupations during the first five years of their career compared to their male counterparts. This effect is most pronounced for engineering and computer science (EngComp) graduates. Fertility, in general, did not prove to contribute significantly to this relative higher female rate of non-transition to STEM or EngComp occupations, but the information if children were born before the graduation shows to be an important control variable. Mediator analysis can explain only some part of the gender difference.

JEL Codes: I23, J13, J16

Keywords: job mismatch, university-to-work transition, non-entry rate

*University of Tübingen, LEAD Tübingen, jakob.schwerter@uni-tuebingen.de

1 Introduction

The under-representation of women in sciences, technology, engineering, and mathematics (STEM) fields of higher education programs and occupations has received considerable attention in the scientific literature but also became more predominant in the public debate for several years. Almost all of the STEM occupations continue to be dominated by men while women are observed not to enter STEM occupations even though they have graduated with a degree in STEM. Recent numbers illustrate this dilemma: In 2016, women accounted for some 28 percent of STEM graduates in Germany. At the same time, only around 19 percent of the STEM workforce was female. In contrast, women represented nearly half of all university graduates (48.5%) as well as almost half of the entire workforce (46%) in Germany in 2016 (BA, 2018). Hence, women are still underrepresented in the STEM workforce and seem to have troubles entering the STEM workforce.

At the same time, the contribution of women to STEM occupations is univocally considered as crucial to the innovative power and the continuous development of the STEM sector. Due to aging and a male-dominated workforce, and in order to meet future demands, the industry needs female workers in STEM (Burke, 2007). Scholars also stress that women (i) bring in new thinking styles and different approaches to problem-solving which can be a boost to any company (Simard et al., 2013) and (ii) the presence of women increase the quality of goods and services produced and that ultimately increases innovations (Burke, 2007).

Since the STEM sector is crucial for a country's economic success, the issue of women not entering STEM occupations is not only important from a gender equality perspective but can also have a substantial impact on the economic performance of companies because of second-tier men take the place of women who would have been better prepared but leaving STEM (Justman and Mendez, 2018). The underlying reasons for the underrepresentation of women in this area remain mostly unknown. It is, however, well-documented that STEM occupations are still overly hostile to female workers (e.g. Simard et al., 2013; Danbold and Huo, 2017).

The importance of the issue is also reflected in a large number of studies in the empirical literature. A considerable part of these studies consists of exit studies that focus on the retention of women who already are in the STEM workforce (e.g. Preston, 1994, 2004; Morgan, 2000; Kahn and Ginther, 2015). Univocally, all these studies find higher exit rates for women compared to men within STEM (apart from Hunt (2016), who also uses non-STEM graduates and occupations). Those exit-studies do not distinguish between a missing entry or an exit during the career. Sassler et al. (2017) are notable exceptions of studies which examine the gender differences in the transition to first jobs. They highlight, however, only differences in transition rates between the STEM subfields (Sassler et al., 2017).

The higher rates of non-transition in those studies could be, however, a mere gender effect and not occupation-specific. Therefore, we consider it to be of great importance to include male and female graduates from both STEM and non-STEM fields of study in our analysis, similar to Hunt (2016). This inclusion of both genders and STEM, as well as non-STEM individuals, enables us to show that a gender difference in transition behavior exists specific in STEM occupations. To the best of our knowledge, the transition behavior of STEM graduates to degree-related occupations has not been examined using both genders and all occupational fields.

One possible explanation of the gender difference is childcare obligations. Unfortunately, current findings are mixed. Kahn and Ginther (2015) find that women with children are less likely to transition into or stay in science and engineering careers. On the contrary, family expectations could not explain the different transition or staying rates in STEM and non-STEM occupations in the study of Preston (1994), Hunt (2016), and Sassler et al. (2017).

Using cohorts 2005 and 2009, and two waves of the Graduate Panel from the German Centre for Higher Education Research and Science Studies (DZHW), regression analysis shows a gender difference for female STEM graduates of around 4 to 5 percentage points compared to male STEM graduates and workers from all other fields. This gender difference is mainly driven by the group of engineers and computer scientists (EngComp), which is why we focus mostly on this group in the later analysis. For them, we find a slightly higher significant gender difference of around 5 to 6 percentage points. We do not find any significant evidence that the relative higher non-entry rate of women in EngComp can be attributed to fertility nor potential fertility proxied by marital status. We do find, however, that the information of born children before graduation is an important control variable. This information, whether children were born before or after university graduation, is, as far as we know, unique within this topic. Intermediate outcomes, e.g. how they value they degree and if they would have studies (the same filed) again, cannot explain the gender difference completely.

The remainder of this paper is organized as follows: Section 2 reviews the relevant literature. Section 3 describes the econometric approach before section 4 describes the dataset and first descriptive results. The results of the empirical analysis, as well as robustness checks, are presented in section 5. Section 6 concludes.

2 Literature Review

In the following, we will review what influences university graduate decisions after the completion of their degree. This helps to put our analysis and results into context, select the control variables and for the discussion of the results later on. The literature has dealt extensively with this question and has come up with different results.

Knauf and Rosowski (2009) stress the importance of private life planning on career orientation. With increasing age, as partnerships become more important for young adults, private life and family planning have a stronger and more significant influence on career and occupational choices. Analyzing data of American Bachelor graduates, Joy (2000) finds that men hold jobs that have a higher self-reported career potential than women. Further, females are much more likely to enter clerical work than any other occupation, regardless of their college major (Joy, 2000). Both results suggest that women could miss entering STEM at a higher rate than males after university. A general exit behaviour is already found empirically by Preston (1994) Preston (2004) Morgan (2000) Kahn and Ginther (2015) and Hunt (2016), while only Hunt (2016) can deny a general gender effect.

Theoretical models, as well as empirical studies, suggest that childbearing can be one possible explanation for women wind up in unrelated occupations. The human capital model of Polachek (1981) predicts that individuals who already know that they will interrupt their career and temporarily exit the labor force will choose occupations that have relatively lower skill depreciation rates since they avoid potentially high losses of income and a costly re-entry. He applies the model to explain occupational segregation of men and women and shows that the different amounts of time spent in the labor force can indeed explain great parts of the differences in professional employment between men and women. Consequentially, Perna (2004) identifies gender to be an essential factor that represents the preferences of an individual in the process of occupational choice, since women plan their career in conjunction with their plans for raising children. The time during which they cannot be an active part of the labor force represents interruptions and delays of career benefits. According to Jansen and Pascher (2013), female students anticipate the potential future problems concerning the reconciliation of work and family life and are thus underrepresented in degree-related employment. Looking at STEM students, in particular, Ivanova and Stein (2013) find that work-family balances might be one of the reasons why more women drop out of academic research in chemistry. This could mean that not current but future possible fertility plays an important role, Then we have only a female effect and no fertility effect.

Ceci and Williams (2010), as well as Wang et al. (2013), argue that women have more often both a high mathematical and a high verbal ability leading to a greater range of both STEM and non-STEM career opportunities for them to choose. Women have, therefore, more possibilities from which they can choose and thus decide whether it will be a STEM or a non-STEM occupation that can fulfill their life goals and values (Wang et al., 2013). Friedman-Sokuler and Justman (2016) confirm that differences in mathematical abilities cannot explain the gender gap in the STEM field. Instead, they highlight the role of cultural and psychological factors as well as social and economic incentives. In line are the results of Lubinski and Benbow (2006), who find that women with high math ability are less interested in pursuing a career in a math-intensive field than are their male peers. Since we look at University graduates, we, however, already have a pre-selected sample. I.e., even though females have a wider range of possibilities, they already selected themselves into, e.g., computer sciences studies.

There is strong evidence that another factor influencing the choice of a career in STEM is self-efficacy, i.e., how much an individual believes in his or her abilities to achieve goals or overcome obstacles (Enman and Lupart, 2000). Among others, Hübner et al. (2017) and Heilbronner (2013) have shown that there are gender differences concerning self-efficacy, with women reporting lower levels than men. Hence, it can be expected that women, when faced with the challenges of a STEM undergraduate degree, may not believe in their abilities to succeed in a STEM environment and discontinue their pathway in the STEM field (Heilbronner, 2013);

or even with a degree question the readiness for a STEM occupation. To check this, we use subjective ability measured after graduation as a proxy (which turns out to be unimportant for the gender difference).

Ardicianono (2004), Zafar (2013), Wiswall and Zafar (2015) and Biewen and Schwerter (2019), however, suggest that not factors like expectations or perceived abilities but preferences explain gender differences in the college major choices. The variety of factors that have been outlined above could raise the presumption that, among STEM graduates, women are especially likely to pursue a career in a non-STEM occupation. The listed results are, however, not specific for university graduates but more general. Thus, our paper contributes to the question if women, even if self-selected into STEM studies with finished degrees, still not transfer into STEM occupation.

Besides the internal factors mentioned previously, external factors such as experiences at the workplace and the availability of role models also play an essential role in the decision to pursue a career in the STEM field (Heilbronner, 2013). Studies unambiguously conclude that women still face substantial barriers and discrimination at STEM workplaces (Danbold and Huo, 2017). The prevailing perception in technological occupations is the belief that being family-oriented is not associated with professional success. This assumption is often described as a *family penalty*, where women who wish to do both, climb the career ladder and raise their children, often experience their family responsibilities as a barrier to advancement (Simard et al., 2013). The external factors are, however, not testable with our data. We will, however, use the results in the discussion of the results.

3 Econometric Model

We aim to examine whether females with a STEM or EngComp degree are less likely to enter the job market with a STEM- or EngComp-occupation or not, and if this is most pronounced for females with children. Thus, we will examine whether female STEM graduates are not transitioning into STEM occupations at a higher rate than men, relative to other professional fields. Further, we investigate if fertility is contributing to this potentially higher non-entry rate for females with a STEM degree. The regression equation is as follow:

$$y_{it} = \alpha + \gamma_1 \cdot female_i + \delta_1 \cdot STEM_i + \rho_1 \cdot STEM_i \cdot female_i + X'_{it} \cdot \beta_1 + \epsilon_{it} \tag{1}$$

where index i stands for the individual and t for the wave. The dependent variable y_{it} is a binary variable that is equal to one if an individual i holds a job unrelated to their degree and equal to zero if the job is related. The construction of the dependent variable is a critical and central issue to this analysis. Using a sample of all university graduates, simply grouping the dependent variable into STEM and non-STEM occupations does not fit since graduates from non-STEM fields are not likely to work in STEM occupations after graduation than the other way around. Consequently, the analysis has to focus on job-relatedness. This definition of a study-job match is the only way to incorporate the career paths of both STEM and non-STEM graduates into the model and to provide a comparison between those groups. However, clear definitions of job-relatedness do not exist. In order to construct a measure of job-relatedness with the available data, two possibilities seemed feasible: One option is to compare the respective degree fields with the current occupation and determine whether an occupation is related to the field of study or not. However, an inquiry at the Federal Employment Agency revealed that there are no official matches of fields of study and occupations (apart from STEM).

Consequently, we rely on a second alternative and utilize a question that is included in both waves of the survey, asking individuals to rate how closely their field of study is related to their current occupation.[1] Although this measure is subjective, the literature on horizontal job mismatch considers it to be sufficiently powerful (e.g. Fehse and Kerst, 2007). This approach is not only in line with what most studies in the overeducation literature do but has also been successfully implemented in studies on both horizontal mismatch (Robst, 2007; Verhaest et al., 2017) and persistence of STEM graduates in STEM occupations (Xu, 2013; Hunt, 2016). Initially, the survey question has a five-level response scale where 1 indicates a very close relationship between university degree field and current occupation, and 5 indicates no match at all. We use this scale to construct the binary dependent variable, and re-code ratings of 5 and 4 as an unrelated job ($y_{it} = 1$) and ratings of 1 - 3 as a related job ($y_{it} = 0$).[2]

On the right-hand side of equation (1), $female_i$ is a dummy variable equal to one if the observed individual is a woman and equal to zero if the individual is a man. The binary variable $STEM_i$ is equal to one if an individual has a degree in STEM and zero for individuals holding degrees from all other fields of study. The classification of STEM degree subjects follows the classification of the Federal Employment Agency (BA, 2018). We include the interaction of the two dummy variables in the regression equation in order to isolate the effect that is specific for female STEM graduates. The coefficient on ρ_1 hence gives the gender difference and adding δ_1 and ρ_1 thus shows the probability of working in an unrelated job specifically for female STEM graduates. A positive value of $\delta_1 + \rho_1$ would indicate female STEM graduates not entering STEM occupations excessively, while a negative value of $\delta_1 + \rho_1 < \delta_1$ would only indicate a lowered entry rate of females compared to males within STEM.

The vector of variables X includes several covariates in order to control for other factors that might affect the attrition from the STEM field. The covariates are derived from both the literature on job mismatch and STEM entry behavior. Table 1 lists the full set of covariates. We include general individual information, experience, socio-cultural, personal and educational

[1] The official wording of the survey question is: "Would you say that your higher education qualification matches your job, concerning the academic qualification (field of study). Rate on a scale from 1 *Yes, definitely* to 5 *No, not at all*."

[2] Although researchers debate whether or not it is sensible to dichotomize Likert-type scales, we see little problem in our case. The loss of information from collapsing the ratings is rather negligible since only the endpoints of the scale are labeled. At the same time, our results are much easier to understand and interpret. Including the indifference rating (3) in the related-category will give quite conservative estimates. We will run robustness checks in which the middle category is sorted to be degree-unrelated and using the Likert-type scale variable.

background, as well as job characteristics, study information, and origin information. Additional, following Hunt (2016), we include dummies for all other areas of studies.

We estimate the regression equation using a random effects model. Even though the outcome variable is binary, the interpretation of (triple) interaction terms of a non-linear Logit model is non-trivial, especially for a random effects model. A fixed effects model is not possible because most of the variables of interest are constant over time. In order to account for serial correlation of the error term, standard errors in (1) are clustered on the individual level. Further, dummies for the two survey waves are added in order to account for changes in the distribution that is not necessarily identical over time. In a robustness check, among others, we look for wave specific estimation results.

To investigate possible fertility effects, we further add a children-dummy into the equation and interact the dummy with the variables for the degree and gender and their interaction:

$$\begin{aligned} y_{it} = \alpha_2 &+ \delta_2 \cdot STEM_i + \rho_2 \cdot STEM_i \cdot female_i + \lambda \cdot STEM_i \cdot female_i \cdot children_{it} \\ &+ \gamma_2 \cdot female_i + \gamma_3 \cdot children_{it} + \gamma_4 \cdot STEM_i \cdot children_{it} \\ &+ \gamma_5 \cdot female_i \cdot children_{it} + X'_{it} \cdot \beta_2 + \xi_{it} \end{aligned} \tag{2}$$

The variable *children* is a binary variable equal to one if an individual has one or more children at time t and zero otherwise. The interaction of *STEM* and *female* still isolates the gender difference that is specific for female STEM graduates, but now only for childless women. Hence, the coefficient of ρ_2 reveals whether childless female STEM graduates differ from childless male STEM graduates in their probability to work in an unrelated job.

The coefficient of the triple interaction, λ, gives the difference in the probability of working in an unrelated job among female STEM graduates having children. With λ being the result of this triple interaction, $\delta_2 + \rho_2 + \lambda > \delta_2 + \rho_2 > 0$ would indicate that women with a STEM degree are over-proportionately not entering their occupational field due to childcare obligations, relative to men, relative to other professional fields and relative to those without children and vice versa.

We further follow Hunt (2016) by separating STEM into the groups engineering and computer sciences (EngComp), and mathematics and natural sciences (MatNat). For this, we replace the STEM dummies with a dummy for EngComp and for MatNat. Without the emphasis on fertility here, the regression equation looks as follows:

$$\begin{aligned} y_{it} = \alpha &+ \gamma_6 \cdot female_i + \delta_3 \cdot EngComp_i + \rho_3 \cdot EngComp_i \cdot female_i \\ &+ \delta_4 \cdot MatNat_i + \rho_4 \cdot MatNat_i \cdot female_i + X'_{it} \cdot \beta_3 + \eta_{it} \end{aligned} \tag{3}$$

We further use a combination of regressions equation (2) and (3), i.e. including the children-dummy and its interactions into equation (3). We will additionally add specific variables and interactions which will be named in the specific subsections to test the robustness of the results.

4 Data

We exploit a relatively underused data source for Germany, the Graduate Panel of the German Centre for Higher Education Research and Science Studies (DZHW). The survey aims at better understanding the career paths of German higher education graduates asking a variety of questions on the course of study, transition to a professional career, further education as well as sociodemographic characteristics. Since 1989, every fourth graduate cohort was surveyed in the Graduate Panel. The survey population consists of all higher education graduates who completed a degree at a German institution of higher education in either the winter or summer semester of the respective year. Due to the unique sample and survey design, the DZHW Graduate Panel offers the best opportunities to comprehensively examine research questions about German university graduates (Baillet et al., 2017a,b).

In the analysis, we pool together the observations of both the 2005 and the 2009 cohort. Both cohorts include graduates of traditional degree courses as well as Bachelor graduates. Including control variables for different types of degrees controls for biases from a potentially different labor market behavior of the respective graduates. We include information from the first and second survey waves, which both contain interviews with individuals one year and five years after graduation. This observation period is regarded as the career start of university graduates. The inclusion of the second wave helps to overcome the problem of individuals going for a gap-year after graduation or possible longer periods for labor market entry decisions, which might also relate to fertility. Thus, only if the career start needs more than five years, we miss this person. In the final data, only a minor part was without a job in general and thus is not dealt with any further.

Initially, the combined dataset contains 22,282 observations of which 11,788 are from the 2005 cohort, and 10,494 are from the 2009 cohort. We drop observations because they might follow a different underlying distribution, e.g., graduates who are younger than 21 and older than 40 by the time they finish their degree. Observations with more than one degree from different fields of study as well as individuals who report having more than one job at the same time are also excluded from the sample to prevent having unclear information in both explanatory and explained variables. Further, we only include those individuals in our analysis that responded to both survey waves[3]. The final sample used for the empirical analysis contains observations of 13,250 individuals observed at two waves. More than half of the individuals (58.92%) in the sample are female, and 13.36% are women with a STEM degree. An overview of the distribution of the nine subject groups and the degree subjects can be found in the appendix.

— Table 1 here —

A detailed overview of the summary statistics of all covariates is shown in table 1. Since the observations are pooled over both survey waves, the number of observations doubles for variables

[3]Results are robust to not dropping those only observed once. There are, however, some constant variables, such as origin, only surveyed in the second wave.

that do not contain any missing values. Most of the covariates are dummy variables that take on the value one if the statement is true and zero otherwise. Except for the variables *children*, *cohort*, *Current occupation in East-Germany* and some family information.

4.1 Dependent Variable

As outlined in chapter 3, information on the dependent variable (having a job unrelated to one's field of study) is taken from a survey question on how adequate an individual's current job is concerning his or her field of study. Originally, the survey question asks participants to rate the adequacy on a scale from 1 (definitely adequate) to 5 (not adequate at all). Since the intermediate points on the scale are not labeled, using all five levels for a multinomial outcome variable would provide little meaningful information, and interpretation of results would be less clear. Therefore, a *related job* is defined as such if individuals rated the match between field of study and current occupation with 1 or 2, while ratings of 3 constitute the middle category of having a *indifferent related* job and ratings of 4 or 5 are taken to define that an individual is currently holding an *unrelated* job.

— Figure 2 here —

Figure 2 gives an overview of the sample distribution of job adequacy as defined for this analysis. This view gives the first insight into the transition behavior of individuals in the sample. The graph shows that seven out of ten (70.23%) individuals in the sample reported that their current job is very closely related to their field of study. Only 15.35% rate their job as not being adequate concerning the field of study they majored.[4] Almost as many, 14.42%, can be considered to be somewhat indifferent or uncertain, stating that their job is somewhat adequate to their degree field. Thus, in general, graduates are more likely to have related than unrelated occupations.

Table 2 provides a more detailed description, which shows the levels of job unrelatedness by STEM and non-STEM fields of study, and by gender. The distribution of the job unrelatedness categories among the different fields of study for all working individuals shows that STEM graduates report more often having a job that matches the field of study than non-STEM graduates (74.41% and 67.96% of individuals respectively). Accordingly, having an unrelated occupation is much more of a problem for graduates of non-STEM fields (17.42%) than for STEM graduates (11.51%). Given that the education in STEM fields of study is very often targeted at a specific occupation (e.g., a degree in mechanical engineering aims at preparing for a career as a mechanical engineer), these results seem plausible. Looking at the two STEM subgroups does not provide surprises either: The reported shares for related and unrelated occupations are all roughly the same as the values for all STEM graduates. Engineering and computer science (EngComp) exhibit the lower shares of individuals with an unrelated occupation (10.76%), compared to mathematics and natural sciences (MatNat) (12.93%).

[4] Official numbers to assess the reliability of our measure are difficult to find. The OECD reports a field-of-study mismatch for Germany of 20%, however, data only exists for the years 2015 and 2016 (OECD, 2017).

Examining the difference between the genders is even more insightful. Among STEM graduates, women are more than three percentage points more likely to report having a job that is not related to their degree field. Although female graduates of non-STEM fields are also more likely to report an unrelated occupation, the gender difference is much smaller in this group. Again, this notion also translates to the opposite category of having a job that is closely related to one's university major. Among both groups, STEM and non-STEM graduates, women are also less likely to have a related job, with the difference again being more substantial for STEM than for non-STEM graduates. When looking at the STEM subgroups, it becomes once again apparent that a more differentiated perspective provides essential information. Within the group of engineering and computer science graduates, women are more likely to have a job that is unrelated to their STEM degree and less likely to have a job that is entirely related. In math and natural sciences, however, the share of males in the sample who report holding an unrelated occupation is higher than the share of females. There is also no gender difference in having a closely related occupation in math and the natural sciences.

For the purpose of the empirical analysis that will follow in chapter 5, the information of the three-categorical unrelatedness measure is further condensed into a binary variable, *unrelated* that is equal to one if a job is not adequate to the field of study and equal to zero otherwise, i.e., if a job is somewhat or entirely adequate. Since table 2 shows that the middle category does not differ much across either field of study nor gender, the gain in interpretability of the results due to the binary dependent variable outweighs the slight loss of information when dichotomizing the variable. A descriptive insight of job-unrelatedness for STEM and gender can be found in the supplementary appendix.

5 Empirical Results

The following chapter presents the results of the random effects regression analysis, as outlined in chapter 3. All models use standard errors that are clustered at the individual level in order to account for serial correlation across the two survey waves. If we use just one wave, we calculate heteroscedasticity robust standard errors. The outcome variable *unrelated* is equal to zero if the occupation is related and equal to one if the occupation is unrelated. A negative effect, thereby, gives an advantage to have a degree-related occupation.

Table 5 shows the basic regression results of equations (1), (2), (3) and a combination of the last two. First, in column one, we see a negative STEM effect of -7.7 percent at the 0.1 percent significance level on unrelatedness. Thus, male STEM-graduates are less likely to have an unrelated job. For females, this STEM-advantage is reduced, because the interaction term has a positive coefficient of 4.4 percentage points, significant at the 1 percent level. Adding both named effects gives the female STEM effect, which is equal to -3.3 percent and included at the bottom of the table. Thus, female STEM graduates, as well as males, have a STEM advantage to have a degree-related occupation, but it is lowered for women by 4.4 percent, from -7.7 to -3.3 percent. If we include the control variables in column (2), which are summarised in table 1, the STEM coefficient δ increases to -6.3 percent and the STEM$\times$female coefficient ρ increases to 5.0 percent, leaving a STEM advantage for women of only -1.3 percent compared to men and women of all other fields. Men, thereby, have a higher probability of working in a degree-related occupation if they have a STEM degree. Women as well, but the effect is much smaller.

— Table 5 here —

Next, we want to investigate if this gender difference can be explained by fertility. As shown in the literature review, this question is so far answered ambiguously and not specifically for the career start. Including the variable *children* as a measure of fertility and its interactions, we see that the effects mentioned above are robust and we cannot detect any fertility-specific effects once including the set of control variables in column (4). Thus, the child-specific effect found in column (3) is captured by the rich set of control variables. Since the field of STEM is heterogeneous, we split STEM into engineering and computer sciences (EngComp) and mathematics and natural sciences (MatNat). Column (5) and (6) show that EngComp drives the STEM effect and that fertility still does not play an essential role in the explanation of an unrelated occupation. We, therefore, find that children do not explain the relative (missing) transformation of STEM degrees into STEM occupations of men or women. This would suggest that current fertility is not essential for the job-entry decision, but we will come back to this more explicitly in the next subsection. The positive interaction effects show a general gender-penalty for women of an unrelated occupation for EngComp. It could also mean that that women just decided not to enter STEM or EngComp occupations. We will come back to this discussion later on. In the following, we will stick to the separation of STEM into EngComp and MatNat, because the main driver of the STEM effect is EngComp.

5.1 Potential fertility and before graduation

The job-entry and fertility decision might also be affected by the relationship-status of individuals. Employees might, even if the female graduates do not already have children, fear that women will get children. Alternatively, women might already plan to get children and thus do not start a career in the field of the degree, but just *a* job. This might be more likely if women are married, which is why we use a marriage-dummy to check for potential-fertility effects, similar to Becker et al. (2019).

Table 6 shows different specifications, including a marriage dummy, further relationship information, and more specific children-variables and their interactions with the gender-dummy. In column (1), replacing *children* with *married*, the male STEM advantage for a degree-related occupation increases marginally. The gender difference decrease, leading to a higher probability for female EngComp graduates to have a degree-related job up to -3.9 percent.

— Table 6 here —

Next, in column (2) and (3), we use both, children and marriage and their interactions, while including the triple interactions with children in column (2) and with marriage in column (3).

Here, δ is robust, while ρ, however, jumps up to 6.1 percentage points in column (2) and down again to 4.7 percentage points in column (3). It does seem that the different triple interaction control for different things while being insignificant themselves. We further note that the interaction of EngComp and marriage is robust significant at the 5 or even 1 percent level with a coefficient of 5.0 and 4.4. Thus, being married and having a degree in EngComp increases the likelihood of an unrelated occupation, independent of the gender of the individual. Next, we include further information if the individual is in a relationship and whether their partner works or not and interact both variables with the gender dummy. Results as well as the overall explanation of the regression is robust.

The regression results presented in table 6 thus show that we cannot explain the gender difference by including marital status information in the model. The coefficient for the general gender difference varies only from 6.1 percentage points to 4.7 percentage points, which is in each other 95% confidence intervals. The marital status as well as having children or not, however, is likely endogenous and, thereby, difficult to interpret and does not even explain a significant portion of the effects of the gender difference. In column (5) and (6) we, therefore, make use of the information when the individuals graduated and the year in which the children were born. Thus, this child-information should not be endogenous anymore, because we know it happened before obtaining the degree. In column (5), leaving the marital status information in the model, we add whether at least one child was born before the year in which individuals graduated and interact this information with the gender dummy and both the gender and degree dummy, while we exclude the marital status in column (6). The gender difference is in both columns now somewhat in-between the results from before at 5.5 percentage points, significant at the 1 percent significance level. The newly included variables are not significant anymore but do impact marginally the gender difference, which is why we will keep those three variables from now on the model in the set of control variables.

The purpose of all subsequent subsections is to see if we can explain this general gender difference between EngComp graduates. Since fertility nor marriage explained the gap and might be endogenous because of simultaneous decisions we cannot separate, we will not include those variables anymore. We only include a variable indicating if at least one child was born before graduation because there we know the sequence of decisions[5]. This is, as far as we know, unique in the data compared to other studies in this field and does have a marginal impact on the variables of interest comparing column (6) in the tables 5 and 6.

5.1.1 Childcare per county

We use five different variables to capture possible regional childcare information. We include the percentage of (i) children under three years in childcare, (ii) children between three to six and (iii) children with a foreign origin at the district of the current occupation[6]. The first two are further

[5] We also tested if all children were born before graduation and if just all children of the first wave were born before graduation and this did not impact the results compared to column (6).

[6] Table 3 in the appendix further shows the summary statistics for those variables

separated into childcare for at least 7 hours and without an hourly-restriction. Additional, we further include a variable about the density of the population in the same district to distinguish rural and urban areas. In table 7, we see that the inclusion of neither of these variables affects the general EngComp effect or the gender difference, nor do the variables themselves add much explanatory power the probability of an unrelated occupation.

5.1.2 Robustness

We additionally check if the basic regression results of equation (1) and (3) with the timing of the childbirth information are robust for different sample restrictions and additional variables. First, we only leave individuals who report their current job either as a mid-term or long-term solution in column (1) and (2) of table 8. Next, we include a continues variable for the share of males in each area of study and interact it with the gender dummy in column (3) and (4). Then, we include the information if individuals desire children and interact this with the female dummy. Problematic here, however, is, that we need to assume that the desire is constant over the observation period because the question is only asked in wave 2. At last, we include the age of the youngest child before graduation instead of the dummy for children born before graduation and its interaction. The gender difference varies from 4.8 percent up to 7.0 percentage points and thus produces nothing new. The newly introduced variable are all insignificant and do not add specific information.

5.1.3 Sensitivity

For the dependent variable, we recoded a 5-Likert scale variable. We run a sensitivity check in which we sort the middle category now the other way to see if this decision influences our results. Further, we use the 5-Likert scale to see if our simplification produces any problems. The regression results in table 9 column (1) to (4) show that the results are very robust to our baseline regression results.

5.1.4 Heterogeneity within waves

Next, we check whether the gender difference is prevalent for both waves. This way, we check for short-term and medium-term effects. In table 10, we include regression specifications following equation (3), just that we include *children born before graduation* and its interaction with the gender in every second column. First, for wave 1 (of 2), we see a pronounced male effect for a degree in EngComp by 11.1 percent at the 0.1 percent significance level and a gender difference which is comparable to the regressions before of 6.2 percentage points. If only the male effect decreases and the gender difference is constant, the actual female effect to have an unrelated occupation decreases as well. Thus, making it more likely that females have a degree-related occupation. If we then include the new children information in column (2), results show that the estimations are robust and do not change as before.

For wave two, however, the story changes. Men have a lowered but still, negative probability having an unrelated job by -4.6, the lowest so far. Further, the gender difference is almost equal in absolute value and just marginal significant at the 10% level by 4.2 percentage points. In column (4), including children born before graduation, the male advantage decreases marginally to -4.5 percent, and the gender difference becomes significant at the 5 percent level again, increasing marginally to 4.4 percentage points. Thus, the absolute value of both the male effect and the gender difference are almost equal, leading to a non-existing female benefit in the second wave, once controlling for children born before graduation.

5.1.5 Mediator Analysis

Why do females not benefit in the same manner as men do from an EngComp degree? To find an answer, we include intermediate outcomes. Problematic here is, however, that not all individuals gave full information on those variables. Table 4 explains the whole set of mediators. First, in column (1) of table 11, we added variables which indicate if individuals studied for a better position in the labor market, whether the jobs during their studies corresponded to their field, how they value their degree and whether they would have studied (the same field) again. Those variables should help to control for believes of the specialization and their degree, and its connection to the labor market. This intake should sort between individuals who are not interested in still pursuing a career within the field and those who do not. Column (2) includes information on the application process, i.e., the number of job applications, interviews, and offers, and if the job was assigned. This information should help to control for higher competition in the job search and the general access to jobs. We, unfortunately, do not know whether job application corresponded to the degree. Next, in column (3), we include information on the difficulties in the job search, which should give more insights into the problems individuals face on the job market. Column (5) then includes all the named mediators, additional including self-rated abilities and the firm sizes of the employees. To check if dropped individuals drive results, column (6) excludes all mediators but uses the same number of observations.

We see that all variables increase the general effect for EngComp graduates from -8.2 percent from table 11 column (5) (in which no mediators are included, but them sample size is equivalent to column (4) with all mediators) to at best -5.2 percent in column (2) and to -5.5 percent in column (4) in which we include all mediators. The mediators thus explain a small portion of the general effect. The gender difference decreases from 6.1 percentage points from column (5), e.g., to 3.8 percent in column (1) and to 4.5 percent in column (4). Due to those changes, the female EngComp effects changes as well. The baseline from table 6 column (7) is set with -0.024. The female effect varies from -2.9 (column (1) up to -0.8 in column (2) or -1.0 in column (4). Thus, if we include all mediators, the male EngComp advantage for a degree-related occupation decreases as well as for women. If we compare this to column (5), we see that the changes are not due to the drop of individuals but due to the inclusion of the mediators. The significance of the gender-difference is for most of the regressions only marginally significant anymore. Thus, also, the certainty of the significance is not as strong as it was before.

5.2 Discussion

All in all, one clear results emerge from the empirical analysis: The regression analysis has shown that there are a higher rate of non-transitions to EngComp (and thereby STEM) occupations for women compared to men. The EngComp work environment, therefore, seems to be significantly less attractive to its female graduates compared to males. The second part of the empirical analysis has shown that people are not more likely to not enter STEM occupations than other professional fields in response to childcare responsibilities. Even though the variable children born before graduation is a helpful control variable, it does not add additional explanation itself. Further, the gender difference in EngComp is mostly driven by the East of Germany and by the first wave, the primary career start period.

The missing fertility effect is in line with, e.g. Polachek (1981) and Jansen and Pascher (2013). Females balancing work and (family) life might not even study STEM in the first place. The literature review showed that not mathematical abilities (e.g. Friedman-Sokuler and Justman, 2016) but differences in self-efficacy (e.g. Heilbronner, 2013), personal interest (e.g. Heilbronner, 2013) and willingness to compete (e.g. Buser et al., 2017) are reason for difference in STEM and EngComp. The graduates in our sample, however, already self-selected themselves into STEM or EngComp studies, showing some confidence and interest in those fields.

In the mediator analysis, we see, that a set of mediators can explain only a small part of the general gender difference in EngComp. The rest of the effect is either driven by a self-selection process or statistical discrimination against women, which we cannot test. The self-selection should be, however, irrespective of the reasons individuals (i) studied for a better preparation for the labor market, (ii) how they value the degree in terms of (a) a possible interesting career, (b) longer lasting learning, (c) usefulness for the career, (d) personal development and (e) knowledge for the career. Further, we control in the mediator analysis for individuals who (iii) disliked studied in general or specifically the field they where. We can further rule out a general gender problem, because when including (iv) a variable for the share of males per area of study into the regression, the gender difference for EngComp is robust. Additionally, we also include information on the (v) application process, (vi) difficulties with the jobs search and (vii) different firm sizes.

The remaining gender difference could, however, also be due to a hostile work environment towards women in EngComp (e.g. Danbold and Huo, 2017) or statistical discrimination (e.g. Riach and Rich, 2006). Unfortunately, given the data, we cannot specifically analyze which reason might be true for the remaining gender-difference. The missing significance for fertility is partly in line with Becker et al. (2019), who, using a correspondence test study, find only worse callback rates for married but childless females applying for part-time jobs. Our results for marriage is, however, only EngComp specific and not gender-specific.

Overall, the gender gap in the EngComp field predicted by the model seems to be small in size but existing nonetheless. Additionally, it might discourage female EngComp graduates from entering EngComp occupations in the future, creating a vicious circle in the attraction of female talent to the STEM sector due to missing role models.

6 Conclusion

We use a sample of German university graduates from all fields of study in order to determine whether there is a specific non-transition for women from STEM fields of study. We, therefore, focus on the job-entry decision only and not general exit studies such as Hunt (2016). Non-transitions are considered as such if a graduate reports holding a job unrelated to his or her field of study within the first five years after graduation.

The regression analysis shows that graduates from EngComp (and thereby STEM) are more likely to enter an occupation related to their study compared to other fields. Second, this higher transition-rate is reduced for women for EngComp occupations compared to men and relative to women of other professional fields. This results is robust to different model specification and more pronounced in the first wave. For the second wave, we even find no STEm advantage for women anymore. The information if individuals have born a child before the graduation year does not add further explanation but serves as an important control variable affecting the estimator of the gender difference. Additionally, the mediator analysis helps to explain a small portion of the gender differences. This shows that we can only partly explain this gender difference by their opinions, believes and difficulties graduates have after their studies. Other reasons, such as statistical discrimination against women in EngComp, could not be tested with the usage of the data and should be focussed in future research.

References

Ardicianono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics 121*, 343–375.

BA (2018). Berichte: Blickpunkt Arbeitsmarkt - MINT-Berufe.

Baillet, F., A. Franken, and A. Weber (2017a). DZHW-Absolventenpanel 2005. Daten- und Methodenbericht zu den Erhebungen der Absolvent(inne)nkohorte 2005 (1. und 2. Befragungswelle). Version 1.0.0. FDZ-DZHW.

Baillet, F., A. Franken, and A. Weber (2017b). DZHW-Absolventenpanel 2009. Daten- und Methodenbericht zu den Erhebungen der Absolvent(inne)kohorte 2009 (1. und 2. Befragungswelle). Version 1.0.0. FDZ-DZHW.

Becker, S. O., A. Fernandes, and D. Weichselbaumer (2019). Discrimination in hiring based on potential and realized fertility: Evidence from a large-scale field experiment. *IZA Discussion Paper Series*, 1–37.

Biewen, M. and J. Schwerter (2019). Does more math in high school increase the share of female STEM workers? Evidence from a curriculum reform. *Mimeo*, 1–42.

Burke, R. J. (2007). Women and minorities in STEM: A primer. In R. J. Burke and M. C. Mattis (Eds.), *Women and Minorities in Science, Technology, Engineering and Mathematics. Upping the numbers*, pp. 3–28. Edward Elgar.

Buser, T., N. Peter, and S. C. Wolter (2017). Gender, Competitiveness, and Study Choices in High School: Evidence from Switzerland. *American Economic Review 107*(5), 125–130.

Ceci, S. J. and W. M. Williams (2010). Sex Differences in Math-Intensive Fields. *Current directions in Psychological Science 19*(5), 275–279.

Danbold, F. and Y. J. Huo (2017). Men's Defense of their Prototypicality Undermines the Success of Women in STEM Initiatives. *Journal of Experimental Social Psychology 72*, 57–66.

Enman, M. and J. Lupart (2000). Talented Female Student's Resistance to Science: An exploratory study of post-secondary achievement motivation, persistance, and epistemological characteristics. *High Abitliy Studies 11*(2), 161–178.

Fehse, S. and C. Kerst (2007). Arbeiten unter Wert? Vertikal und horizontal inadäquate Beschäftigung von Hochschulabsolventen der Abschlussjahrgänge 1997 und 2001. *Beiträge zur Hochschulforschung 29*(1), 72–98.

Friedman-Sokuler, N. and M. Justman (2016). Gender streaming and prior achievement in high school science and mathematics. *Economics of Education Review 53*, 230–253.

Heilbronner, N. N. (2013). The STEM Pathway for Women: What Has Changed? *Gifted Child Quarterly 57*(1), 39–55.

Hübner, N., E. Wille, J. Cambria, K. Oschatz, B. Nagengast, and U. Trautwein (2017). Maximizing Gender and Equality by Minimizing and Course Choice and Options? Effects and of Obligatory. *Journal of Educational Psychology 109*(7), 993–1009.

Hunt, J. (2016). Why do women leave science and engineering? *ILR Review 69*(1), 199–226.

Ivanova, M. and P. Stein (2013). Bachelorabschluss: Endstation von Chemikerinnen? Einstellungen von Studierenden der Chemie zum Studium und zur beruflichen Karriere - Ergebnisse einer Befragung an zwölf ausgewählten Hochschulen in Deutschland. In U. Pascher and P. Stein (Eds.), *Akademische Karrieren von Naturwissenschaftlerinnen gestern und heute*, pp. 125–149. Springer Fachmedien.

Jansen, K. and U. Pascher (2013). "Und dann hat man keine Zeit mehr für Familie oder so." - Wissenschaftsorientierung und Zukunftsvorstellungen von Bachelorstudentinnen chemischer Studiengänge. In U. Pascher and P. Stein (Eds.), *Akademische Karrieren von Naturwissenschaftlerinnen gestern und heute.*, pp. 151–192. Springer Fachmedien.

Joy, L. (2000). Do Colleges Shortchange Women? Gender Differences in the Transition from College to Work. *American Economic Review 90*(2), 471–475.

Justman, M. and S. J. Mendez (2018). Gendered choice of STEM subjects for matriculation are not driven by prior differences in mathematical achievement. *Economics of Education Review 64*, 282–297.

Kahn, S. and D. K. Ginther (2015). Are recent cohorts of women with engineering bachelors less likely to stay in engineering? *Frontiers in Psychology 6*, 1–15.

Knauf, H. and E. Rosowski (2009). Wie tragfähig ist die Studien- und Berufswahl? Biographische Verläufe und Orientierungsprozesse nach dem Abitur. In M. Oechsle, H. Knauf, C. Maschetzke, and E. Rosowski (Eds.), *Abitur und was dann? Berufsorientierung und Lebensplanung junger Frauen und Männer und der Einfluss von Schule und Eltern*, pp. 283–324. VS Verlag für Sozialwissenschaften.

Lubinski, D. and C. P. Benbow (2006). Study of Mathematically Precocious Youth After 35 Years: Uncovering Antecedents for the Development of Math-Science Expertise. *Perspectives on Psychological Science 1*(4), 316–345.

Morgan, L. A. (2000). Is Engineering Hostile to Women? An Analysis of Data from the 1993 National Survey of College Graduates. *American Sociological Review 65*(2), 316–321.

OECD (2017). Getting Skills Right: Skills for Jobs Indicators. resreport, OECD Publishing.

Perna, L. W. (2004). Understanding the Decision to Enroll in Graduate School: Sex and Racial/Ethnic Group Differences. *The Journal of Higher Education 75*(5), 487–527.

Polachek, S. W. (1981). Occupational Self-Selection: A Human Capital Approach to Sex Differences in occupational structure. *The Review of Economics and Statistics 63*(1), 60–69.

Preston, A. E. (1994). Why Have All the Women Gone? A Study of Exit of Women from the Science and Engineering Professions. *The American Economic Review 84*(5), 1446–1462.

Preston, A. E. (2004). *Leaving science. Occupational exit from scientific careers.* Russel Sage Foundation.

Riach, P. and J. Rich (2006). An experimental investigation of sexual discrimination in hiring in the english labor market. *The B.E. Journal of Economic Analysis & Policy 5*, 1–22.

Robst, J. (2007). Education and job match: The relatedness of college major and work. *Economics of Education Review 26*, 397–407.

Sassler, S., J. Glass, Y. Levitte, and K. M. Michelmore (2017). The missing women in STEM? Assessing gender differentials in the factors associated with transition to first jobs. *Social Science Research 63*, 192–208.

Simard, C., A. D. Henderson, S. K. Gilmartin, L. Schiebinger, and T. Whitney (2013). Climbing the technical ladder: Obstacles and Solutions for mid-level women in technology. resreport, Michelle R. Clayman Institute for Gender Research at Stanford University and Anita Borg Institute for Women and Technology.

Verhaest, D., S. Sellami, and R. van der Velden (2017). Differences in horizontal and vertical mismatches across countries and fields of study. *International Labour Review 156*(1), 1–23.

Wang, M.-T., J. S. Eccles, and S. Kenny (2013). Not Lack of Ability but More Choice: Individual and Gender Differences in Choice of Careers in Science, Technology, Engineering, and Mathematics. *Psychological Science 24*(5), 770–775.

Wiswall, M. and B. Zafar (2015). Determinants of College Major Choice: Identification using an Information Experiment. *Review of Economic Studies 82*, 791–824.

Xu, Y. J. (2013). Career Outcomes of STEM and Non-STEM College Graduates: Persistence in Majored-Fields and Influential Factors in Career Choices. *Research in Higher Education 54*, 349–382.

Zafar, B. (2013). College major choice and the gender gap. *Journal of Human Resources 48*, 546–593.

Appendix

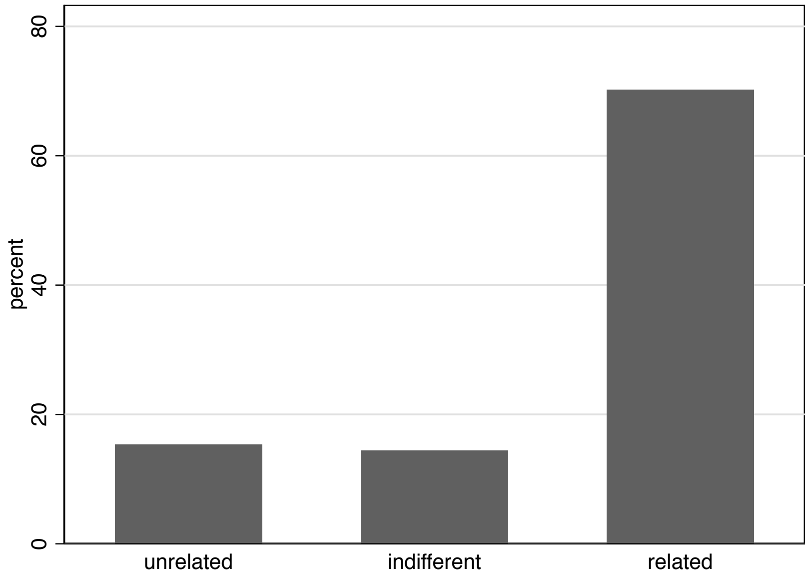
Table 1: Summary statistics of covariates

	Obs	Mean	Std.Dev.	Min	Max
Unrelated	13181	0.153	0.3605	0	1
Female	13181	0.589	0.492	0	1
STEM	13181	0.350	0.4771	0	1
Math	13181	0.026	0.1592	0	1
Natural Sciences	13181	0.096	0.2952	0	1
Math & Natural Sciences	13181	0.122	0.3271	0	1
Computer Sciences	13181	0.055	0.2274	0	1
Engineering	13181	0.174	0.379	0	1
Engineering & Computer Sciences	13181	0.229	0.4199	0	1
Children	13181	0.200	0.3997	0	1
At least one child born before graduation	13181	0.044	0.2061	0	1
Experience					
Vocational training before university	13181	0.268	0.4429	0	1
Employment before university	13181	0.317	0.4654	0	1
Voluntary internship	13181	0.388	0.4873	0	1
Mandatory internship	13181	0.535	0.4988	0	1
Student assistant	13181	0.366	0.4818	0	1
Working student	13181	0.332	0.471	0	1
Parental background					
At least one parent with Abitur	13181	0.226	0.4184	0	1
At least one parent with a university degree	13181	0.274	0.4461	0	1
At least one parent with a blue-collar occupation	13181	0.061	0.2385	0	1
Personal					
Age at degree completion	13181	26.209	2.7941	21	40
Age at degree completion squared	13181	694.719	159.4631	441	1600
Birthyear	13181	1979.864	3.5809	1964	1988
Cohort	13181	0.390	0.4877	0	1
Wave	13181	0.508	0.4999	0	1
HEEQ in East-Germany	13181	0.291	0.4544	0	1
University degree in East-Germany	13181	0.300	0.4582	0	1
Current occupation in East-Germany	13181	0.249	0.4324	0	1
Foreign	13181	0.032	0.1748	0	1
Educational Background					
Grade of HEEQ	13181	2.213	0.6143	1	4
Year of HEEQ	13181	1999.677	3.3588	1983	2006
Field-specific HEEQ	13181	0.001	0.0246	0	1
HEEQ from vocational school	13181	0.020	0.1411	0	1
Foreign HEEQ	13181	0.107	0.3095	0	1
High School at vocational school	13181	0.046	0.2093	0	1
University Information					
Grade of University degree	13181	2.2123	0.6143	0.8	4
Type of degree: Diplom	13181	0.5830	0.4931	0	1
Type of degree: Magister	13181	0.067	0.2495	0	1
Type of degree: Bachelor	13181	0.165	0.3708	0	1
Type of degree: State Examination	13181	0.086	0.2796	0	1
Type of degree: Teaching degree	13181	0.097	0.2962	0	1
Type of degree: Other	13181	0.003	0.0557	0	1
University	13181	0.677	0.4676	0	1

Applied University	13181	0.323	0.4676	0	1
Family Information					
Partner with a job	13181	0.620	0.4854	0	1
Partner without a job	13181	0.119	0.3238	0	1
In a relationship	13129	0.488	0.4999	0	1
Married	13129	0.256	0.4363	0	1

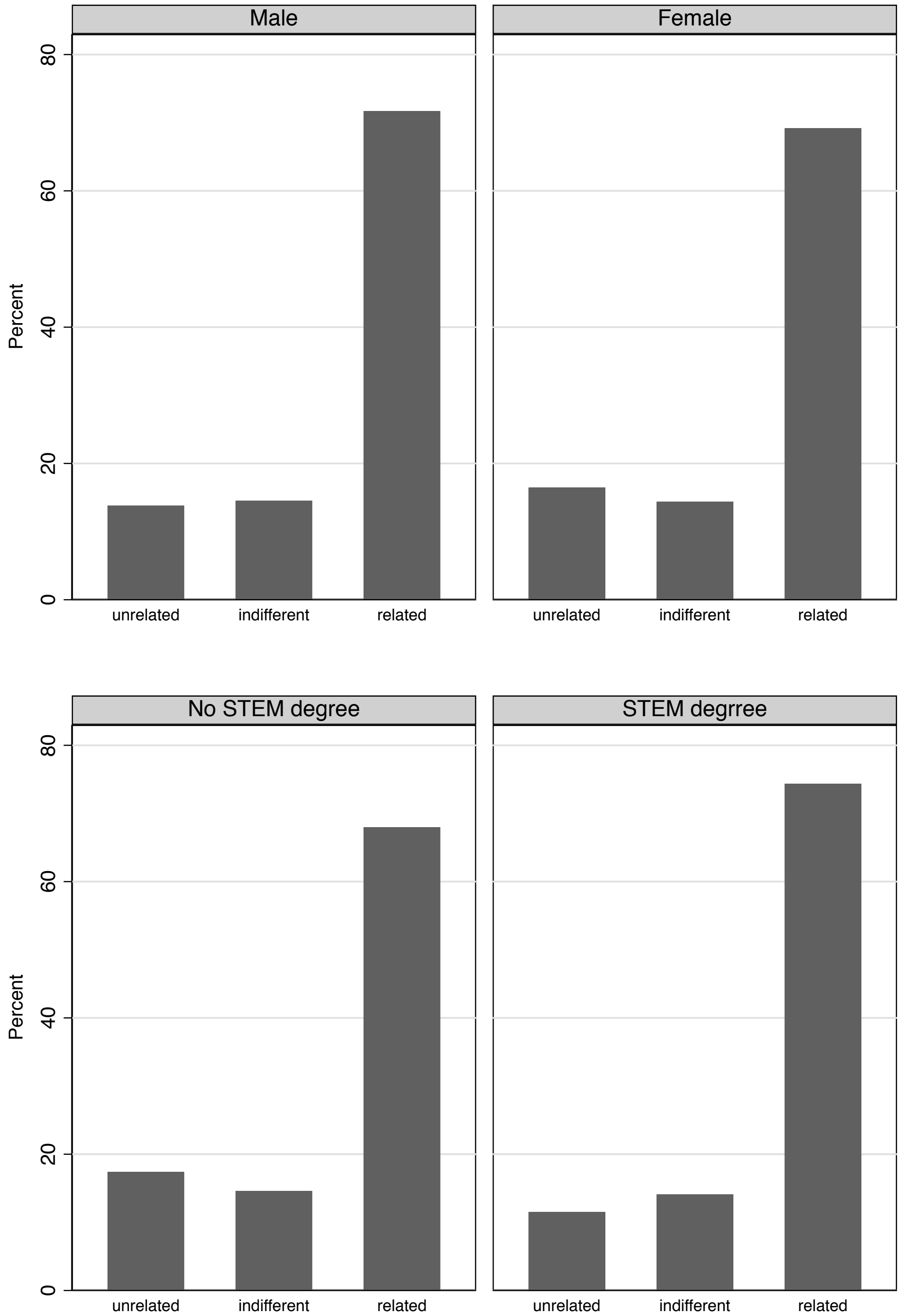
Notes: The table shows summary statistics for all covariates used in the empirical analysis, pooled over both survey waves. In the regressions, we further use state specific effects for the state in which the higher education entrance qualification (HEEQ) was obtained. Source: DZHW Graduate Panel 2005 and 2009, own calculations.

Figure 1: Reported levels of job unrelatedness



Note: The graph shows the reported levels of job unrelatedness in the sample, pooled over both waves. Job unrelatedness is proxied with the survey question on how closely the field of study matches the current job. For the regression analysis, we construct a binary variable in which we include the group of *indifferent* in *related.* A sensitive check including the middle category in *unrelated* confirms the results but with higher coefficients. For 2016, official OECD-data find a job-unrelatedness for Germany of around 20 percent, which is similar to the data in use.

Figure 2: Reported levels of job unrelatedness



Note: The graph shows the above graphic conditional on gender and STEM-degree. In general, males and STEM graduates are less likely to have an unrelated job.

Table 2: Job Adequacy by Field of Study and Gender

Current job is adequate to field of study	Gender Male %	Female %	All %
all STEM			
Unrelated	10.15	13.78	11.54
Indifferent	14.04	14.24	14.12
Related	75.81	71.98	74.35
Math & Natural Sciences			
Unrelated	12.83	13.03	12.95
Indifferent	12.01	12.73	12.45
Related	75.16	74.25	74.60
Engineering & Computer Science			
Unrelated	9.43	14.77	10.79
Indifferent	14.59	16.21	15.00
Related	75.98	69.02	74.21
Other fields of study			
Unrelated	17.84	17.22	17.40
Indifferent	15.07	14.42	14.61
Related	67.10	68.37	67.99

Note: The table shows the relationship between job relatedness and degree of study, pooled over both survey waves. The corresponding survey question was: "Would you say that your higher education qualification matches your job, with respect to the academic qualification (field of study). Rate on a scale from 1 *Yes, definitely* to 5 *No, not at all.*" Ratings of 1 and 2 have been converted to *related job*, a rating of 3 to *indifferent*, and ratings of 4 and 5 to *unrelated job*. Refers to page 8.

Table 3: Summary statistics of childcare variables

Share of children...	Obs	Mean	Std.Dev.	Min	Max
... under 3 years in care for at least 7 hours	12432	0.1422	0.12222	0.0010	0.5500
... under 3 years in care	12432	0.2525	0.1341	0.0300	0.6280
... between 3 and 6 years in care for at least 7 hours	12432	0.3855	0.2276	0.0125	0.9475
... between 3 and 6 years in care	12432	0.2830	0.1511	0.757	1
... with foreign origin under 3 in care	12432	0.2830	0.1511	0	.7110
Urbanization	12432	2.2317	0.7447	1	3

Notes: The information required on-site usage of the data because of a lowered anonymization restriction with as well allowed us to the urbanisation criteria, which is why this variable is here included as well.

Table 4: Summary statistics of mediators

	Obs.	Mean	St. Dev.	Min	Max
study for the labor market:					
role of labor market: in choice of studies	13155	0.316	0.4649	0	1
course of study: career preparation	13072	0.262	0.4396	0	1
Value of degree:					
interesting career	13095	0.836	0.3705	0	1
longer lasting learning	13084	0.683	0.4655	0	1
usefulness for career	13079	0.596	0.4907	0	1
personal development	13102	0.830	0.3755	0	1
knowledge for career	13066	0.515	0.4998	0	1
Job during study corresponds to subject					
elf-employed	11871	0.158	0.3645	0	1
does not correspond to subject	11810	0.561	0.4963	0	1
In retrospect:					
study again	12946	0.914	0.2802	0	1
study the same subject again	12999	0.687	0.4636	0	1
Application process					
number of applications	11111	15.449	25.6499	0	500
number of interviews	11111	2.701	3.0608	0	35
number of job offers received	11111	1.499	1.6968	0	31
no application, job assigned to me	11111	0.095	0.2933	0	1
Present subjective ability					
specialised subject knowledge	13015	2.453	0.9682	1	5
broad basic knowledge	13024	2.103	0.8307	1	5
knowledge of academic methods	12977	2.402	0.9621	1	5
foreign languages	13001	2.938	1.2455	1	5
ability to communicate	13018	2.180	0.9229	1	5
negotiation skills	12997	3.313	1.0397	1	5
organisational skills	13015	2.144	0.9609	1	5
IT skills	13024	2.207	1.0272	1	5
flexibility	12990	2.261	0.9323	1	5
written expression skills	13031	2.056	0.9175	1	5
oral expression skills	13029	2.161	0.8729	1	5
recognise and fill gaps in knowledge	12996	2.025	0.8097	1	5
leadership qualities	13001	3.307	1.0145	1	5
business knowledge	13010	3.503	1.2067	1	5
ability to cooperate	12992	2.107	0.8641	1	5
time management	13015	2.300	0.9882	1	5
application of knowledge	12993	2.165	0.8195	1	5
interdisciplinary thinking	12987	2.314	0.9060	1	5
intercultural competencies	12999	3.001	1.2424	1	5
independent working	13023	1.639	0.7569	1	5
taking responsibility	13001	2.231	0.9512	1	5
conflict management	13011	2.941	1.0001	1	5
problem solving abilities	13019	2.289	0.8474	1	5
analytical skills	13004	2.215	0.9289	1	5

Continued on next page...

... table 4 continued

	Obs.	Mean	Std. Dev.	Min	Max
difficulties with job search:					
few job vacancies	11212	0.381	0.4857	0	1
other specialisation required	11212	0.158	0.3643	0	1
other degree required	11212	0.081	0.2725	0	1
other salary expectations	11212	0.136	0.3431	0	1
other work expectations	11212	0.106	0.3081	0	1
professional experience required	11212	0.564	0.4960	0	1
too far away	11212	0.192	0.3942	0	1
knowledge deficit	11212	0.134	0.3409	0	1
expectations of contents	11212	0.210	0.4076	0	1
family/partner	11212	0.137	0.3440	0	1
other problems	11212	0.068	0.2525	0	1
no problems	11650	0.212	0.4085	0	1
Firm size:					
firm size above 100	11697	0.179	0.3831	0	1
firm size below 100	11697	0.442	0.4967	0	1
freelancer	11697	0.037	0.1878	0	1

Notes: In the regressions later on, the first three application variables are categorised as follows: No application, 1 to 10 applications, 11 to 50; 50 to 500. No interview; 1 to 5 interviews; 5 to 35. No offers; 1 to 3; 3 to 31.

Table 5: Main Regression Results

	(1)	(2)	(3)	(4)	(5)	(6)
STEM	-0.077***	-0.063***	-0.085***	-0.070***		
	(0.0105)	(0.0138)	(0.0116)	(0.0146)		
EngComp					-0.078***	-0.082***
					(0.0142)	(0.0149)
MatNat					-0.007	-0.008
					(0.0195)	(0.0195)
STEM×female	0.044**	0.050**	0.044**	0.049**		
	(0.0150)	(0.0153)	(0.0163)	(0.0164)		
EngComp×female					0.052**	0.063**
					(0.0184)	(0.0200)
MatNat×female					0.012	0.013
					(0.0215)	(0.0215)
female	-0.007	0.011	-0.010	0.005	0.011	0.005
	(0.0102)	(0.0104)	(0.0113)	(0.0113)	(0.0104)	(0.0112)
children			-0.041*	-0.009		0.000
			(0.0169)	(0.0168)		(0.0156)
STEM×children			0.040+	0.032		
			(0.0225)	(0.0222)		
EngComp×children						0.020
						(0.0218)
female×children			0.021	0.023		0.027
			(0.0203)	(0.0199)		(0.0185)
STEM×female×children			0.002	0.009		
			(0.0333)	(0.0331)		
EngComp×female×children						-0.055
						(0.0393)
Control variables	No	Yes	No	Yes	Yes	Yes
Female degree effect	-0.033	-0.013	-0.041	-0.020	-0.026	-0.019
R^2	0.007	0.087	0.008	0.088	0.089	0.089
N	13181	13181	13181	13181	13181	13181

Notes: Cluster robust standard errors in parenthesis. Female degree effect is either for STEM or for EngComp, depending on the degree variable in the regression specification. Regression results showing all coefficients is available upon request.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Fertility, potential fertility and birth of children before graduation

	(1)	(2)	(3)	(4)	(5)	(6)
EngComp	-0.086***	-0.089***	-0.084***	-0.085***	-0.088***	-0.079***
	(0.0153)	(0.0152)	(0.0154)	(0.0154)	(0.0150)	(0.0143)
EngComp×female	0.047*	0.061**	0.047*	0.048*	0.054**	0.055**
	(0.0199)	(0.0200)	(0.0199)	(0.0199)	(0.0188)	(0.0187)
female	0.008	0.004	0.007	0.009	0.007	0.011
	(0.0118)	(0.0117)	(0.0118)	(0.0156)	(0.0116)	(0.0106)
children		0.009	0.016	0.017		
		(0.0183)	(0.0172)	(0.0174)		
married	-0.009	-0.018	-0.016	-0.023	-0.011	
	(0.0136)	(0.0153)	(0.0159)	(0.0179)	(0.0130)	
female×children		0.022	0.012	0.012		
		(0.0215)	(0.0196)	(0.0197)		
female×married	0.012	0.010	0.007	0.001	0.016	
	(0.0162)	(0.0177)	(0.0188)	(0.0220)	(0.0150)	
EngComp×children		-0.008	-0.025	-0.024		
		(0.0255)	(0.0217)	(0.0218)		
EngComp×married	0.031	0.050*	0.044*	0.044+	0.038*	
	(0.0195)	(0.0206)	(0.0225)	(0.0225)	(0.0174)	
EngComp×female×children		-0.053				
		(0.0395)				
EngComp×female×married	0.019		0.019	0.017		
	(0.0393)		(0.0393)	(0.0393)		
in a relationship				-0.009		
				(0.0119)		
in a relationship×female				-0.008		
				(0.0160)		
partner without work				-0.006		
				(0.0125)		
partner without work×female				0.031		
				(0.0202)		
At least one child born before graduation					-0.018	-0.021
					(0.0400)	(0.0397)
At least one child born before graduation×female					0.000	0.006
					(0.0470)	(0.0467)
EngComp×At least one child before graduation					0.009	0.024
					(0.0536)	(0.0535)
EngComp×At least one child before grad.×female					-0.065	-0.074
					(0.0859)	(0.0858)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
MatNat and interaction with female	Yes	Yes	Yes	Yes	Yes	Yes
Female-EngComp Effect	-0.039	-0.028	-0.037	-0.037	-0.034	-0.024
R^2	0.089	0.089	0.089	0.089	0.089	0.089
N	13129	13129	13129	13129	13129	13181

Notes: Cluster robust standard errors in paranthesis. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Childcare information

	(1)	(2)	(3)
EngComp	-0.080***	-0.079***	-0.079***
	(0.015)	(0.015)	(0.015)
EngComp×female	0.055**	0.053**	0.052**
	(0.019)	(0.020)	(0.020)
female	0.013	0.015	0.189
	(0.011)	(0.011)	(0.175)
At least one child born before graduation year		0.014	0.010
		(0.024)	(0.024)
child before graduation year×female		-0.021	-0.013
		(0.028)	(0.029)
EngComp×child before graduation year		-0.022	-0.022
		(0.031)	(0.031)
EngComp×child before graduation year×female		0.033	0.031
		(0.058)	(0.058)
MatNat	-0.008	-0.008	-0.009
	(0.021)	(0.021)	(0.021)
MatNat×female	0.014	0.012	0.012
	(0.023)	(0.023)	(0.023)
share of children			
... under 3 years in care for at least 7 hours	-0.055	-0.047	0.060
	(0.078)	(0.079)	(0.153)
... under 3 years in care	-0.095	-0.103	-0.130
	(0.081)	(0.081)	(0.111)
... between 3 & 6 years in care for at least 7 hours			-0.012
			(0.058)
.... between 3 & 6 years in care			0.184
			(0.146)
... with foreign origin under 3 in care			-0.001
			(0.059)
urbanisation			-0.009
			(0.012)
Control variables	Yes	Yes	Yes
Interaction with share of children variables	No	No	Yes
Female-EngComp Effect	-0.025	-0.026	-0.027
R^2	0.090	0.090	0.091
N	12149	12131	12131

Notes: Cluster robust standard errors in parenthesis. Small reduction in observations because some individuals did not gave their zip-code information which are needed to merge the childcare information. We run a test-regression with just this subset of individuals without including any other variable and got robust results.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Robustness Check

	(1)	(2)	(3)	(4)
EngComp	-0.062***	-0.098***	-0.080***	-0.080***
	(0.0146)	(0.0195)	(0.0146)	(0.0143)
EngComp×female	0.050*	0.070**	0.048*	0.055**
	(0.0195)	(0.0251)	(0.0190)	(0.0187)
female	0.008	0.021	0.029^{+}	0.011
	(0.0106)	(0.0194)	(0.0171)	(0.0106)
At least one child born before graduation	-0.015	-0.021	-0.024	-0.001
	(0.0405)	(0.0398)	(0.0410)	(0.0487)
At least one child born before graduation×female	-0.006	0.005	-0.002	-0.011
	(0.0478)	(0.0468)	(0.0479)	(0.0617)
EngComp×At least one child born before graduation	0.006	0.026	0.023	0.033
	(0.0513)	(0.0535)	(0.0541)	(0.0548)
EngCom×At least one child born before graduation×female	-0.036	-0.076	-0.068	-0.080
	(0.0928)	(0.0860)	(0.0865)	(0.0862)
MatNat	0.017	-0.011	-0.009	-0.009
	(0.0211)	(0.0199)	(0.0198)	(0.0194)
MatNat×female	-0.004	0.018	0.009	0.013
	(0.0232)	(0.0222)	(0.0219)	(0.0215)
share of males		0.049		
		(0.0370)		
share of males ×female		-0.029		
		(0.0476)		
desire to have children			-0.009	
			(0.0123)	
desire too have children ×female			-0.021	
			(0.0168)	
childage				-0.003
				(0.0040)
childage×female				0.002
				(0.0048)
Control variables	Yes	Yes	Yes	Yes
Designation subsample	Yes	No	No	No
Female-EngComp Effect	-0.012	-0.028	-0.032	-0.024
R^2	0.079	0.089	0.088	0.089
N	10808	13177	12716	13129

Notes: Cluster robust standard errors in paranthesis. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Sensitivity Check

	(1)	(2)	(3)	(4)
EngComp	-0.097***	-0.100***	-0.309***	-0.317***
	(0.0186)	(0.0188)	(0.0487)	(0.0492)
EngComp×female	0.063**	0.069**	0.183**	0.200**
	(0.0237)	(0.0243)	(0.0631)	(0.0646)
female	0.017	0.016	0.017	0.011
	(0.0127)	(0.0130)	(0.0352)	(0.0359)
MatNat	-0.031	-0.031	-0.096	-0.098
	(0.0251)	(0.0251)	(0.0680)	(0.0679)
MatNat×female	0.011	0.011	0.055	0.059
	(0.0275)	(0.0275)	(0.0770)	(0.0771)
At least one child born before graduation		-0.025		-0.165
		(0.0437)		(0.1179)
At least one child born before graduation×female		0.025		0.141
		(0.0516)		(0.1425)
EngComp×At least one child before graduation		0.060		0.207
		(0.0685)		(0.1766)
EngComp×At least one child before grad.×female		-0.185+		-0.422
		(0.0958)		(0.2669)
Control variables	Yes	Yes	Yes	Yes
Specification	Changed middle category		5-Likert scale	
R^2				
Female-EngComp Effect	-0.034	-0.030	-0.125	-0.117
R^2	0.097	0.097	0.133	0.133
N	13181	13181	13181	13181

Notes: Cluster robust standard errors in paranthesis. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Heterogeneity within Wave 1 and 2

	(1)	(2)	(3)	(4)
EngComp	-0.111***	-0.114***	-0.046**	-0.045**
	(0.0200)	(0.0201)	(0.0165)	(0.0166)
EngComp×female	0.062*	0.066*	0.042+	0.044*
	(0.0260)	(0.0264)	(0.0218)	(0.0222)
female	0.006	0.004	0.017	0.018
	(0.0143)	(0.0146)	(0.0121)	(0.0122)
MatNat	-0.050+	-0.051+	0.037	0.037
	(0.0269)	(0.0269)	(0.0249)	(0.0249)
MatNat×female	0.023	0.025	-0.003	-0.003
	(0.0285)	(0.0286)	(0.0275)	(0.0275)
At least one child born before graduation		-0.061		0.013
		(0.0524)		(0.0485)
At least one child born before graduation×female		0.041		-0.027
		(0.0613)		(0.0568)
EngComp×At least one child before graduation		0.086		-0.031
		(0.0772)		(0.0642)
EngComp×At least one child before grad.×female		-0.092		-0.059
		(0.1427)		(0.0980)
Control variables	Yes	Yes	Yes	Yes
Wave	1	1	2	2
Female-EngComp Effect	-0.049	-0.048	-0.004	-0.001
R^2	0.093	0.094	0.087	0.088
N	6479	6479	6702	6702

Notes: Cluster robust standard errors in paranthesis. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Mediator Analysis

	(1)	(2)	(3)	(4)	(5)
EngComp	-0.067***	-0.052***	-0.070***	-0.055**	-0.082***
	(0.0150)	(0.0153)	(0.0152)	(0.0178)	(0.0177)
EngComp×female	0.038+	0.044*	0.045*	0.045+	0.061**
	(0.0200)	(0.0206)	(0.0202)	(0.0233)	(0.0233)
female	0.055	0.043+	-0.005	0.130	0.011
	(0.0419)	(0.0227)	(0.0226)	(0.0797)	(0.0134)
Control variables	Yes	Yes	Yes	Yes	Yes
MatNat and interaction	Yes	Yes	Yes	Yes	Yes
At least one child before graduation and interactions	Yes	Yes	Yes	Yes	Yes
Mediators×female					
study for labor market	Yes	No	No	Yes	No
job corresponds to study	Yes	No	No	Yes	No
value of degree	Yes	No	No	Yes	No
in retrospect	Yes	No	No	Yes	No
application process	No	Yes	No	Yes	No
difficulties with job search	No	No	Yes	Yes	No
firm size	No	No	No	Yes	No
subjective ability	No	No	No	Yes	No
Female-EngComp Effect	-0.029	-0.008	-0.025	-0.010	-0.021
R^2	0.117	0.111	0.113	0.149	0.089
N	11280	11111	11212	8058	8058

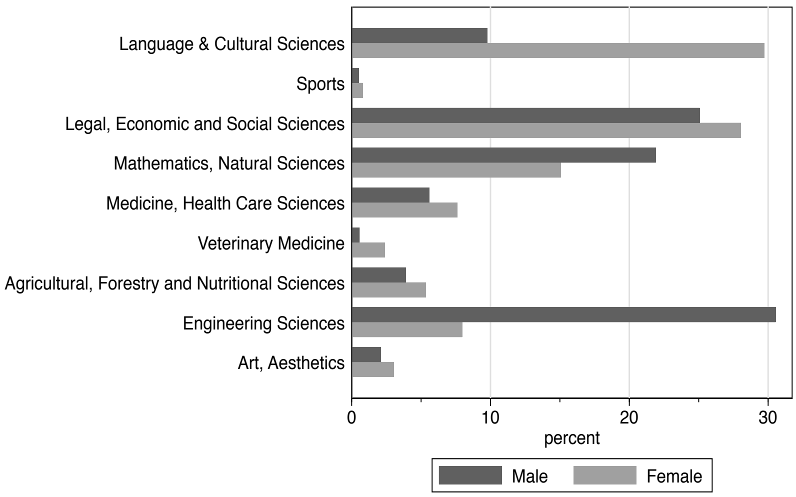
Notes: Cluster robust standard errors in paranthesis. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Supplementary Appendix

Distribution of subject groups and field of studies

Figure 3 gives an overview of the distribution of the nine subject groups across gender and reveals that huge gender differences exist. Only around 8% of women graduate with a degree in Engineering, compared to around 30% of men. A similar trend can be seen in Mathematics and Natural Sciences. Even though the difference is not as big as in Engineering, it is still considerable: Only 15% of women compared to 22% of men graduate with a degree in this field. This gives a first hint to a potential underrepresentation of women in these occupational fields. The lack of role models and the fear of discriminatory work environments might consequently discourage women from actually entering the occupational field. By contrast, figure 3 also identifies subject groups that seem to be typically feminine, such as the field of Language and Cultural Sciences. For the analysis, we will group the fields of study of math and natural sciences to MatNat and engineering and computer sciences to EngComp.

Figure 3: Distribution of Degree Subject Groups across Gender



Note: The graph shows the distribution of subject groups across gender. Except for the differences in Sports and in Legal, Economic and Social Sciences, all differences are statistically highly significant at the 1-percent level ($p < 0.01$). Source: DZHW Graduate Panel 2005 and 2009, own calculations.

Table 12 gives more detailed insights into the distributions of the degree subjects across gender. It becomes especially clear that the situation within the different STEM fields is much more nuanced as suggested by the broader classification. While almost as many men as women graduate with a degree in Mathematics (1.99% and 2.98%, respectively), a higher share of men (10.67%) than women (1.97%) majors in Computer Science. Male dominance also becomes apparent in all Engineering disciplines, except for Architecture and Spatial Planning. In the Natural Sciences, on the other hand, the numbers reverse in some degree subjects: Biology, Pharmaceutics, and Geosciences seem to be more popular among women than among men.

Physics and Astronomy, however, see a higher share of male graduates than females. With respect to the the research question of this paper, it seems to be possible that the different shares of female and male graduates in the respective STEM subfields may also influence female transition behavior. It will, therefore, be of great importance to check for a possible heterogeneity of effects for the specific fields.

Table 12: Degree Area by Gender

	Gender		
Area of Study	Male %	Female %	Total %
Mathematics			
Mathematics, natural sciences, generally	0.07	0.10	0.09
Mathematics	1.92	2.88	2.49
Computer Science			
Computer science	10.67	1.97	5.52
Natural Sciences			
Physics, astronomy	2.79	0.55	1.47
Chemistry	2.65	2.13	2.34
Pharmaceutics	0.29	1.45	0.98
Biology	1.68	3.93	3.02
Geosciences (without geography)	0.29	0.59	0.47
Geography	1.58	1.41	1.47
Engineering			
Engineering in general	1.06	0.07	0.47
Mining, metallurgy	0.14	0.02	0.07
Mechanical engineering, process engineering	12.46	2.63	6.64
Electrical engineering	6.81	0.42	3.02
Traffic engineering, nautical science	2.15	0.20	0.99
Architecture, interior design	2.27	2.70	2.53
Spatial planning	0.32	0.27	0.29
Civil engineering	3.99	1.04	2.24
Surveying	1.43	0.62	0.95
Others			
Linguistics and cultural sciences, generally	0.59	2.45	1.69
Prot. theology/religious doctrine	0.47	0.52	0.50
Cath. theology/religious doctrine	0.54	0.64	0.60
Philosophy	0.29	0.35	0.32
History	1.76	1.32	1.50
Library science, documentation, publishing	0.91	2.00	1.56
General and comparative literature studies and linguistics	0.29	0.89	0.64
Classical philology, modern Greek	0.07	0.10	0.09
German studies (German, Germanic languages without Anglistics)	1.45	5.73	3.99
Anglistics, Americanistics	0.57	2.52	1.72
Romanistics	0.16	1.13	0.74
Slavistics, Baltistics, Finno-Ugristics	0.00	0.05	0.03
Non-European linguistics and cultural studies	0.04	0.07	0.06
Cultural studies in a wider sense	0.07	0.32	0.22

Psychology	1.24	3.87	2.80
Educational sciences	1.00	5.28	3.53
Special pedagogy	0.32	2.49	1.61
Sports, sports science	0.52	0.79	0.68
Business and social studies, generally	0.57	1.13	0.91
Political sciences	1.76	1.21	1.43
Social sciences	1.47	2.39	2.02
Social services	1.40	8.00	5.31
Legal studies	0.45	0.62	0.55
Administrative science	0.72	0.54	0.61
Economic sciences	13.31	12.67	12.93
Industrial engineering	5.35	1.62	3.14
Health sciences, generally	0.64	1.33	1.05
Human medicine	4.07	5.22	4.75
Dental medicine	0.86	1.04	0.96
Veterinary medicine	0.57	2.37	1.64
Landscape management, environmental design	1.24	2.07	1.73
Agricultural sciences, food and beverage technology	1.97	1.64	1.77
Forestry, wood industry	0.57	0.35	0.44
Food sciences and home economics	0.11	1.27	0.80
Art, aesthetics, generally	0.18	0.73	0.50
Fine art	0.07	0.10	0.09
Design	0.90	1.31	1.14
Performing art, film and television, theatre studies	0.11	0.26	0.20
Music, musicology	0.84	0.64	0.72
Total	100.00	100.00	100.00

Note: The table shows the distribution of degree subjects across gender. Source: DZHW Graduate Panel 2005 and 2009, own calculations. Refers to page 33.

Descriptive insights

Table 13 shows the differences in job inadequacy[7] between both field of study as well as gender. The upper panel displays the gender differences in job inadequacy by STEM and non-STEM degree. As outlined before, female STEM graduates are more likely than their male counterparts to have an unrelated job. While the relatively low shares of both men and women leaving the field (10.11% and 13.79% respectively) as well as the small gender difference might not sound too worrisome, the existence of a gender gap may already be enough to make STEM occupations unattractive for women. Among non-STEM graduates, women are also more likely to have a job unrelated to their field of study. For the STEM-subgroups, the descriptive results in table 13 point to different directions. With 5.43 percentage points, the gender difference in engineering and computer sciences seems alarmingly high. This gender difference might very

[7] Entries into the occupational field that does not match the field of study are proxied with having an inadequate job.

well induce a vicious circle where the existing shortage of female coworkers further discourages female university graduates from taking a position in this field. For the subjects math and natural sciences the gender difference is almost zero.

Table 13: Differences in Job Inadequacy between Field of Study and Gender

Current job is not adequate to field of study	Male %	Female %	Difference
all STEM	10.11	13.79	3.68
Math & Natural Sciences	12.83	12.99	0.16
Engineering & Computer Science	9.39	14.82	5.43
Other fields of study	17.85	17.24	-0.61
Double differences			
all STEM vs. all non-STEM			4.29
Math & Natural Sciences vs. all others			0.77
Engineering & Computer Science vs. all others			6.04

Notes: The table shows the relationship between having a job not adequate to field of study and gender, pooled over both survey waves. Source: DZHW Graduate Panel 2005 and 2009, own calculations.

The lower panel of table 13 gives further insights into the different non-transition behavior of STEM and non-STEM graduates. Not only does it compare female (male) STEM graduates with a female (male) non-STEM graduates, but it also shows the difference of the gender difference. The results in the lower panel of table 13 are the raw descriptive counterparts of the coefficient ρ in equation (1). As outlined in chapter 3, a positive and significant coefficient would indicate that the non-transition rate to STEM occupations is higher for women than for men, relative to all other professional fields. Hence, we expect at least a negative gender difference or a even a negative female effect in the regressions analysis.