Underwhelming results on field of study choices and gender horizontal segregation

Carlo D'Ippoliti*and Svenja Flechtner†

October 18, 2022

Preliminary version; please do not quote or circulate.

Abstract

We study fields of study aspirations of secondary school graduates in Germany. We use data from the German school-leavers panel, administrative data from the German Ministry of Education, and income data from the German micro-census. Our novel dataset allows us to simultaneously test several explanations emphasized in the literature, including the role of students' skills and preferences, the prospective occupations' expected returns and risks, and gender-specific social norms. We find evidence that all explanations considered are statistically relevant but substantially only marginally important, in the sense that they contribute very little to explaining the overall variance of educational aspirations between students. The bulk of the observed variance between boys and girls is left unexplained by the main theses put forward in the economics literature, calling for further research on as yet understudied determinants of gender educational segregation, and possibly more interdisciplinary studies.

JEL Codes: I24, J16, J24

Keywords: Fields of study choice, gender wage gap, wage inequality

*Sapienza University of Rome, Italy. Email: carlo.dippoliti@uniroma1.it.

†University of Siegen, Germany. Email: svenja.flechtner@uni-siegen.de. Corresponding author.

1 Introduction

Labour market segregation is a significant determinant of individuals' career prospects and of wage inequality, and the reduction of specifically gender-based segregation is a policy priority in the European context (Bettio and Verashchagina, 2009). In so far as occupational segregation depends on educational segregation, young people's decisions on how much and what to study are consequential for the gender pay gap and other measures of gender-based inequality in the labour market. In this paper, we consider educational aspirations of German secondary school graduates.

We use data from the German school-leavers panel 2008, containing a large sample of school leavers: about 27,000 students in their last year of upper secondary education. We combine data on their characteristics, expectations and aspirations with administrative data from the German Statistical Office (on aggregate trends by fields of study) as well as with our own results on new estimates concerning incomes and their variability, from the German Microcensus data. The resulting original dataset allows us to simultaneously consider - for the first time, to our knowledge - a very wide set of possible explanations for the observed field of study aspirations, spanning both individual characteristics and the characteristics of the fields of study.

Indeed, an interdisciplinary literature has developed a number of arguments to explain young adults' choices, developing into largely disconnected streams of research. Within the economics literature, the most widely investigated hypotheses concern: (i) different abilities between boys and girls (the comparative advantage argument); (ii) different preferences and/or expectations concerning the prospective occupations; and (iii) gender-specific social norms. Focusing on the German case, we are able to consider all three sets of explanations simultaneously, while controlling for individual socio-economic and background characteristics.

The students in our sample face a decision on: entering the labour market without further training; going into vocational training; or enrolling into tertiary education within a specific field of study and kind of institution. We use nested logistic regressions to study the intended field of study choices that students declare a few months before leaving secondary school.

We find that all explanations are marginally important, but they contribute little to explaining the overall variance of aspirations among students. Among up to ten possible alternatives the students face, our model correctly predicts 36-37% of individual fields of study aspirations. We run a series of alternative specifications to try to quantify the impact of each sort of argument, finding that background characteristics have the largest weight in prediction. Ignoring field-specific skills reduces the level of correct predictions only by about 2 percentage points; the reduction is even smaller when we ignore expected returns (and their variability), preferences, or social norms.

Furthermore, even accounting for all these explanations, a gender dummy variable remains statistically highly significant, and using it as the only explanatory variable results in a predictive power comparable to the more sophisticated models. Indeed, even dropping it from the more complete models leads to high losses in our predictive ability. This means that substantially more research efforts should be devoted to explain this as yet unexplained residual, possibly in terms of factors such as school effects, teacher effects, or behavioural aspects such as students' personality or the availability of role models.

The paper is organized as follows. Section 2 provides information about the institutional background of secondary and tertiary education in Germany, as well as descriptive statistics about gender segregation in tertiary career choices. Section 3 discusses the literature on gendered career choices. Section 4 introduces our dataset and empirical model. Section 5 presents and discusses our results, and section 6 concludes.

2 Institutional background and gender labour market segregation in Germany

2.1 The German university system

In this paper, we are interested in educational choices students make after finishing upper secondary education with a university entrance permit. This section explains relevant details of the secondary and tertiary educational systems in Germany.

There are two main types of universities, and two different, corresponding entrance permits in Germany. The general university entrance permit (*Abitur*) is obtained after 13 years of schooling and opens the door to studying at regular universities (*Universität*) as well as at universities of applied sciences (*Fachhochschule*, FH henceforth). After 12 years of schooling, the Fachabitur offers the possibility to study at universities of applied sciences only. While *Abitur* and *Fachabitur* do not pose, in principle, any limits on the field of study a student can choose in tertiary education – although there might be entrance constraints imposed by the universities, for instance in terms of grades required for specific fields –, there is further the possibility of obtaining field-specific entrance permits. Different pathways lead to such entrance permits. For example, secondary degrees obtained at academies and colleges that specialise in specific fields, such as administration or education, only give entrance to tertiary education in these fields. Further, students with a vocational degree can obtain a university entrance permit in the field of their degree.

Overall, we distinguish four types of university entrance permits: entrance permits to either university and FH or FH alone, and both types can be either general or field-specific. In 2021, 20.4% of all university entrance permits obtained were limited to universities of applied sciences (22.5% among male students, 18.6% among female students), while 79.6% of all degrees gave entrance to universities (77.5% among male students, 81.4% among female students) (Statistisches Bundesamt, 2022a).

2.2 Field of study segregation by gender

In winter term 2020/21, women and men had roughly equal shares among first semester students, across all university types. While women have been slightly over-represented (with 50-51%) at universities for years, they represented only 46% at universities of applied sciences in 2021, up from 37% in 2004. Relative to women's shares among school-leavers with a university entrance certificate, however, they are *under*-represented at universities: in 2021, 55.3% of all general university entrance

Table 1: Share of female students in different fields of study and average incomes of job starters

Field of study	Percentage of women, 2021	Income of job starters in €, 2018
Humanities	67.5	34,700
Agriculture, environment, nutrition	66.7	40,900
Law, economics and social sciences	58.0	45,100
Mathematics and sciences	52.2	47,600
Medicine and life sciences	70.9	48,850
Engineering	28.0	51,350

Source: own table, based on Statistisches Bundesamt (2022b) for student numbers and Statista (2018) for income data. Incomes are average gross earnings. Categories do not overlap exactly.

permits were obtained by female students, and hence their transition rate to university studies is lower than that of men. The same holds true for students with FH-only permit, which is roughly at gender parity (Statistisches Bundesamt, 2022a).

Besides this vertical gender segregation in transition to university, there is considerable horizontal segregation: table 1 shows the gender distribution in different fields of students who started their degrees in summer term 2021. As average earnings of job starters in these fields show, this segregation is associated with considerable wage inequality. Overall, as the share of women in a field goes up, average income goes down, with the exception of medicine. While the share of women in careers such as engineering, math and natural sciences has increased over the past, there is still considerable horizontal segregation with bearing on later income earnings.

3 Literature

The literature has considered different explanations of the vertical and horizontal inequalities in career choices and earnings outcomes documented in section 2. Among them, the most prominent ones are (1) gender-specific abilities and comparative advantages, (2) gender-specific preferences and expectations, (3) the internalization of gender-specific social norms and stereotypes, and (4) (anticipated) discrimination and prejudice.

3.1 Abilities and comparative advantage

Many studies have documented how girls and boys tend to develop different comparative advantages and personal strengths in terms of academic abilities in different fields. According to these studies, boys tend to show comparative advantages in math- and science-related fields, whereas girls are stronger in languages (see e.g. Kahn and Ginther, 2017, for a recent overview with a focus on the US). Evidence of these gender differences in abilities usually comes from large international studies such as the OECD's Programme for International Student Assessment (PISA) study. We consider data from the latter study for Germany because it tests students at the age of 15, whereas some other international tests like the German TIMSS study worked with students at the end of primary school only, and gender differences in abilities tend to vary considerably with student age (e.g. Winkelmann et al. 2008). In the 2007 PISA study in Germany, average scores of boys and girls showed a statistically significant difference in science-related competencies and in reading, but not in mathematics. Germany was amongst the countries where boys outperformed girls markedly when it came to explaining phenomena scientifically (by 21 score points; OECD average was 15 score points) (OECD, 2009). There were also strong differences in science-related abilities within schools. [present more data]

Even though differences in abilities are a dominant explanation of gender-specific career choices in economics, the comparative advantage argument is not undisputed, for a number of reasons. First, some authors have questioned the empirical existence of gender ability gaps. Stoet and Geary (2018) refer to 2016 PISA data to make the point that girls perform better or as good as boys in science-related tasks in two thirds of the countries. Guiso et al. (2008) argue, using PISA data as well, that the gender math gap has disappeared in more gender-equal countries. Fryer Jr and Levitt (2010) failed to detect a gender gap in math in Muslim countries. Nevertheless, in these countries, fields of study choices are segregated. This calls into question whether boys and girls simply sort themselves into different fields according to their abilities. Considering international comparative data, Stoet and Geary (2018) find that enrollment-wise, girls are underrepresented in university STEM fields relative to their tests scores, suggesting that there must be something else than abilities driving fields of study choices. Using US data, Wang et al. (2013) found that students with high mathematical and high verbal skills are less likely to go into STEM fields than students with a strength in mathematics alone. Since this was particularly true for female students, they interpret that females, on average, had more options, implying that the lower share of females in STEM fields is not due to their lack of abilities (see also Ceci et al., 2009). Aucejo and James (2021) study the formation of verbal and mathematical skills among students in England over the duration of compulsory education, and find that women developed large advantages in verbal skills, whereas there were only small gender differences in mathematical skills.[1]

Second, it has been questioned whether test score differentials, where they are found, really reflect differences in abilities. There are doubts both about abilities measured through tests and through self-assessments. Niederle and Vesterlund (2010) argue that test score differentials indicate differences in performance in competitive test-taking environments, rather than in abilities (see also Mann et al., 2015). Beilock et al. (2010) conducted a study with US-American school students showing that math-anxious female teachers could transmit their math-anxiety to female students, reducing their test performance.

Another group of studies has compared grades and test scores with students' self-assessments and emphasized that ability gaps may appear to be large when boys and girls diverge in their subjective assessments of their abilities, relative to their grades or test scores. Boys tend to overestimate their abilities in mathematics relative to their grades, as opposed to girls (for studies using German data, see Goetz et al., 2013; Weinhardt, 2017). Gender stereotypes about boys being more talented in mathematics and girls in languages are pervasive in students from early ages on, as a consequence of parental and teacher attitudes, and influence students' self-concepts (see e.g. Tiedemann, 2000). If this is so, it might well be the case that subjective assessments of one's own abilities are an important driver of fields of study choices, rather than abilities as such. Lörz et al. (2011) analyse German survey data and find that for the choice of technical majors, subjectively perceived abilities

[1] In their study, men were much more sensitive to skills: they responded more positively to mathematical skills and enrolled in STEM more often, and reacted strongly negatively to worse verbal skills.

matter much more than objective ability. When girls underestimate their abilities in certain fields, this may create gender segregation even despite objective abilities. Another study using role model interventions also provides support to the idea that subjective assessments of abilities may prevent girls from going into STEM fields, here specifically because of a lack of role models. Breda et al. (2018) analyse the effect of classroom interventions of female role models with a science background with regards to female students' subsequent high-school course choices in France. The intervention was able to raise the probability of girls to enrol in STEM courses in the subsequent school-year by 30%, and by 20% for male-dominated STEM courses. Among other effects that may have channelled this outcome, girls' math self-concept had increased.

Finally, it should be mentioned that recent research has emphasized how gender gaps in abilities, where they exist, may be the result not of innate gender differences in cognitive abilities, but of gender-specific upraising and socialisation (OECD, 2009). For example, Hyde and Linn (2006) study cognitive abilities and find that these do not vary systematically between boys and girls, and therefore argue that factors other than cognitive abilities must be at the origin of different performances of girls and women in mathematical fields, whether in tests or career-wise (see also Spelke, 2005). Brenøe (2018) showed how Danish girls without male siblings developed stronger abilities in STEM fields as compared to girls with brothers, and interprets that gender-specific parenting is responsible for some girls developing a different set of skills than boys: when no sons are available in a family, girls are confronted with rather male activities, too. Likewise, many parents tend to expect better performance in math from sons than from daughters (OECD, 2015).

3.2 Preferences and expectations

In the 2007 PISA study, Germany was among the countries where more boys than girls stated a future-oriented motivation to learn science because they might be interested in science jobs (OECD, 2007). Does this reflect a difference in preferences of females and males, which may then explain fields of study choices?

There has been an important tradition in economics that links career choices of girls and boys to work-family goals and gender-specific preferences. Starting with Becker (1981, 1985), it has been argued that men and women maximize their utilities by specializing in market-activities (wage work) and domestic production (care work), respectively. In such a setting, due to increasing returns to market activities, girls who (expect to) specialize in care work at least partly – for instance, they could expect to work part-time for longer parts of their working life –, would not rationally invest in careers that are incompatible with being a primary caregiver, but follow their preferences for high involvement in child-rearing and care work (Hakim, 2002). Careers could be incompatible with being a primary caregiver because, for instance, part-time work is unusual, or because part-time work is (believed to be) incompatible with demanding careers. Such demanding careers, in turn, are typically better-paid positions and careers in science and management.

Empirically, however, such preferences for specialization have not found unambiguous support. Using US data, Morgan et al. (2013) study how fields of study choices relate to work-family goals and occupational plans. They criticise the above-mentioned literature for being ‘based entirely on convenience samples’ that are used to interpret the existence of ‘stable and deeply ingrained preferences for caring or nurturing [in women], that, when coupled with beliefs about the incompatibility of science with caring or nurturing, make women less likely to choose STEM occupations' (p. 992). Their own study found no gender differences in stated work-family goals, but in occupational plans (i.e. jobs in which the students wanted to work). These gendered job aspirations can easily be sorted by being more or less comfortable for primary caregivers. Nevertheless, given their failure to detect gendered work-family preferences, the authors argue that girls do not choose these jobs because they have a preference for work-family compatibility, but because they have a preference for the job itself.[2] For the case of Germany, Ochsenfeld (2016) did not find evidence for the hypothesis that girls choose fields and jobs that are more compatible with family care work.

If preferences about work-family compatibility are not relevant for girls' career choices, they may still have systematically different preferences regarding other characteristics of fields of study than boys. Here, the literature has differentiated pecuniary and non-pecuniary factors. As to the former, Ochsenfeld (ibid.) found that vocational preferences are a dominant explanation of career choices of German school-leavers. Because these preferences are highly gendered, strong segregation results. While his study is able to differentiate preferences and the conformance to social expectations – and finds that social approval is much less important than stated preferences–, it often remains unclear to what extent male and female students follow their preferences, and to what extent they conform to social expectations and gender norms (see next subsection). The experiment by Breda et al. (2018), in which the

[2] It might well be the case that jobs are more compatible with family life as a consequence of more women working in them.

effects of role model interventions on French high-school students' course choices were analysed, is another study that corroborates the idea that gendered preferences have a role to play: the students' interest in science-related careers was raised by the presence of the role models, and those with a professional background were more effective than young researchers in STEM fields. This suggests that girls caught interest in STEM courses because they could relate these to jobs that corresponded to their preferences. The female role models worked in the cosmetics industry, which is arguably well compatible with feminine stereotypes. Jobs in the cosmetics industry are compatible both with gendered job preferences and with social expectations, where women are expected not to go into masculine fields. It is thus not completely straightforward to know whether girls caught in interest because these jobs meet their preferences, or because they allowed them to align with social norms.

Another gender-specific difference in preferences that has received attention in the literature relates to competition in fields and jobs. According to Buser et al. (2017), girls prefer to be in less competitive jobs, which predicts their less frequent choice of math-intensive fields, because these tend to be more competitive. It could also be the case that women do not want to be the only or among the very few women in male-dominated fields and professions. In a study among French students enrolled in highly selective and competitive high schools, Landaud et al. (2020) found that girls in competitive environments turned away from the competitive science fields, whereas the choices of boys were not affected.

Preferences can also refer to expected income earnings associated with different fields of study. The evidence is ambiguous here. Using US data, Arcidiacono (2004) identifies large income premiums of some college majors compared to others. However, using data from the US National Longitudinal Survey's school leavers in 1972, he finds that these were not driving students' choices, but field and job preferences were. Using US data as well, partly from the same survey, Montmarquette et al. (2002) found that expected income was an essential determinant of college major choices, but with considerable differences by gender and race. Women were less sensitive to anticipated income, while non-whites were more sensitive than whites. Zafar (2013) also found that among Northwestern undergraduates, males valued pecuniary outcomes of majors much more than females. Using data for Duke students, Arcidiacono et al. (2012) found that students' choices were informed both by the income earnings they expected, given their field of study, and their subjective assessment of their abilities in their field of choice. For French students, Beffy et al. (2012) analysed that expected income played a minor role in major choices for both genders, but that non-pecuniary factors were valued more. Boudarbat (2008) studied major choices of community college students in Canada (cohorts 1990 and 1995) and found that the students' college major choices had been guided by expected earnings, especially if they had been on the labour market before. Women were less sensitive to expected income levels. [3]

Gender differences related with expectations may also stem from different assessments of labour market outcomes and returns to an academic education in general, as well as to particular fields of study. Lörz et al. (2011) studied expected labour market

[3] It is not always clear how students take expected income levels into account when making their choices. For example, it could be that male students are drawn into rather 'male' fields (see next section), while girls go into presumably feminine fields. The latter can be associated with lower expected income, even when income considerations are not dominating girls' choices.

outcomes, subjective success probabilities and expected study costs and of German school leavers and found that vertical segregation at the transition from secondary to tertiary education was due to different perceived returns and labour market outcomes. Comparing male and female students of the same ability (measured through test scores), females tended to have a relatively lower subjective assessment of their abilities and success probabilities. Another results was that girls were more sensitive to the cost associated with a university degree. Further they had lower expectations of their labour market outcomes and returns to a university degree as compared to a vocational training, such that a higher share of women expected to maximize their outcomes through vocational training.

If students may well base their choices on beliefs about expected incomes, these beliefs need not be accurate. There is a literature showing how student beliefs about expected incomes of college degrees and different fields of study are often wrong (could elaborate).

3.3 Social expectations and norms

The role of social expectations and norms has been brought into the fields of study choice literature in various ways. The common point of departure is that some fields of study are commonly perceived as rather masculine and others as rather feminine. There is vast evidence illustrating how boys and girls, as of young ages, associate different fields and tasks with gender, for example when they associate math skills with boys and language skills with girls (Cvencek et al., 2011; Steffens and Jelenec, 2011; Tiedemann, 2000). These internalized stereotypes may then influence their self-concepts and fields of study choices.

Based on this, one idea has been that students do not simply follow their own preferences, but also seek to fulfil stereotyped gender roles and social expectations related to them, for example to obtain approval by parents and peers (see e.g. Nollenberger et al., 2016). For the case of Germany, however, Ochsenfeld (2016) does not find support for this hypothesis. In his study, peers were indeed critical of gender-untypical choices (more often than parents), but these reactions had very little explanatory power for the choices students made.

Regardless of other people's approval or criticism, social norms and expectations may operate through making boys and girls themselves feel belonging into specific fields or not. For instance, girls may be more skeptical about themselves belonging into allegedly masculine fields. Some studies have argued that when such feelings are present, girls may be more sensitive to low grades and rule out STEM careers more quickly as a consequence of low grades than boys. Kugler et al. (2017) studies sensitivity to low grades in a sample of US students and does not identify general behavioural differences between boys and girls, but girls were more sensitive when they obtained low grades in a supposedly masculine field. Girls dropped out more quickly of male-dominated STEM fields as a consequence of low grades, leading the author to interpret that when several signals of "lack of fit" combine, girls react to them. Recent studies on the effects of female role models from STEM fields provide some additional evidence illustrating how a perceived lack of fit may indeed be an important determinant of girls' choices (Breda et al., 2018).

Social norms may have repercussions on other factors that influence fields of study choice. For example, women may be less sensitive to anticipated income earnings (see above) because of the so-called breadwinner norm. One reason for girls apparently caring less about expected income could be that they do not see themselves in the responsibility of being a primary wage earner who sustains a family with a dependent partner.

3.4 Anticipated prejudice and discrimination

Labour market discrimination against women, for instance through wages but also through hiring or promotion decisions, has been an important explanation of different labour market outcomes of men and women. How women are still discriminated against on the labour market nowadays is a matter of debate. While some studies reject the idea that women are discriminated against in STEM and math-intensive fields (Ceci and Williams, 2011), others finds discrimination at least in traditionally male-oriented high-qualified fields Neumark (2016). The important question for our study is whether girls anticipate prejudice and discrimination or not, regardless of whether they have good reasons to expect such experiences or not.

Here, the evidence is scarce and ambiguous. Using German data, Ochsenfeld (2016) did not find support for the hypothesis that girls avoid professions in which they expect discrimination. In contrast, the role model study by Breda et al. (2018) suggests otherwise. As an unintended effect of the female role model interventions in French classrooms, the under-representation and potential discrimination of women in STEM was made more salient to the students, even though they could decrease their own stereotypes about women in STEM. The researchers interpret that this is the reason why only high-achieving students were steered towards STEM fields after the intervention, while low-achieving students were incited to fear discrimination. Folke and Rickne (2022) found that women and men in the US are deterred from jobs where they are the gender minority and relate this to risk of harassment.

4 Data and model

4.1 Data

Table 2: Share of students by potential and actual choices

Choice	Field of study	Potential choices	Times actually chosen	% of potential
University	Humanities	19,699	3,530	17.9
	Life Sciences	19,677	1,265	6.4
	Social Sciences	19,761	1,690	8.6
	Natural Sciences	19,738	2,348	11.9
FH	Humanities	27,316	691	2.5
	Life Sciences	27,294	315	1.2
	Social Sciences	27,378	2,480	9.1
	Natural Sciences	27,355	2,277	8.3
Vocational training		27,634	6,939	25.1
No further study	-	27,634	988	35.8
Doesn't know	-	27,634	5,111	18.5
Total		271,120	27,634	100

We use panel data from a survey of secondary school-leavers with a university entrance permit (*Studienberechtigtenpanel*) provided by the German Centre for

Higher Education Research and Science Studies (DZHW). We use the first wave, collected in 2008, in which final-year students in upper secondary schools were asked about their educational aspirations and plans for their futures.

Students enrolled in the last year of upper-secondary school were asked what, if any, tertiary studies they were seeking for the next year. They were asked to indicate the specific career (e.g. veterinary medicine, archaeology, or law) as well as the type of college (university or FH). Table 2 reports the fields of study aspirations students in our sample report with regards to their post-secondary life. We compare students' stated aspirations with their potential choices. Potential choices comprise all those options a person could access, given the scope of their university entrance permit (see section 2).

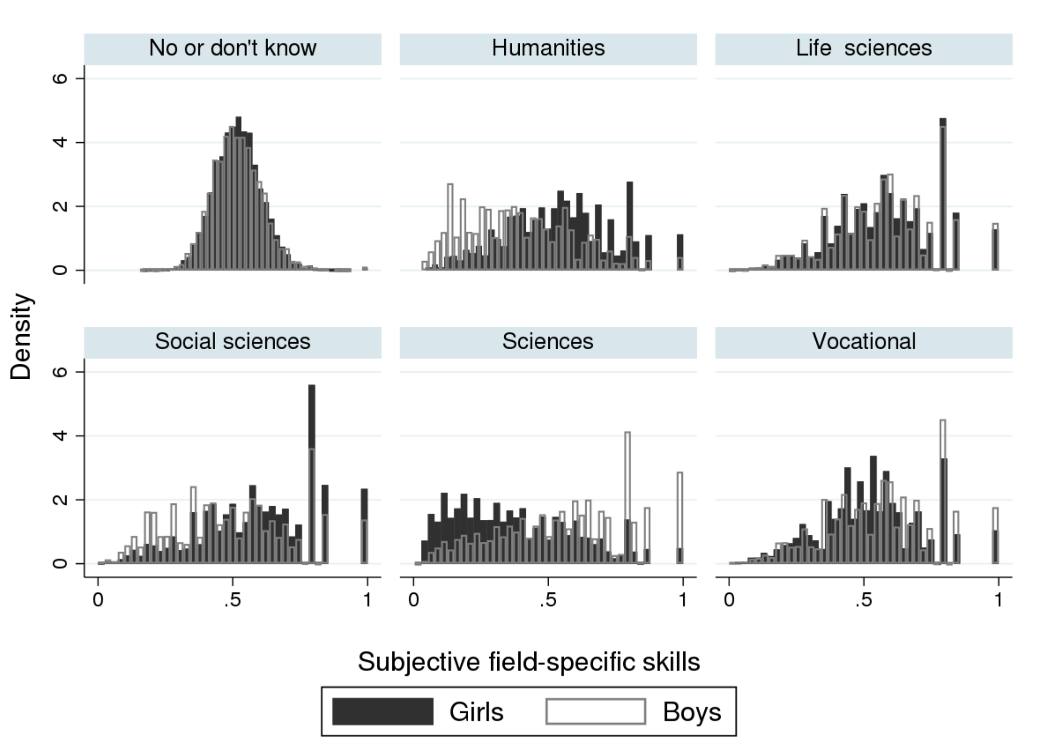
To analyse the determinants of the students' choices, we consider the following potential explanatory factors:

- **Individual characteristics**: field-specific self-assessed abilities, school grades, and preferences
- **Field characteristics**: field feminization, expected income
- **Control variables**: gender, migration background, socio-economic family background, information by parents, limited choice through bad grades

Individual characteristics

Self-assessed field-specific skills For the skills involved or required within each field, we use a question on students' self assessment on their relative skills in the social sciences, humanities, natural sciences, life sciences, and craft skills. For each

Figure 1: Distribution of self-assessed field-specific skills, by gender



person, we compute the cumulative frequency of answers 1, 2, 3, 4, 5 in order to standardize by people's self-confidence. Specific skills are then an average of the standardized questions concerning self-perceived strength in each of the five main fields: social sciences, humanities, sciences, life sciences, vocational activities.[4] Figure 1 shows how self-assessed strengths are distributed across fields and gender. Note that self-assessed strengths tend to overlap in a range of fields, but diverge in the humanities, were more girls perceive a relative strength as compared to boys,

[4]We match: social-communicative and foreign language skills for social sciences; humanities: foreign languages, literature, and arts and music; sciences: technical, mathematics and natural sciences; life sciences: natural sciences and social-communicative; vocational: social-communicative and handcraft.

and in the sciences, where the opposite is the case.

Grades Our survey data contain information about the average grade students had obtained in the current year. School grades were made comparable across school types by converting them into centiles of the empirical grade distribution for each type of school.

Preferences For a number of secondary school disciplines (German, English, and mathematics), students were asked how much they found the respective courses stimulating, multifaceted, understandable, and comprehensible. On this basis, we take the average of how much they liked German, English, and mathematics and create a dummy variable for math inclination in case people liked mathematics more than both English and German, and 0 otherwise.

Field characteristics

Further, we consider three characteristics pertaining to the choices available, rather than to the individual: expected income earnings associated with a field, their variability, and field feminization.

Field-specific expected returns Field-specific expected returns are considered in terms of expected income. Using data from the German Microcensus, we estimated the current incomes of German workers, controlling for their level and field of study, age, and a set of other individual and household characteristics (since only brackets of income levels are provided, we run interval regressions). We then computed a gender-specific rate of return to education by field of study, and the associated variability – which we interpret as a measure of income risk. As a robustness check, we alternatively computed average gender- and field-specific predicted incomes for each year of age, and used these values to estimate a Net Present Value (NPV) of each field of study choice, by gender.

Field feminization The feminization of the various fields was computed by considering the share of workers in the Microcensus who hold a university degree in a certain field. As a robustness test, we also considered the share of women among full professors for each field of study (data from Destatis (2008)), using the previous measure (share of workers in the Microcensus) for the field "no further study".

Control variables

Migration background Students reported evidence of a possible migration background in their family both by reporting on the languages spoken at home and their nationality; in the following analyses we consider the former, but results do not significantly change when using the latter. Students are considered as having a migration background if German is not the only language spoken at home.

Socio-economic family background Information about the socio-economic family background was obtained by factor analysis on the matrix reporting the tetrachoric correlations among the answers to the following questions (all dummy variables): if the student's father has tertiary level educational attainment, if the mother does, if the student considers family finances an important factor in deciding whether to continue studying, if s/he thought about how to finance her studies, if s/he collected information on financing higher education (through three possible channels), if s/he discussed it with her parents, and if s/he expects them to contribute, and if s/he considers finances a constraints for field choice. [5] The results index has been standardized to take on values between 0 and 1.

We include three further controls as dummy variables. First, we include whether students had discussed their career choices with and received information by their parents. Second, whether students perceived that their career choice was constrained by too low grades. This may be the case for some very selective fields, especially medicine. Even though students may in principle possess the entry requisites in terms of the university entrance permit, some fields select on the basis of the students' final grade in secondary school. In practice, this means that only a small share of all students with a general university entrance permit will be admitted in programs such as Medicine. Third, whether students had already completed a vocational degree prior to obtaining their secondary degree. This could be the case, in particular, for students obtaining their secondary degree in the context of further vocational qualification and obtaining a field-specific entry permit. We control for previous vocational training because the likelihood that these students would complete another vocational degree should be substantially reduced.

4.2 Model

Conditional and multinomial regression models, which many studies in the field have used, rely on the assumption of independently distributed errors: alternative choices are assumed to be independent of each other. In reality, however, the alternative

[5] With the traditional method of taking those variable with an eigenvector > 1, we would have a second factor that could be interpreted as an index of financial concerns. However we decided not to include this in the analysis.

fields of study choices students face are not fully independent; for instance, choosing between business administration and economics is a choice between rather similar alternatives as compared to choosing between economics and biology. Students inclined towards the social sciences will likely make a choice among alternatives from within the social science disciplines, while students with a preference for natural science will choose within this category, and so forth. To account for this nesting structure of all alternative study programs available to students, we estimate students' fields of study choice by a full information nested logistic regression. This model relaxes (and allows us to test) the assumption of independence of irrelevant alternatives (IIA). The nested logistic model does not need IIA because rather than assuming independently distributed errors, it allows for clustering similar alternatives into nests. Note that the IIA assumption is typical of the simple expected utility model, which experimental studies generally regard as unrealistic.[6]

Aggregating study courses by broad field of study, it emerges that students have up to 10 options: no further education; vocational training; or one of the following four fields within a university or FH program: humanities, social sciences, natural sciences, or life sciences. We consider as separate nests the options of no further study and of vocational training; the choices of field within university or university of applied sciences studies are assumed to be correlated within a nest. Choosing not to study any further is our reference category.

Let $S = \{s_1, s_2, s_3, s_4\}$ be the set of kind of study choices, with

[6]For a demonstration of how this model can be derived under the assumption of utility maximization, see Amemiya (1985).

- s_1 = no further study,
- s_2 = undecided,
- s_3 = vocational training,
- s_4 = university of applied sciences (FH),
- s_5 = university.

s_4 and s_5 are nests of field of study choices. Let $F = \{f_1, f_2, f_3, f_4\}$ be the field choices within $S = \{s_4, s_5\}$ with

- f_1 = humanities,
- f_2 = life sciences,
- f_3 = social sciences,
- f_4 = natural sciences.

s_1 to s_3 are 'degenerate' nests: they only contain one field choice.

Let $C_1 = s$ be the choice of kind of study made (first level choice), and $C_2 = f$ the choice of field within the kind of study (second level choice). The probability that a certain field $C_2 = j$ is chosen, given the kind of study choice $C_1 = t$, is

$$Pr(C_2 = j | C_1 = t) = \frac{e^{(x_t \beta_j)}}{\sum_{f \in S_t} e^{(x_t \beta_j)}} \tag{1}$$

where x is a vector of second-level explanatory variables and β a coefficient; and Pr is distributed as a multinomial logistic function. Define the (non-normalized) inclusive function as

$$I_t = ln[\sum_{f \in S_t} e^{(x_t \beta_j)}] \tag{2}$$

Then the probability of the first-level choice is:

$$Pr(C_t = 1) = \frac{e^{\left(z_t a_t + I_t\right)}}{\sum_{s \in S} e^{\left(z_s a_s + I_s\right)}} \tag{3}$$

where z is a vector of first-level explanatory variables and α a coefficient.

5 Results and discussion

We run the nested logistic models of field study choice, taking the decision not to continue studying in any form as the reference category. In all estimations, the standard errors within the nested grouping fields of university studies, and within FH studies, are positively correlated, with correlation coefficients significantly different from zero. Accordingly, the IIA hypothesis is rejected and the baseline model of expected utility maximization does not apply, justifying our use of nested logistic regressions.

Table A3 shows the results of our main estimation on the whole sample. Appendix A provides additional results. To better understand gender differences, tables A1 and A2 provide estimates on the boys' and girls' samples separately.

Further, in the appendix we report the main results from estimations using a different variable for the fields' expected income (namely NPV, as a robustness check, in table 3) and using a gender dummy as the only explanatory variable (table A4).

As shown in table A3, all field-specific characteristics seem to exert a statistically significant effect on students' choices, as well as many individual factors. Concerning the former, unsurprisingly students have a higher probability of choosing a field that matches their subjectively assessed comparative strengths (the odds ratio of field-specific skills is significantly larger than 1). On the contrary, both the feminization of a field and the variability of expected incomes (that we interpret as a measure of income risk) negatively impact on students' choice for that field (i.e., their odds ratios are significantly lower than 1). Finally, expected income too - measured by the field-specific returns to education in a specific field, with respect to choosing not to continue studying - exhibits a negative impact, suggesting that students appear to prefer coeteris paribus those fields of education from which they can actually expect to earn comparatively less in the future.

Table 3: Nested logistic regression on whole sample, odds ratios

VARIABLES	All fields	University of applied sciences (FH)				University				Vocational training	doesn't know
		Humanities	Life sciences	Social sciences	Sciences	Humanities	Life sciences	Social sciences	Sciences		
Field-specific skills (self-assessed)	8.619***										
	[0.353]										
Field-specific NPV of expected incomes	1.000***										
	[3.94e-06]										
Field-specific variability of expected income	0.338***										
	[0.0128]										
Field feminization	0.988***										
	[0.00185]										
Math-inclined = 1		1.485***	1.497***	1.384***	1.415***	1.977***	2.368***	2.067***	1.721***	2.246***	2.410***
		[0.126]	[0.206]	[0.107]	[0.113]	[0.147]	[0.237]	[0.182]	[0.146]	[0.159]	[0.170]
Grade		1.008***	0.991***	1.018***	1.011***	1.020***	1.061***	1.035***	1.035***	1.000	1.019***
		[0.00134]	[0.00246]	[0.00119]	[0.00122]	[0.00114]	[0.00200]	[0.00149]	[0.00129]	[0.00109]	[0.00105]
Woman		2.069***	1.641***	0.969	0.340***	3.440***	1.084	0.527***	0.643***	3.440***	2.118***
		[0.179]	[0.232]	[0.0811]	[0.0336]	[0.273]	[0.133]	[0.0561]	[0.0607]	[0.255]	[0.154]
Migration background = 1		0.576**	1.352	0.743**	0.458***	0.490***	1.162	1.124	1.273*	0.626***	3.148***
		[0.0830]	[0.353]	[0.0956]	[0.0614]	[0.0638]	[0.182]	[0.156]	[0.173]	[0.0772]	[0.365]
Socio-economic background		0.999***	0.999**	0.999***	0.999***	1.000	1.003***	1.001***	1.001***	0.998***	1.000
		[0.000362]	[0.000599]	[0.000328]	[0.000318]	[0.000318]	[0.000436]	[0.000380]	[0.000354]	[0.000302]	[0.000304]

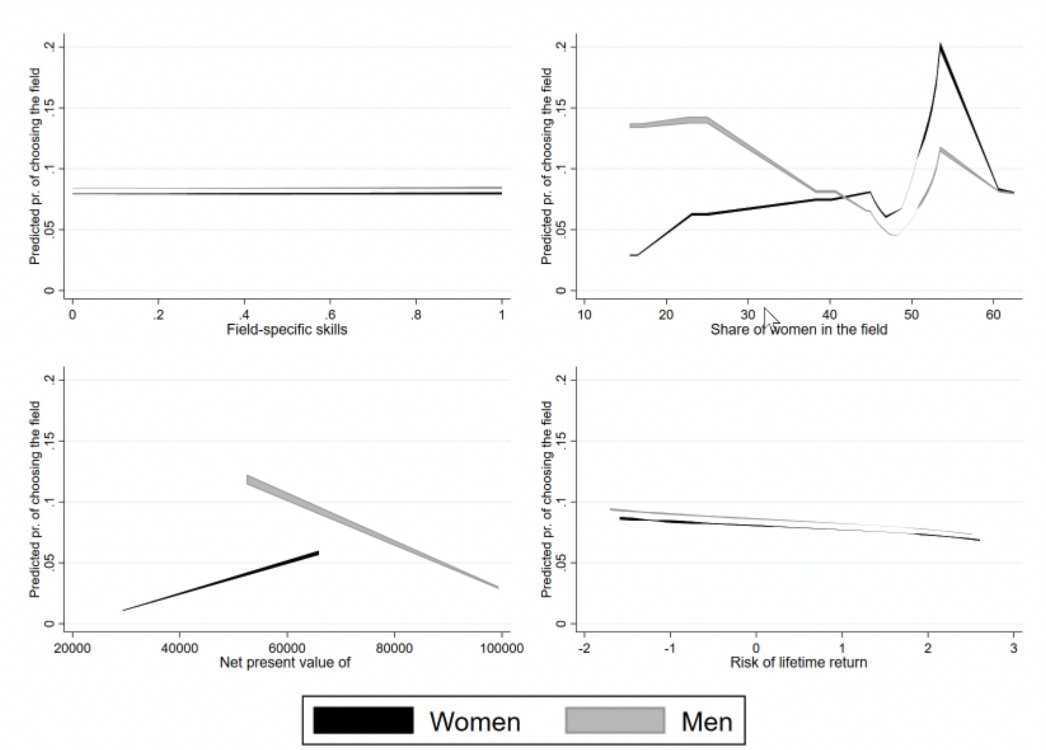
*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Table reports odds ratios. n=27,634. Constants not reported.
Further controls (odds-ratios not reported) include if respondents indicated their parents were an important source of information and advice; if their choice set was constrained by too low grades; and if they had already completed a vocational training.

This last unexpected result should not be given too much thought, however, because as is the case for income risk and field-specific skills, the estimated coefficients are not large enough to produce any relevant trend in terms of predicted student choices. As shown in figure 2, the only variable associated with relevant changes in the students' predicted probability to choose a certain field is the field's feminization. Even that, however, seems to be associated with predicted probabilities in a highly non-linear way that can hardly be conceptualized, except for noting that women tend to prefer highly feminized fields, and men tend to prefer avoiding them as predicted by the literature (but again, with significant exceptions, as visualized by the large swings along both curves). For the other three field-specific variables, statistical significance does not imply substantial relevance.

Concerning individual-level variables, having high grades has almost invariably a positive coefficient, denoting a tendency to continue studying in whatever field; as does a preference for mathematics over English and German. In both cases, the differences between university studies (perceived as more rigorous and prestigious) and FH studies seem to be larger than those between disciplinary fields.

Individual characteristics, such as family socio-economic background, migrant status, and sex, all turn out to be statistically significant despite the control variables included. Migrants tend to shy away from FH (except for the life sciences) and to prefer university studies (except for the humanities). Recalling that we only consider students with a university entrance permit, and that we control for constrained choices due to grades, this result might imply that students with a migrant background - independently of prospective incomes, risk, and feminization - tend to prefer

Figure 2: Predicted probabilities by field characteristics (correlations)



acquiring more education in the fields more traditionally associated with prominent social status. Family socio-economic background appear to exert an opposite effect, with students from better-off families preferring university studies over FH and vocational training. Instead, for girls the choice over disciplinary fields seem to be more salient than that between FH, university, and vocational training. Indeed, ceteris paribus girls prefer to study humanities and life sciences, especially over the hard sciences (however, girls are also more likely to be undecided, or to enter into vocational training).

Sex-specific estimates, reported in tables A1 and A2 in appendix, do not dramatically alter the picture obtained from pooled estimations. As expected given the above-mentioned results on feminization, this variable takes on opposite values between boys (for which it is negative) and girls (positive). The coefficient associated with field-specific skills remains very small for both sexes, while girls seem to be slightly less influenced by their self-perceived skills, and slightly more (negatively) by the riskiness of predicted incomes.

The impact of background characteristics does not qualitatively differ (girls with a lower socio-economic background are less likely to enroll into a FH, boys with a higher socio-economic background are more likely to enroll into a university). Instead, preferences seem to play different roles for boys' and girls' choices. A preference for mathematics more strongly induces girls to continue studying (all coefficients tend to be higher than for boys) and specifically it directs them towards university studies of all kinds (including the humanities). Notably, we do not find that, controlling for field-specific abilities and expected incomes and risks, a preference for math does not induce students to more likely opt for a career in the sciences.

Overall, our main findings denote that many theories reviewed before often find support in the data, in the sense that the associated regressors are statistically significant. However, except for feminization (where anyway the relation with student choices is non-linear and highly variable), none seem to substantially contribute to our understanding of the determinants of horizontal gender segregation. To highlight this point, in table 4 we report how the log-Likelihood and the percentage of correctly predict student choices change, depending on the model specification.

The main comparison is between our comprehensive baseline model, discussed up to here, and a simple model in which only a "woman" dummy variable is included.

As shown in the table, the full model is associated with a lower log-Likelihood and it is the preferred model according to the Akaike Information Criterion. However, the difference in log-Likelihood is not as large as might have been expected, and indeed the simpler, basic model even leads to a higher number of correctly predicted student choices (defined as the cases in which the choices with the higher predicted probability is that which the students answered is their choice). Considering alternative model specifications that exclude a variable each time, in most cases the log-Likelihood and the predictive ability do not change noticeably. The two cases in which excluding a variable leads to a reduction in the ability to predict the correct choice by about 1.5-2 percentage points are field-specific skills and grades (which however, as we saw, impact on university enrollment rather than disciplinary field).

Table 4: Contribution to predictive ability by different explanatory variables

Specification	# obs.	AIC	log likelihood	% correct predictions
Full model (table 3)	27,634	96.699,35	-48.261,67	36.50
Only regressor: dummy woman variable	27,881	114.163,5	-57.066,73	39.33
Full model (table 3), without each of the following:				
Field-specific skills	27,811	100.440,5	-50.133,25	34.53
Field feminization	27,634	96.759,65	-48.291,82	36.45
Field-specific returns to education	27,634	97.051,43	-48.436,72	36.25
Field-specific variability of expected income	27,634	97.495,87	-48.659,94	36.06
Math inclination	27,695	97.259,61	-48.550,80	36.43
Grade	27,634	99.481,92	-49.660,96	34.25
Woman	27,634	98.537,59	-49.189,80	35.16
Migration background	27,634	97.528,92	-48.686,46	36.17
Socio-economic background	27,634	97.196,89	-48.518,45	36.14

In general, it seems that by considering the three most widely discussed economic theories of field of study choice we are unable to adequately capture the determinants of students' choices in Germany.

6 Conclusions

We studied the determinants of German secondary school-leavers' choice about their further education: to go into vocational training or to study in different types of universities and different fields. These choices are highly consequential for their expected income earnings. Hence systematic patterns in the choice processes of women – or other groups – may contribute to the persistence of income earnings inequality. We find that such patterns exist. Overall, our preliminary results indicate that circumstance variables, such as gender, interact with individual characteristics and shape choices in ways that may contribute to the perpetuation of the gender wage gap.

However, we also find that the characteristics of the prospective fields, which together with preferences and individual circumstances should regulate a rational choice between fields, exert a statistically significant but economically negligible impact on student choices.

Our study has some important limitations. Our dataset was obtained from several different sources with the aim to make it as comprehensive as possible. Yet, it does not allow us to consider all possible explanations put forward in the literature. Specifically, due to data limitations we cannot control for school- and teacher-level variables, as well as we must ignore most behavioural explanations. Future research should seek to overcome these limitations.

Nonetheless, we deem it relevant to highlight the limited explanatory power in our sample, of several prominent economic explanations of gender horizontal segregation. In light of the evidence on p-hacking and publication bias against null and negative results in scientific publications, it seems useful to warn economists about complacency over some theories with limited explanatory value. In this sense, at least for the German case our study highlights the risk of some dead ends that might make necessary to consider more interdisciplinary efforts, and in general efforts towards greater creativity in explaining young people's choices. Our negative or almost null results suggest that different strategies of data creation and data collection are needed to explain horizontal gender segregation in a large, relatively advanced economy.

References

Amemiya, Takeshi (1985). *Advanced Econometrics*. Cambridge: Harvard University Press.

Arcidiacono, Peter (2004). "Ability sorting and the returns to college major". In: *Journal of Econometrics* 121.1-2, pp. 343–375.

Arcidiacono, Peter et al. (2012). "Modeling college major choices using elicited measures of expectations and counterfactuals". In: *Journal of Econometrics* 166.1, pp. 3–16.

Aucejo, Esteban and Jonathan James (2021). "The path to college education: The role of math and verbal skills". In: *Journal of Political Economy* 129.10, pp. 2905–2946.

Becker, Gary S (1981). *A Treatise on the Family*. Cambridge, MA: Harvard University Press.

— (1985). "Human capital, effort, and the sexual division of labor". In: *Journal of labor economics* 3.1, Part 2, S33–S58.

Beffy, Magali et al. (2012). "Choosing the field of study in postsecondary education: Do expected earnings matter?" In: *Review of Economics and Statistics* 94.1, pp. 334–347.

Beilock, Sian L et al. (2010). "Female teachers' math anxiety affects girls' math achievement". In: *Proceedings of the National Academy of Sciences* 107.5, pp. 1860–1863.

Bettio, Francesca and Alina Verashchagina (2009). *Gender segregation in the labour market: Root causes, implications and policy responses in the EU*. Luxembourg: Publications Office of the European Union Luxembourg.

Boudarbat, Brahim (2008). “Field of study choice by community college students in Canada”. In: *Economics of Education Review* 27.1, pp. 79–93.

Breda, Thomas et al. (2018). “Can female role models reduce the gender gap in science? Evidence from classroom interventions in French high schools”. In: *PSE Working Papers* 2018-06.

Brenøe, Anne Ardilla (2018). “Sibling Gender Composition and Participation in STEM Education”. In: *University of Copenhagen, Working Paper*. URL: https://pdfs.semanticscholar.org/84ab/89db4dca0808f89853aeed42f3f1b62a79ec.pdf.

Buser, Thomas et al. (2017). “Gender, competitiveness, and study choices in high school: Evidence from Switzerland”. In: *American economic review* 107.5, pp. 125–30.

Ceci, Stephen J and Wendy M Williams (2011). “Understanding current causes of women’s underrepresentation in science”. In: *Proceedings of the National Academy of Sciences* 108.8, pp. 3157–3162.

Ceci, Stephen J et al. (2009). “Women’s underrepresentation in science: sociocultural and biological considerations.” In: *Psychological bulletin* 135.2, p. 218.

Cvencek, Dario et al. (2011). “Math–gender stereotypes in elementary school children”. In: *Child development* 82.3, pp. 766–779.

Destatis (2008). *Bildung und Kultur. Studierende an Hochschulen. Fachserie 11 Reihe 4.1*. Wiesbaden: Destatis.

Folke, Olle and Johanna Rickne (2022). “Sexual harassment and gender inequality in the labor market”. In: *The Quarterly Journal of Economics* 137.4, pp. 2163–2212.

Fryer Jr, Roland G and Steven D Levitt (2010). “An empirical analysis of the gender gap in mathematics”. In: *American Economic Journal: Applied Economics* 2.2, pp. 210–40.

Goetz, Thomas et al. (2013). “Do girls really experience more anxiety in mathematics?” In: *Psychological science* 24.10, pp. 2079–2087.

Guiso, Luigi et al. (2008). “Culture, gender, and math”. In: *SCIENCE-NEW YORK THEN WASHINGTON-* 320.5880, p. 1164.

Hakim, Catherine (2002). “Lifestyle preferences as determinants of women’s differentiated labor market careers”. In: *Work and occupations* 29.4, pp. 428–459.

Hyde, Janet S and Marcia C Linn (2006). “Gender similarities in mathematics and science”. In: *Science* 314.5799, pp. 599–600.

Kahn, Shulamit and Donna Ginther (2017). “Women and STEM”. In: *NBER Working Paper* 23525.

Kugler, Adriana D et al. (2017). *Choice of Majors: Are Women Really Different from Men?* Tech. rep. National Bureau of Economic Research.

Landaud, Fanny et al. (2020). “Competitive schools and the gender gap in the choice of field of study”. In: *Journal of Human Resources* 55.1, pp. 278–308.

Lörz, Markus et al. (2011). “Gender inequalities in higher education: extent, development and mechanisms of gender differences in enrolment and field of study choice”. In: *Irish educational studies* 30.2, pp. 179–198.

Mann, Allison et al. (2015). "The role of school performance in narrowing gender gaps in the formation of STEM aspirations: a cross-national study". In: *Frontiers in psychology* 6, p. 171.

Montmarquette, Claude et al. (2002). "How do young people choose college majors?" In: *Economics of Education Review* 21.6, pp. 543–556.

Morgan, Stephen L et al. (2013). "Feeding the pipeline: Gender, occupational plans, and college major selection". In: *Social Science Research* 42.4, pp. 989–1005.

Neumark, David (2016). *Experimental research on labor market discrimination*. Tech. rep.

Niederle, Muriel and Lise Vesterlund (2010). "Explaining the gender gap in math test scores: The role of competition". In: *Journal of Economic Perspectives* 24.2, pp. 129–44.

Nollenberger, Natalia et al. (2016). "The math gender gap: The role of culture". In: *American Economic Review* 106.5, pp. 257–61.

Ochsenfeld, Fabian (2016). "Preferences, constraints, and the process of sex segregation in college majors: A choice analysis". In: *Social science research* 56, pp. 117–132.

OECD (2007). *PISA 2006: Science Competencies for Tomorrow's World*. Vol. Vol. I: Analysis. OECD Publishing.

— (2009). *Equally prepared for life? How 15-year-old boys and girls perform in school*. OECD Publishing.

— (2015). *The ABC of gender equality in education: Aptitude, behaviour, confidence*. OECD Publishing.

Spelke, Elizabeth S (2005). “Sex differences in intrinsic aptitude for mathematics and science?: a critical review.” In: *American Psychologist* 60.9, p. 950.

Statistisches Bundesamt (2022a). *Schnellmeldungsergebnisse zu Studienberechtigten der allgemeinbildenden und beruflichen Schulen - vorläufige Ergebnisse. Abgangsjahr 2021*. Wiesbaden: Destatis.

— (2022b). *Studierende an Hochschulen - Fachserie 11 Reihe 4.1 - Sommersemester 2021*. Wiesbaden: Destatis.

Steffens, Melanie C and Petra Jelenec (2011). “Separating implicit gender stereotypes regarding math and language: Implicit ability stereotypes are self-serving for boys and men, but not for girls and women”. In: *Sex Roles* 64.5-6, pp. 324–335.

Stoet, Gijsbert and David C Geary (2018). “The gender-equality paradox in science, technology, engineering, and mathematics education”. In: *Psychological science* 29.4, pp. 581–593.

Tiedemann, Joachim (2000). “Parents’ gender stereotypes and teachers’ beliefs as predictors of children’s concept of their mathematical ability in elementary school.” In: *Journal of Educational psychology* 92.1, p. 144.

Wang, Ming-Te et al. (2013). “Not lack of ability but more choice: Individual and gender differences in choice of careers in science, technology, engineering, and mathematics”. In: *Psychological science* 24.5, pp. 770–775.

Weinhardt, Felix (2017). “Ursache für Frauenmangel in MINT-Berufen? Mädchen unterschätzen schon in der fünften Klasse ihre Fähigkeiten in Mathematik”. In: *DIW-Wochenbericht* 84.45, pp. 1009–1014.

Zafar, Basit (2013). “College major choice and the gender gap”. In: *Journal of Human Resources* 48.3, pp. 545–595.

A Additional results

Table A1: Nested logistic regression - girls only

VARIABLES	All fields	University of applied sciences (FH)				University				Vocational training	doesn't know
		Humanities	Life sciences	Social sciences	Sciences	Humanities	Life sciences	Social sciences	Sciences		
Estimation with net present value											
Field-specific skills	7.054***										
	[0.410]										
Field-specific NPV of expected incomes	1.000***										
	[6.80e-06]										
Field-specific variability of expected income	0.224***										
	[0.0148]										
Field feminization	1.020***										
	[0.00246]										
Math-inclined = 1		1.746***	2.525***	1.777***	1.630***	3.163***	3.243***	3.037***	2.345***	3.135***	2.880***
		[0.221]	[0.516]	[0.208]	[0.244]	[0.362]	[0.495]	[0.400]	[0.318]	[0.345]	[0.316]
Grade		1.016***	0.995	1.025***	1.017***	1.035***	1.086***	1.043***	1.048***	1.010***	1.026***
		[0.00190]	[0.00368]	[0.00176]	[0.00226]	[0.00178]	[0.00289]	[0.00211]	[0.00209]	[0.00166]	[0.00161]
Socio-economic background		0.998**	0.995***	0.997**	0.998***	1.001***	1.000	1.000	1.000	0.996***	0.999**
		[0.000541]	[0.000927]	[0.000500]	[0.000614]	[0.000491]	[0.000685]	[0.000569]	[0.000576]	[0.000471]	[0.000479]
Migration background = 1		0.378***	1.788	0.609***	0.282***	0.386***	0.949	1.081	1.530***	0.998***	3.691***
		[0.0792]	[0.647]	[0.115]	[0.0643]	[0.0732]	[0.223]	[0.216]	[0.320]	[0.0912]	[0.637]
Estimation with lifetime returns											
Field-specific skills	7.106***										
	[0.412]										
Field-specific returns to education	0.00204***										
	[0.000754]										
Field-specific variability of expected income	0.212***										
	[0.0155]										
Field feminization	1.010***										
	[0.00262]										
Math-inclined = 1		1.774***	2.773***	1.681***	1.567***	3.002***	3.135***	3.114***	2.375***	3.250***	2.883***
		[0.223]	[0.596]	[0.197]	[0.236]	[0.344]	[0.504]	[0.419]	[0.325]	[0.358]	[0.317]
Grade		1.016***	0.994	1.023***	1.015***	1.034***	1.090***	1.044***	1.049***	1.011***	1.026***
		[0.00189]	[0.00386]	[0.00175]	[0.00226]	[0.00179]	[0.00318]	[0.00223]	[0.00213]	[0.00167]	[0.00161]
Socio-economic background		0.998***	0.995***	0.997***	0.998***	0.999***	1.002***	1.000	1.000	0.996***	0.999**
		[0.000479]	[0.000537]	[0.000979]	[0.000499]	[0.000616]	[0.000490]	[0.000734]	[0.000577]	[0.000581]	[0.000479]
Migration background = 1		0.396***	1.421	1.139	0.249***	0.439***	1.632**	2.409***	2.963***	0.531***	3.696***
		[0.0829]	[0.534]	[0.218]	[0.0579]	[0.0837]	[0.398]	[0.498]	[0.646]	[0.0995]	[0.638]

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Table reports odds ratios. n=14,988. Constants not reported.

Further controls (odds-ratios not reported) include if respondents indicated their parents were an important source of information and advice; if their choice set was constrained by too low grades; and if they had already completed a vocational training.

Table A2: Nested logistic regression - boys only

VARIABLES	All fields	University of applied sciences (FH)				University				Vocational training	doesn't know
		Humanities	Life sciences	Social sciences	Sciences	Humanities	Life sciences	Social sciences	Sciences		
Estimation with net present value											
Field-specific skills	9.682***										
	[0.570]										
Field-specific NPV of expected incomes	1.000***										
	[5.74e-06]										
Field-specific variability of expected income	0.332***										
	[0.0183]										
Field feminization	0.988***										
	[0.00234]										
Math-inclined = 1		1.557***	1.156	1.367***	1.499***	2.039***	3.148***	2.067***	1.921***	2.195***	2.644***
		[0.189]	[0.242]	[0.145]	[0.145]	[0.205]	[0.498]	[0.251]	[0.210]	[0.202]	[0.236]
Grade		1.005***	0.989**	1.017***	1.011***	1.019***	1.066***	1.040***	1.037***	1.000	1.020***
		[0.00212]	[0.00354]	[0.00164]	[0.00148]	[0.00153]	[0.00309]	[0.00210]	[0.00165]	[0.00145]	[0.00132]
Socio-economic background		0.998***	1.001	0.999	1.000	1.000	1.004***	1.002***	1.001***	0.999***	1.001***
		[0.000527]	[0.000881]	[0.000444]	[0.000391]	[0.000432]	[0.000722]	[0.000525]	[0.000448]	[0.000388]	[0.000385]
Migration background = 1		0.649**	0.649**	1.538	0.708*	0.431***	415***	1.087	0.690*	1.115	3.629***
		[0.136]	[0.622]	[0.127]	[0.0753]	[0.0783]	[0.265]	[0.144]	[0.203]	[0.188]	[0.561]
Estimation with lifetime returns											
Field-specific skills	9.953***										
	[0.586]										
Field-specific returns to education	0.00680***										
	[0.00250]										
Field-specific variability of expected income	0.344***										
	[0.0200]										
Field feminization	0.982***										
	[0.00294]										
Math-inclined = 1		1.735***	1.283	1.277**	1.527***	2.018***	3.226***	2.167***	1.991***	2.377***	2.650***
		[0.215]	[0.272]	[0.135]	[0.148]	[0.203]	[0.521]	[0.269]	[0.219]	[0.218]	[0.236]
Grade		1.008***	0.991**	1.014***	1.011***	1.019***	1.068***	1.042***	1.039***	1.002	1.020***
		[0.00215]	[0.00357]	[0.00161]	[0.00148]	[0.00154]	[0.00345]	[0.00235]	[0.00173]	[0.00145]	[0.00132]
Socio-economic background		0.998***	1.001	0.999*	0.999	1.000	1.004***	1.002***	1.002***	0.999***	1.001***
		[0.000541]	[0.000905]	[0.000439]	[0.000391]	[0.000431]	[0.000738]	[0.000530]	[0.000451]	[0.000387]	[0.000383]
Migration background = 1		0.943	1.402	1.999***	0.638***	0.631**	2.254***	1.825***	2.477***	1.117	3.638***
		[0.197]	[0.571]	[0.356]	[0.109]	[0.117]	[0.549]	[0.376]	[0.464]	[0.188]	[0.562]

*** p<0.01, ** p<0.05, * p<0.1. Table reports odds ratios. n=12,646. Constants not reported.
Further controls (odds-ratios not reported) include if respondents indicated their parents were an important source of information and advice; if their choice set was constrained by too low grades; and if they had already completed a vocational training.

Table A3: Nested logistic regression on whole sample, odds ratios

VARIABLES	All fields	University of applied sciences (FH)				University				Vocational training	doesn't know
		Humanities	Life sciences	Social sciences	Sciences	Humanities	Life sciences	Social sciences	Sciences		
Field-specific skills (self-assessed)	8.825***										
	[0.362]										
Field-specific returns to education	0.0138***										
	[0.00350]										
Field-specific variability of expected income	0.352***										
	[0.0140]										
Field feminization	0.982***										
	[0.00228]										
Math-inclined = 1		1.613***	1.606***	1.331***	1.436***	1.983***	2.420***	2.155***	1.774***	2.386***	2.411***
		[0.138]	[0.224]	[0.103]	[0.115]	[0.148]	[0.246]	[0.194]	[0.152]	[0.169]	[0.169]
Grade		1.010***	0.992***	1.016***	1.011***	1.020***	1.063***	1.036***	1.036***	1.002*	1.019***
		[0.00136]	[0.00251]	[0.00118]	[0.00123]	[0.00115]	[0.00220]	[0.00162]	[0.00132]	[0.00109]	[0.00105]
Woman		1.621***	1.622***	1.093	0.322***	3.443***	1.512***	0.640***	0.731***	3.191***	2.162***
		[0.143]	[0.234]	[0.0906]	[0.0337]	[0.277]	[0.181]	[0.0674]	[0.0689]	[0.242]	[0.157]
Migration background = 1		0.705**	1.193	1.456***	0.567***	0.624***	1.854***	2.287***	2.372***	0.830	3.318***
		[0.101]	[0.315]	[0.189]	[0.0755]	[0.0807]	[0.291]	[0.326]	[0.335]	[0.102]	[0.384]
Socio-economic background		0.999***	0.999**	0.999***	0.999***	1.000	1.003***	1.001***	1.001***	0.998***	1.000
		[0.000366]	[0.000611]	[0.000326]	[0.000328]	[0.000317]	[0.000443]	[0.000384]	[0.000356]	[0.000302]	[0.000303]

*** p<0.01, ** p<0.05, * p<0.1. n = 27,637. Constants not reported.
Further controls (odds-ratios not reported) include: if respondents indicated their parents were an important source of information and advice; if their choice set was constrained by too low grades; and if they had already completed a vocational training.

Figure 3: Predicted probabilities by field characteristics - estimation using net present value

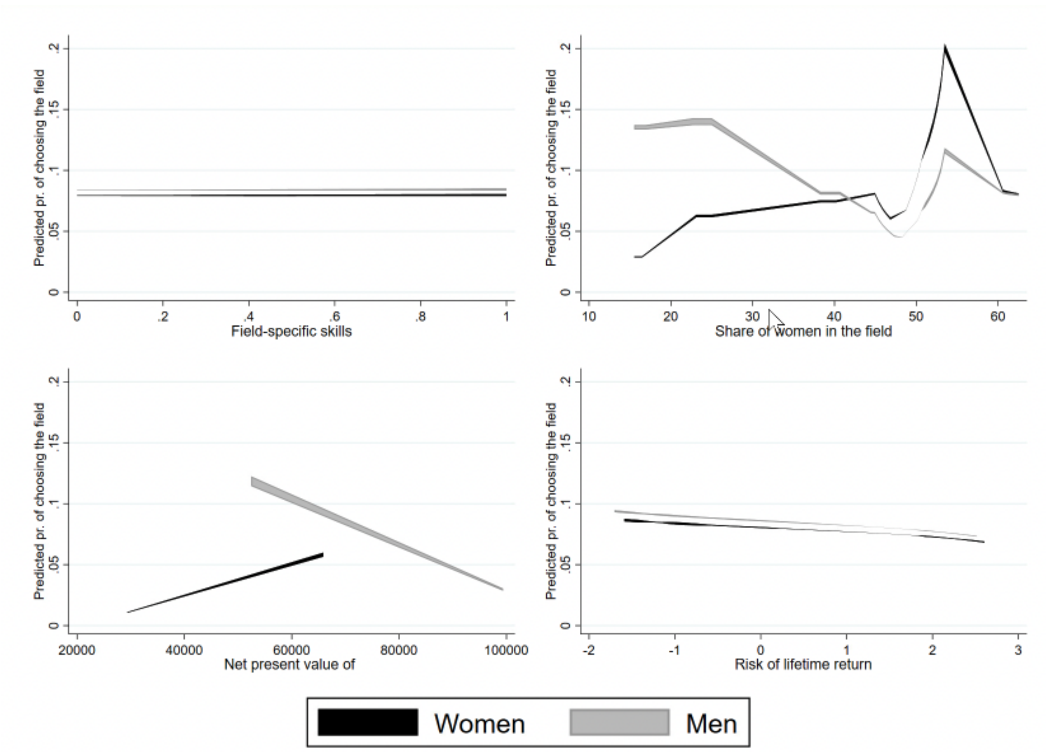


Table A4: Nested logistic regression on whole sample - using only the gender dummy

		University of applied sciences (FH)				University					
VARIABLES	All fields	Humanities	Life sciences	Social sciences	Sciences	Humanities	Life sciences	Social sciences	Sciences	Vocational training	doesn’t know
Woman		7.297***	2.513***	2.041***	4.840***	0.973	10.98***	3.583***	4.553***	2.870***	12.29***
		[0.193]	[0.172]	[0.294]	[0.0778]	[0.623]	[0.226]	[0.298]	[0.198]	[0.676]	[0.412]

*** p<0.01, ** p<0.05, * p<0.1. Table reports odds ratios. n=27,881. Constants not reported.