

# Fields of Study Choices and the Reproduction of Gender Inequalities

Carlo D'Ippoliti\* and Svenja Flechtner<sup>†</sup>

January 4, 2019

*Poster presentation at the AEA Meeting 2019 - Background paper. Incomplete draft; please do not quote. Comments welcome!*

## Abstract

We study fields of study choices of students in Germany who hold a university entrance permit. Fields of study choices of males and females are associated with different expected income earnings, making this choice a relevant factor for explaining gender income segregation. We use data from the German school-leavers panel 2008, featuring data of 28,000 students in their last year of upper secondary education. We use nested logistic regressions to study the choices students make out of up to 10 groups of options. We consider a broad range of explanatory factors, including individual skills and grades, expected income, preferences, gender, diverse household characteristics, and several behavioural variables. In contrast to the vast majority of the literature, we do not only consider a gender differences but interact gender with other background factors. We find that diversity matters indeed: choice patterns of women are different when controlling for migration background and household characteristics.

---

\*Sapienza University of Rome. Email: [carlo.dippoliti@uniroma1.it](mailto:carlo.dippoliti@uniroma1.it).

<sup>†</sup>University of Siegen. FMM Fellow and ZOE Fellow. Email: [svenja.flechtner@uni-siegen.de](mailto:svenja.flechtner@uni-siegen.de).

# 1 Motivation

A secondary school graduate's choice about if and how to continue their education has important implications for their labour market prospects. There is considerable vertical and horizontal labour market segregation due to students' choices of tertiary education and field of study. Future increases of female labour participation may come along with further segregation: there is evidence of a trade-off between sectoral gender segregation on the one hand and the raise of women's employment (Bettio and Verashchagina, 2009). Understanding the reasons of this labour market segregation is therefore of primary importance for researchers and policy-makers alike.

The importance of fields of study choices have already inspired growing bodies of literature in the social sciences, most of which have been concerned with European and US students. Thanks to a diverse range of empirical studies, using diverse datasets and methodologies, important insights into the determinants of fields of study choices of secondary school-leavers have been gained.

The empirical literature has tested a number of hypothesized explanations stemming from different theoretical backgrounds, and different hypotheses have found empirical support. First, there is evidence that students' abilities direct them into fields where they can rely on their comparative advantage Ayalon, 2003; Bettio and Verashchagina, 2009. At the same time, however, students' self-assessment of their personal strengths and weaknesses may be biased: Goetz et al. (2013) conducted an experimental study with upper secondary school students (*Gymnasium*) in Germany and found that female students reported lower perceived competence than male students, despite having the same average grades in math.

Second, students seem to follow their preferences for different fields. Third, it has been discussed to what extent students conform to social expectations, for example when girls do not go into professions perceived as male. Fourth, several papers have studied if anticipated prejudice prevents certain groups, such as women, from going into male-dominated fields.

The literature has focused on specific groups of students separately, for example women or migrants, without providing direct comparisons. At the same time, claims about factors influencing fields of study choices are usually made in a generalizing way, without taking into account enough that different mechanisms may matter for different people. In this paper,

we analyze whether different determinants matter in different ways for explaining labour market segregation of specific groups of students. In particular, we study to what extent students of, say, women are shaped by similar factors across all women, or if choices differ between women from different socio-economic or ethnic backgrounds.

We use data from the German school-leavers panel 2008, featuring data of 28,000 students in their last year of upper secondary education. We use nested logistic regressions to study the choices students make out of up to 10 groups of options. We consider a broad range of explanatory factors, including individual skills and grades, expected income, preferences, gender, diverse household characteristics, and several behavioural variables. We find that expected income earnings are only significantly positive for boys but not for girls. There is a possible self-confirmatory bias: self-perceived field-specific skills are always important and labor market optimism is correlated with studying more. Field feminization exerts opposed effects: field feminization is negative for boys but positive for girls. Migrants behave like students from richer families in that they prefer university studies to vocational training and universities of applied sciences. These differences could arise both from factors on the demand side of labour, such as direct or indirect discrimination, and from the supply side, through people's own choices.

The paper advances the literature through addressing socio-economic backgrounds, migration backgrounds and gender not only separately, but also jointly. Other studies have typically focused on one of these factors alone, usually without indicating explicit reasons to do so. Ayalon (2003) is an exception when she argues that sources of gender-based inequality are fundamentally different from those based on socio-economic background and ethnic origin, because unequal educational achievements primarily stem from lower participation rates in postsecondary schooling, whereas girls obtain unequal labour market outcomes primarily as a consequence of their fields of study choice. While we do not debate these differences (see also descriptive statistics in section ??), we consider these different factors together with potential interactions between them in mind. For example, societal stereotypes that affect women's career choices may be intertwined with discrimination facing women from migrant backgrounds. To consider these channels jointly, we analyse determinants of one vertical moment of inequality - to proceed to university after secondary school or not - and one horizontal moment of inequality - fields of study choices.

Our results vindicate that intersectional perspectives of diverse channels of discrimination and disadvantage are important to understand fields of study choices in their specific social contexts.

## **2 Background: Horizontal and vertical labour market segregation and income inequalities in Germany**

Fields of study choices and expected incomes related with them are segregated to considerable degrees, both by gender and by socio-economic and migration background.

According to data from the German Federal Statistical Office, the share of female first semester students has slowly risen over the past years, but there is still no gender parity with 48.2% in winter term 2016 (Destatis, 2018). Table 1 shows the gender distribution in different fields of students having started their degrees in 2012 and 2016, alongside average earnings of job starters in these fields from 2018. It is clear that females are overrepresented in fields with lower expected earnings, with the exception of medicine.

Besides this horizontal gender segregation, there is additional vertical segregation: while there is no gender parity among undergraduates, the share of girls finishing upper secondary school has exceeded that of boys. In 2015, 54.3% of all school-leavers with a university entrance certificate were female. In all other (lower) types of school and certificate - including those with permit to enter applied science universities, and the group without any certificate - males are overrepresented (Destatis, 2017).

Vertical segregation according to socio-economic background begins much earlier than at the entry to university: many choices are pre-determined with transitions from primary to lower secondary and from lower to upper secondary. Upper secondary schooling (school years 11-12 or 11-13, depending on track) leads to a university entrance permit. Students who leave school after 10 years - to work, or to do vocational training - cannot access university without further additional preparatory trainings. In turn, some lower secondary tracks are supposed to lead into vocational training rather than into upper secondary schooling. The transition from primary school into different tracks thus has important implications. At the same time, this

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

Field of study	Share of females, 2016	Income of job starters, 2018, in Euro
Humanities	0.67	34,700
Agriculture, environmental science and nutrition	0.58	40,900
Law, economics and social sciences	0.57	45,100
Mathematics and sciences	0.47	47,600
Medicine and life sciences	0.66	48,850
Engineering	0.23	51,350

Source: own table, based on Destatis (2018), Fachserie 11, Reihe 4.1 for student numbers and Statista (2018) for income data. Income are average gross earnings. Categories do not overlap exactly.

early transition is segregated according to socio-economic background and migration history to considerable degrees. In 2006, 96% of all children of parents with upper secondary schooling obtained at least this same level of qualification (Bildungsberichterstattung, 2009).

These early disadvantages are mirrored in the composition of students enrolled in upper secondary school: in 2016, the majority of students enrolled in a school leading to university entrance permit (Gymnasium) had parents with a university degree. The share of students whose parents had no degree or vocational training was below 10% (Destatis, 2017). Statistically, even when controlling for socio-economic background, a family history of migration is an additional source of disadvantage (Bildungsberichterstattung, 2009). Correspondingly, the educational enrolment rate of young people between 20 and 25 years (i.e. the age of tertiary education) is lower among students with migration background, in particular those with Turkish family backgrounds (ibid.).

## 3 Methodology and data

### 3.1 Data

We use panel data from a representative survey of school-leavers (with a university entrance permit) provided by the German Centre for Higher Education Research and Science Studies (DZHW). The survey (*Studienberechtigtenpanel*) follows school-leavers with a university entrance permit. In this paper, we use the first wave, in which final-year students in upper secondary schools are asked about their educational aspirations and plans for the future.

Our dataset contains responses from 28.281 individuals, of which 53.8% are female. Given that the share of females among all students who obtained a university entrance permit in 2015 was 54% (Destatis, 2017), this corresponds well. 8% of our respondents were born outside Germany. About 12 % speak German as well as a second language at home, 1.4% speak no German in their families. On this basis, we consider among 8-13% of our sample to have a migration background. According to microcensus data from 2011, 22% of all students attending Gymnasium in 2011 had a migration background widely defined (9.8% when second or third-generation students with German citizenship are excluded).

Students enrolled in the last year of upper-secondary school are asked what, if any, tertiary studies they are seeking for the next year (then the subsequent waves of the panel followed these students in the following years). 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 (Universität, henceforth “uni”) or a university of applied sciences (Fachhochschule, henceforth “FH”) program: humanities, social sciences, natural sciences, or life sciences. In Germany, successful completion of some kinds of secondary school only grants access to selected tertiary level study careers (e.g. from vocational training to FH in the same field of study). For each student, we only consider those options that are actually open to him or her.

Table 2 reports the number of students in our sample, by potential and actual choice. The answers of 24,249 respondents are considered because almost 14% of the sample did not answer to some crucial questions in the survey. Among our respondents, 70.6% seek to achieve a general university

Table 2: Share of students by potential and actual choices

Nested choice	Field of study	Observations (potential choices)	Times actu- ally chosen	% of po- tential
No education/ don't know	-	24,249	4792	19.8%
Vocational train- ing	Vocational	24,249	6545	27%
University	Humanities	18,620	2806	15.1%
	Life Sciences	18,610	1321	7.1%
	Social Sciences	18,668	2713	14.5%
	Natural Sciences	18,645	2532	13.6%
FH	Humanities	24,144	475	2%
	Life Sciences	24,134	215	0.9%
	Social Sciences	24,192	1040	4.3%
	Natural Sciences	24,169	1810	7.5%
Total		219,680	24,249	100%

*Note: the choices that could potentially be made depend on the type of secondary school in which students are enrolled.*

entrance permit, and 27.7% seek a general entrance permit to universities of applied sciences. The remaining 1.7% will obtain a limited entrance permit.

### 3.2 Method

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) typical of the simple expected utility model, which experimental studies generally regard as unrealistic.<sup>1</sup> The nested logistic does not need IIA because rather than assuming independently distributed errors (such as conditional or multinomial logistic models do) it allows for clustering similar alternatives into nests.

<sup>1</sup>For a demonstration of how this model can be derived under the assumption of utility maximization, see Amemiya, 1985.

As shown in table 2, 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, together with being unsure (“don’t know” voluntary answer) are our reference category. We consider the following potential explanatory factors:

- classical utility maximization: field-specific skills and preferences, school grades, and field-specific expected returns;
- individuals’ diversity: gender, migrant background, and family income;
- behavioral aspects: gender roles (feminization of the various fields), and individuals’ optimism.

For a number of secondary school disciplines (German, English, mathematics, and vocational-related activities) students were asked how much they found the respective courses stimulating, multifaceted, understandable, and comprehensible. We aggregated the answers to obtain a measure of students’ preference for these disciplines. In the subsequent analysis, we abstain from positing unambiguous links between these three disciplines and the various fields of study, and we estimate the impact of preference for each discipline on the propensity to choose each field. 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.

Field-specific expected returns are considered in terms of expected income. The relevant data was obtained from (Glocker and Storck, 2014). They provide income data for different university career choices, differentiating men and women as well as general universities and universities of applied sciences. Estimates are based on data from the German Micro Census.

Students reported their migrant background both by reporting their nationality and their place of birth; in the following analyses we consider the former, but results do not significantly change when using the latter.

Family income was not directly asked in the survey, so it was obtained by factor analysis on the matrix reporting the tetrachoric correlations among the answers to the following questions: 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.

School grades were made comparable across school types by converting them into centile of the empirical grade distribution for each school.

Finally, the feminization of the various fields is computed by considering the share of women among full professors for each field of study; and students' optimism was computed with respect to their labour market prospects. Specifically, students were asked how easy, on a scale from 1 to 5, they think it is in general to find employment after finishing studying, and separately how easy it will be for them personally. Our measure of optimism is the ratio of the two answers.

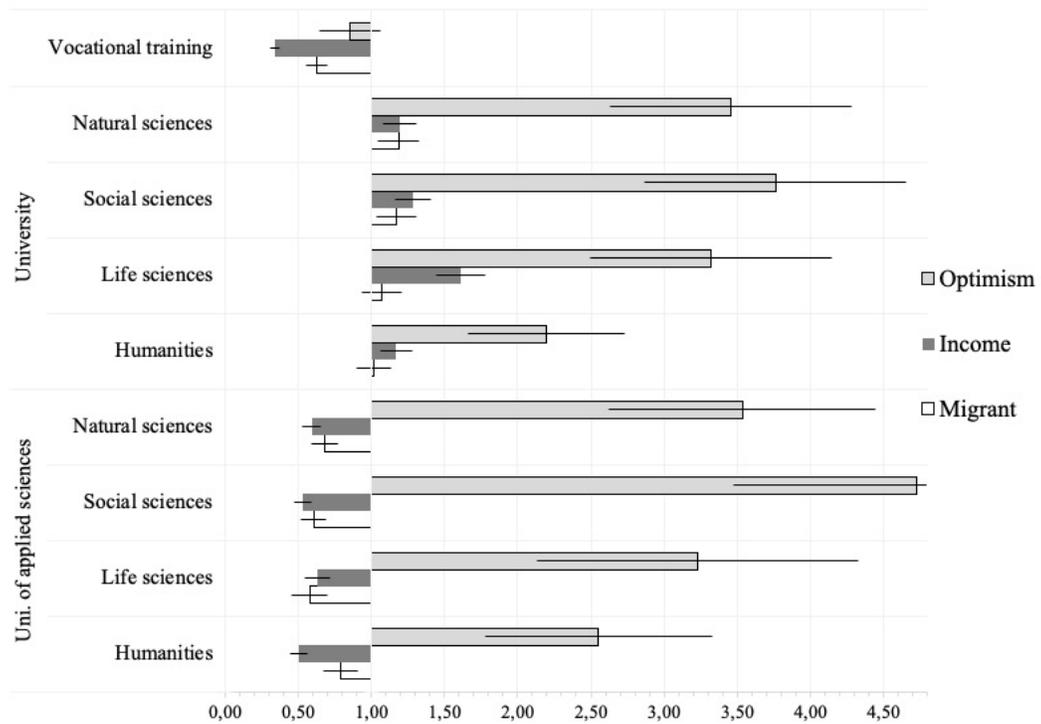
## 4 Results and discussion

We run the nested logistic model on the whole sample and for boys and girls separately. The main results are reported in tables 3 to 8. Since field feminization and the expected income associated to each study field are highly correlated, they could not be included together in the estimates due to multicollinearity.

In all estimations, the standard errors within the nest grouping fields of university studies, and within that grouping fields of university of applied sciences (FH) studies, are positively correlated and the correlation coefficient is significantly different from zero. Accordingly, the IIA hypothesis is rejected and the baseline model of expected utility maximization does not apply.

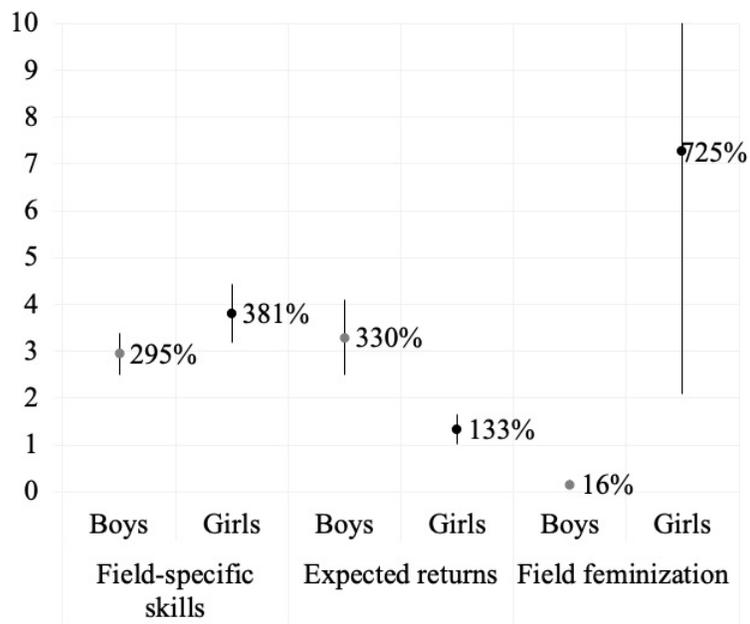
As shown in figure 1, migrants seem to behave like the students from higher socio-background families, preferring university studies to vocational training and universities of applied sciences. A possible interpretation of this finding is in terms of signaling, in so far as students of migrant background who enroll in tertiary education could wish to clearly distinguish themselves from other non-German residents in the country, by pursuing studies such as the natural or social sciences that attach social prestige.

Figure 1: Relevance of the individual-level variables in the choice process, odds ratios



Note: Full results are shown in tables 3 and 4.

Figure 2: Relevance of the field-specific variables in the choice process, odds ratios



Note: Full results are shown in tables 5-8.

Labour market optimism induces students to continue studying, either in a university or FH, rather than entering in vocational training. Furthermore, as shown in figure 2 field-specific skills are positively associated with choosing a certain field both for boys and girls, as predicted by extant literature. Since these skills are self-perceived, this finding and that on optimism may denote a certain confirmatory bias on the side of students.

Finally, we find significant gender differences in the impact of field characteristics on students' choice. On the one hand, the coefficient of expected earnings is only significantly positive for boys and not for girls, that is, only boys seem to choose a certain field of study because of the expected returns associated to it. On the other hand, field feminization even exerts opposed effects: a negative one for boys, and a positive one for girls. This last finding may be compatible both a pollution theory of discrimination, whereby men do not wish to be in a profession where many women are employed, as well as with the hypothesis that women may be encouraged to follow role models of the same gender and/or wish to work in an environment perceived to be women-friendly.

Figure 3: Nested logistic regression on whole sample, odds ratios (includes feminization)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(13)	(14)
	Field-spec vars	FH humanities	FH life sciences	FH social sciences	FH sciences	uni humanities	uni life sciences	uni social sciences	uni sciences	vocational	$\beta$	uni fan
Woman		0.954 [0.0663]	1.045 [0.0859]	0.857*** [0.0504]	0.580*** [0.0363]	1.367*** [0.0675]	1.510*** [0.0825]	1.137*** [0.0545]	1.013 [0.0527]	1.807*** [0.0865]		
Migrant		0.790 [0.116]	0.577** [0.126]	0.606*** [0.0869]	0.679*** [0.0917]	1.020 [0.119]	1.071 [0.133]	1.170 [0.134]	1.185 [0.137]	0.625*** [0.0717]		
Family income		0.504*** [0.0605]	0.628*** [0.0867]	0.528*** [0.0581]	0.591*** [0.0638]	1.169* [0.110]	1.611*** [0.165]	1.284*** [0.121]	1.197* [0.114]	0.341*** [0.0320]		
School grades		1.022 [0.270]	0.408* [0.191]	1.009 [0.244]	0.837 [0.205]	6.029*** [1.385]	8.519*** [1.940]	6.455*** [1.468]	8.174*** [1.847]	8.70e-05*** [5.34e-05]		
Preferences: math		0.895 [0.0802]	0.808*** [0.0827]	1.165* [0.0924]	1.267*** [0.102]	0.980 [0.0719]	0.828** [0.0720]	1.184** [0.0887]	1.302*** [0.103]	1.299*** [0.0797]		
Preferences: German		0.756*** [0.0658]	0.670*** [0.0676]	0.732*** [0.0671]	0.748*** [0.0711]	1.186** [0.0894]	0.921 [0.0793]	1.014 [0.0804]	0.989 [0.0840]	1.417*** [0.0949]		
Individual optimism		2.553*** [0.778]	3.231*** [1.098]	4.729*** [1.250]	3.535*** [0.914]	2.198*** [0.537]	3.317*** [0.825]	3.761*** [0.893]	3.454*** [0.824]	0.851 [0.210]		
Preferences: English		0.783*** [0.0603]	0.634*** [0.0602]	0.682*** [0.0556]	0.809*** [0.0671]	0.983 [0.0660]	0.777*** [0.0616]	0.884* [0.0626]	0.878* [0.0667]	1.204*** [0.0718]		
Field-specific skills		3.311*** [0.398]										
Field feminization		0.636 [0.282]										
Constant											1.338*** [0.0461]	1.333*** [0.0423]
Observations	162,090	162,090	162,090	162,090	162,090	162,090	162,090	162,090	162,090	162,090	162,090	162,090
Log-Likelihood	-31201	-31201	-31201	-31201	-31201	-31201	-31201	-31201	-31201	-31201	-31201	-31201

Note: Nested tree structure is reported in table 2. Results including expected income and field feminization are reported separately due to multicollinearity.

Figure 4: Nested logistic regression on whole sample, odds ratios (includes expected income)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(13)	(14)
	Field-spec vars	FH humanities	FH life sciences	FH social sciences	FH sciences	uni humanities	uni life sciences	uni social sciences	uni sciences	vocational	fh tau	uni tau
Women		0.942 [0.0637]	0.952 [0.0758]	0.956 [0.0598]	0.679*** [0.0440]	1.361*** [0.0666]	1.431*** [0.0768]	0.963 [0.0568]	1.171*** [0.0650]	1.785*** [0.0852]		
Migrant		0.787 [0.115]	0.580** [0.125]	0.607*** [0.0867]	0.678*** [0.0915]	1.022 [0.119]	1.072 [0.133]	1.170 [0.136]	1.185 [0.133]	0.624*** [0.0717]		
Family income		0.494*** [0.0591]	0.609*** [0.0832]	0.511*** [0.0562]	0.565*** [0.0607]	1.147 [0.108]	1.555*** [0.156]	1.235** [0.116]	1.152 [0.109]	0.329*** [0.0309]		
School grades		1.018 [0.267]	0.416* [0.190]	0.997 [0.240]	0.827 [0.202]	6.037*** [1.382]	8.509*** [1.932]	6.468*** [1.466]	8.168*** [1.841]	8.26e-05*** [5.07e-05]		
Preferences: math		0.869 [0.0753]	0.765*** [0.0771]	1.018 [0.0749]	1.061 [0.0781]	0.921 [0.0605]	0.726*** [0.0600]	1.008 [0.0686]	1.112 [0.0731]	1.119* [0.0716]		
Preferences: German		0.702*** [0.0571]	0.614*** [0.0603]	0.629*** [0.0535]	0.607*** [0.0548]	1.083 [0.0685]	0.800*** [0.0634]	0.860** [0.0647]	0.834** [0.0656]	1.214*** [0.0832]		
Individual optimism		2.504*** [0.755]	3.089*** [1.036]	4.433*** [1.163]	3.287*** [0.848]	2.126*** [0.517]	3.127*** [0.772]	3.509*** [0.826]	3.227*** [0.764]	0.793 [0.196]		
Preferences: English		0.736*** [0.0533]	0.596*** [0.0556]	0.602*** [0.0471]	0.679*** [0.0529]	0.914 [0.0525]	0.689*** [0.0511]	0.768*** [0.0517]	0.757*** [0.0533]	1.056 [0.0642]		
Field-specific skills		3.245*** [0.362]										
Expected income		2.430*** [0.411]										
Constant											1.330*** [0.0416]	1.348*** [0.0393]
Observations	162,090	162,090	162,090	162,090	162,090	162,090	162,090	162,090	162,090	162,090	162,090	162,090
Log-Likelihood	-31187	-31187	-31187	-31187	-31187	-31187	-31187	-31187	-31187	-31187	-31187	-31187

Note: Nested tree structure is reported in table 2. Results including expected income and field feminization are reported separately due to multicollinearity.

Figure 5: Nested logistic regression on boys sample, odds ratios (includes feminization)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(13)	(14)
	Field-spec vars	FH humanities	FH life sciences	FH social sciences	FH sciences	uni humanities	uni life sciences	uni social sciences	uni sciences	vocational	fh tau	uni tau
Migrant		0.761 [0.158]	0.665 [0.203]	0.726* [0.135]	0.714* [0.126]	1.083 [0.182]	1.143 [0.200]	1.158 [0.189]	1.156 [0.188]	0.674** [0.114]		
Family income		0.438*** [0.0812]	0.730 [0.152]	0.595*** [0.0910]	0.595*** [0.0866]	1.218 [0.172]	1.695*** [0.255]	1.374** [0.189]	1.304** [0.179]	0.419*** [0.0597]		
School grades		0.730 [0.378]	0.754 [0.383]	0.941 [0.354]	0.650 [0.240]	7.287*** [2.499]	8.990*** [3.072]	7.057*** [2.406]	8.062*** [2.730]	1.51e-05*** [1.81e-05]		
Preferences: math		0.902 [0.123]	0.813 [0.129]	1.111 [0.116]	1.162 [0.118]	0.951 [0.0944]	0.756** [0.0942]	1.080 [0.104]	1.100 [0.109]	1.218** [0.104]		
Preferences: German		0.702*** [0.0876]	0.533*** [0.0839]	0.567*** [0.0737]	0.578*** [0.0754]	0.999 [0.107]	0.786* [0.0979]	0.849 [0.0970]	0.816* [0.0988]	1.293*** [0.127]		
Individual optimism		2.228* [0.925]	3.896*** [1.692]	3.666*** [1.253]	2.788*** [0.916]	2.823*** [0.917]	3.673*** [1.199]	4.213*** [1.318]	3.761*** [1.173]	0.406** [0.144]		
Preferences: English		0.719*** [0.0792]	0.574*** [0.0817]	0.563*** [0.0663]	0.676*** [0.0750]	0.792** [0.0757]	0.634*** [0.0719]	0.700*** [0.0711]	0.710*** [0.0749]	1.131 [0.0985]		
Field-specific skills		2.777*** [0.467]										
Field feminization		0.158*** [0.0899]										
Constant											1.283*** [0.0600]	1.241*** [0.0450]
Observations	72 416	72 416	72 416	72 416	72 416	72 416	72 416	72 416	72 416	72 416	72 416	72 416
Log-Likelihood	-13963	-13963	-13963	-13963	-13963	-13963	-13963	-13963	-13963	-13963	-13963	-13963

Note: Nested tree structure is reported in table 2. Results including expected income and field feminization are reported separately due to multicollinearity.

Figure 6: Nested logistic regression on boys sample, odds ratios (includes expected income)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(13)	(14)
	Field-spec vars	FH humanities	FH life_sciences	FH social_sciences	FH sciences	uni_humanities	uni_life_sciences	uni_social_sciences	uni_sciences	vocational	fh_tau	uni_tau
Migrant		0.766 [0.163]	0.660 [0.213]	0.727* [0.136]	0.713* [0.126]	1.080 [0.182]	1.142 [0.202]	1.158 [0.189]	1.156 [0.188]	0.674** [0.114]		
Family income		0.426*** [0.0807]	0.732 [0.158]	0.587*** [0.0904]	0.581*** [0.0846]	1.216 [0.172]	1.704*** [0.254]	1.349** [0.186]	1.284* [0.176]	0.406*** [0.0580]		
School grades		0.729 [0.389]	0.753 [0.397]	0.953 [0.360]	0.642 [0.237]	7.247*** [2.486]	9.054*** [3.095]	7.021*** [2.391]	8.077*** [2.733]	1.48e-05*** [1.78e-05]		
Preferences: math		0.905 [0.120]	0.771 [0.124]	1.054 [0.106]	1.067 [0.105]	0.975 [0.0875]	0.723*** [0.0868]	1.007 [0.0929]	1.044 [0.0937]	1.089 [0.0987]		
Preferences: German		0.699*** [0.0822]	0.508*** [0.0819]	0.534*** [0.0662]	0.516*** [0.0656]	1.011 [0.0941]	0.754** [0.0872]	0.781** [0.0866]	0.758** [0.0868]	1.138 [0.118]		
Individual optimism		2.162* [0.921]	3.894*** [1.743]	3.632*** [1.246]	2.678*** [0.881]	2.778*** [0.903]	3.612*** [1.184]	4.128*** [1.289]	3.677*** [1.145]	0.383*** [0.136]		
Preferences: English		0.721*** [0.0764]	0.551*** [0.0804]	0.534*** [0.0617]	0.619*** [0.0661]	0.805** [0.0681]	0.613*** [0.0655]	0.653*** [0.0648]	0.670*** [0.0664]	1.017 [0.0927]		
Field-specific skills	2.948*** [0.447]											
Expected income	3.297*** [0.792]											
Constant											1.309*** [0.0573]	1.257*** [0.0414]
Observations	72,416	72,416	72,416	72,416	72,416	72,416	72,416	72,416	72,416	72,416	72,416	72,416
Log-Likelihood	-13953	-13953	-13953	-13953	-13953	-13953	-13953	-13953	-13953	-13953	-13953	-13953

Note: Nested tree structure is reported in table 2. Results including expected income and field feminization are reported separately due to multicollinearity.

Figure 7: Nested logistic regression on girls sample, odds ratios (includes feminization)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(13)	(14)
	Field-spec vars	FH humanities	FH life_sciences	FH social_sciences	FH sciences	uni_humanities	uni_life_sciences	uni_social_sciences	uni_sciences	vocational	fh_tau	uni_tau
Migrant		0.778 [0.167]	0.506** [0.159]	0.487*** [0.110]	0.662* [0.145]	0.970 [0.158]	1.021 [0.179]	1.180 [0.189]	1.257 [0.210]	0.589*** [0.0923]		
Family income		0.542*** [0.0909]	0.596*** [0.114]	0.496*** [0.0786]	0.624*** [0.105]	1.152 [0.147]	1.630*** [0.231]	1.238* [0.160]	1.104 [0.151]	0.305*** [0.0383]		
School grades		1.209 [0.402]	0.199** [0.159]	1.118 [0.352]	1.153 [0.380]	5.114*** [1.586]	8.211*** [2.519]	5.966*** [1.851]	9.015*** [2.779]	0.000202*** [0.000143]		
Preferences: math		0.974 [0.118]	0.908 [0.122]	1.364** [0.178]	1.370** [0.189]	1.353*** [0.152]	1.292** [0.167]	1.615*** [0.197]	1.879*** [0.247]	1.818*** [0.170]		
Preferences: German		0.895 [0.108]	0.872 [0.117]	0.997 [0.132]	0.895 [0.129]	1.694*** [0.183]	1.371*** [0.168]	1.392*** [0.159]	1.354** [0.170]	2.006*** [0.186]		
Individual optimism		3.307*** [1.515]	3.236** [1.716]	6.979*** [2.896]	5.679*** [2.480]	2.029* [0.746]	3.512*** [1.338]	3.655*** [1.339]	3.499*** [1.340]	1.847* [0.656]		
Preferences: English		0.911 [0.101]	0.750** [0.0956]	0.865 [0.103]	0.864 [0.112]	1.420*** [0.139]	1.169 [0.133]	1.246** [0.129]	1.136 [0.133]	1.603*** [0.135]		
Field-specific skills		3.806*** [0.625]										
Field feminization		7.253*** [5.156]										
Constant											1.392*** [0.0673]	1.466*** [0.0725]
Observations	89,674	89,674	89,674	89,674	89,674	89,674	89,674	89,674	89,674	89,674	89,674	89,674
Log-Likelihood	-37177	-37177	-37177	-37177	-37177	-37177	-37177	-37177	-37177	-37177	-37177	-37177

Note: Nested tree structure is reported in table 2. Results including expected income and field feminization are reported separately due to multicollinearity.

Figure 8: Nested logistic regression on girls sample, odds ratios , odds ratios (includes feminization)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(13)	(14)
	Field-spec vars	FH_humanities	FH_life_sciences	FH_social_sciences	FH_sciences	uni_humanities	uni_life_sciences	uni_social_sciences	uni_sciences	vocational	fh_tau	uni_tau
Migrant		0.761 [0.161]	0.516** [0.154]	0.498*** [0.111]	0.661* [0.142]	0.979 [0.158]	1.027 [0.178]	1.177 [0.188]	1.249 [0.207]	0.589*** [0.0921]		
Family income		0.525*** [0.0865]	0.572*** [0.106]	0.477*** [0.0751]	0.586*** [0.0965]	1.110 [0.141]	1.514*** [0.209]	1.170 [0.150]	1.051 [0.142]	0.295*** [0.0370]		
School grades		1.167 [0.383]	0.228** [0.169]	1.071 [0.355]	1.091 [0.355]	5.134*** [1.582]	8.021*** [2.443]	5.932*** [1.809]	8.748*** [2.676]	0.000186*** [0.000132]		
Preferences: math		0.835 [0.0957]	0.775* [0.101]	1.009 [0.107]	0.984 [0.111]	1.055 [0.0977]	0.925 [0.105]	1.143 [0.127]	1.332*** [0.127]	1.480*** [0.136]		
Preferences: German		0.743*** [0.0814]	0.725** [0.0912]	0.742*** [0.0823]	0.646*** [0.0822]	1.317*** [0.111]	1.006 [0.108]	1.023 [0.114]	0.996 [0.105]	1.641*** [0.149]		
Individual optimism		3.208*** [1.436]	3.061** [1.556]	6.050*** [2.478]	4.966*** [2.130]	1.858* [0.676]	3.049*** [1.140]	3.153*** [1.138]	3.027*** [1.136]	1.688 [0.595]		
Preferences: English		0.770*** [0.0770]	0.656*** [0.0801]	0.676*** [0.0701]	0.654*** [0.0763]	1.142* [0.0890]	0.888 [0.0906]	0.950 [0.0953]	0.866 [0.0879]	1.351*** [0.111]		
Field-specific skills	3.488*** [0.570]										1.349*** [0.0606]	1.433*** [0.0698]
Expected income	1.328 [0.312]											
Constant												
Observations	89,674	89,674	89,674	89,674	89,674	89,674	89,674	89,674	89,674	89,674	89,674	89,674
Log-Likelihood	-17181	-17181	-17181	-17181	-17181	-17181	-17181	-17181	-17181	-17181	-17181	-17181

Note: Nested tree structure is reported in table 2. Results including expected income and field feminization are reported separately due to multicollinearity.

## 5 Conclusion

This paper studies fields of study and labour market segregation of secondary school-leavers in Germany, considering in particular the intersectional influences of gender, socio-economic and migration background. Our analysis finds that different channel seem indeed to matter differently for different people. This work in progress allows us to highlight the following results.

Expected earnings are only significantly positive for boys, not for girls. Field feminization exerts opposed effects: negative for boys, positive for girls. Migrants behave like students from richer families, preferring university studies to vocational training and universities of applied sciences. This finding is consistent with Kristen et al. (2008), who has emphasized that students from families with a Turkish background appreciate university studies more than vocational careers. Signalling may be at the origin of these preferences.

## References

- Amemiya, Takeshi (1985). *Advanced Econometrics*. Cambridge: Harvard University Press.
- Ayalon, Hanna (2003). “Women and men go to university: Mathematical background and gender differences in choice of field in higher education”. In: *Sex Roles* 48.5-6, pp. 277–290.
- 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.
- Bildungsberichterstattung, Autorengruppe (2009). *Bildung in Deutschland 2008. Ein indikatorengestützter Bericht mit einer Analyse zu Übergängen im Anschluss an den Sekundarbereich I*. Wiesbaden: Destatis.
- Destatis (2017). *Statistisches Jahrbuch 2017*. Wiesbaden: Destatis.
- (2018). *Statistisches Jahrbuch 2018*. Wiesbaden: Destatis.
- Glocker, Daniela and Johanna Storck (2014). “Risks and returns to educational fields—A financial asset approach to vocational and academic education”. In: *Economics of Education Review* 42, pp. 109–129.
- Goetz, Thomas et al. (2013). “Do girls really experience more anxiety in mathematics?” In: *Psychological science* 24.10, pp. 2079–2087.

Kristen, Cornelia et al. (2008). "Higher education entry of Turkish immigrant youth in Germany". In: *International Journal of Comparative Sociology* 49.2-3, pp. 127–151.