Earnings Expectations of "First-in Family" University Students and Their Role for Major Choice*

Katharina Adler,[†] Fabian Kosse,[‡] Markus Nagler,[§] Johannes Rincke[¶]

May 15, 2025

Abstract

How do students' earnings expectations differ by being the first in their family to attend university (FiF) and how do they affect field of study choice? We leverage unique survey and administrative data to document sizable gaps in expected earnings between FiF and non-FiF students. Our data can explain two-thirds of this gap, with the largest share attributable to field of study choice. We show that FiF students sort less into study fields based on their earnings expectations. Investigating potential explanations, we find that in high-earning fields, FiF students expect lower own ability and worse non-wage amenities than non-FiF students.

*We thank the editor, David Jaeger, as well as three helpful referees for comments and suggestions. Katja Kaufmann as well as participants at various seminars and conferences provided helpful comments and suggestions. Nagler and Rincke thank the German Science Foundation for generous funding (NA 1722/1-1 and RI 1959/4-1). Nagler gratefully acknowledges funding by the Joachim Herz Foundation through an Add-On Fellowship. Kosse expresses his thanks to the Jacobs Foundation for generous funding.

†University of Würzburg; katharina.adler@uni-wuerzburg.de

‡University of Würzburg, CESifo, IZA; fabian.kosse@uni-wuerzburg.de

§University of Erlangen-Nuremberg, CESifo, IZA, and LASER; markus.nagler@fau.de

¶University of Erlangen-Nuremberg and CESifo; johannes.rincke@fau.de

1 Introduction

Students who would be first-in-their-family (FiF) to attend university face a series of disadvantages that may translate into reduced social mobility (Boneva and Rauh, 2018; Adamecz-Völgyi et al., 2020; Blanden et al., 2023). FiF students choose different fields of study and are more likely to drop out of university (Henderson et al., 2020; Edwards et al., 2022; Adamecz-Völgyi et al., 2023). That being said, conditional on having attended university, FiF students have different characteristics than non-FiF students: In particular, FiF students seem to enter university with lower cognitive (but similar or even higher non-cognitive) skills (Edwards et al., 2022; Adamecz et al., 2024). In the labor market, female FiF students have lower earnings on average, for example (Adamecz-Völgyi et al., 2023). The finding that FiF students face challenges in the labor market or that they are more likely to study less prestigious subjects (and at less prestigious universities) is even true in high-ability samples (Shure and Zierow, 2023; Stansbury and Rodriguez, 2024).

Expectations and beliefs about own ability, earnings, and non-wage amenities are prime candidates to explain some of the issues that disadvantaged students face. For example, in the context of gender gaps, the decisive role of expectations and beliefs is well documented. The idea is that women's lower wage expectations may translate into lower reservation wages, which in turn affect actual earnings (e.g., Kiessling et al., 2024). In standard human capital models, earnings expectations and expectations about related non-wage amenities should also affect human capital investments such as the choice of study field. However, there is no evidence on whether FiF students at university have different labor market expectations than non-FiF students. There is also no evidence to date on whether FiF students react differently to such expectations when deciding on their human capital investments.

In this paper, we use unique survey data on over 3,500 first-year undergraduate students from a large German university to inform on FiF differences in labor market expectations and to shed light on the causes and consequences of these. To do so, we elicit earnings expectations at different points in subjects' future lives and across various (counterfactual) fields of study as well as information on motives for study program choice, perceived ability ranks, and expectations about non-wage amenities, again across various fields of study. We classify an individual as an FiF student if they report in the survey that none of their parents have a university degree. We acknowledge that this classification may also capture differences between students with and without parents with a university degree that go beyond parents' educational attainment. We then link these survey data to administrative data on student characteristics and their actual field of study. The survey took place in students' second week at university (i.e., at a very early stage of students' university life).

We first study FiF differences in expected earnings. We document a sizable FiF gap in expected earnings both at age 30 and 40: FiF students expect to earn 5 percent less than their non-FiF classmates when they are 30 and 6 percent less when they are 40, respectively. This amounts to around half of the gender gap in earnings expectations in our data and is comparable to, albeit lower than, the actual earnings gap by FiF status we find in a complementary analysis of the German Socioeconomic Panel (SOEP). A heterogeneity analysis by ability (measured via high school GPA) and by gender suggests that the gap in our survey is mainly driven by low-ability and by male FiF students. Decomposing the FiF gap via Oaxaca–Blinder shows that our data can explain close to two thirds of the observed gap, which is more than twice the share that can be explained when assessing gender differences. Next to student characteristics, the largest share of the FiF gap in expected earnings is attributable to students' fields of study.

In a second step, we therefore investigate whether differences in earnings expectations matter for field of study choice. To study the relation between expected earnings and selection, we elicited expected earnings at the age of 30 conditional on obtaining a degree in eight counterfactual potential study fields. This allows us to study whether expected earnings predict the actual choice of study field *within* students. This approach is similar in spirit to Arcidiacono et al. (2020), who estimate the role of ex-ante treatment effects on expected earnings for major choice and occupational choice. Although in our setting the sorting into fields of study has already taken place, we think it is a useful exercise to gauge the consequences of differences in earnings expectations for field of study choice because the survey took place at the very beginning of the students' time at university and hence was arguably unaffected by their field of study environment. We find that within-person differences in earnings expectations across fields are less predictive for the field of study choice of FiF students than of non-FiF students. FiF students' lower self-perceived ability and their expectations about non-wage amenities in high-earning fields explain this difference in the relationship between expected earnings and field of study choice. This suggests that focusing on FiF students' expectations about earnings, own ability, and non-wage amenities is key to close important gaps in field of study choice between FiF and non-FiF students.

Our paper contributes to several strands of the literature. We primarily contribute to the literature on social mobility. Because of the key role of education, several papers have studied the role of socioeconomic status and FiF status on education outcomes, finding a variety of disadvantages connected to the latter. FiF students' lower educational attainment is partly due to lower perceived pecuniary and non-pecuniary benefits of enrolling at university, but also due to lower early human capital development (Boneva and Rauh, 2018; Adamecz-Völgyi et al., 2020; Blanden et al., 2023).[1] Many papers find important differences in field-of-study choices, with FiF students on average attending less prestigious universities and studying for degrees with lower labor market returns (e.g., Adamecz-Völgyi et al., 2023). This is even true when conditioning on high-ability samples (Shure and Zierow, 2023). For example, Stansbury and Rodriguez (2024) show that FiF students face strong disadvantages in the academic labor market in the US.[2]

To our knowledge, our paper is the first to document a sizable FiF gap in earnings expectations already at the start of university studies. We find that the FiF gap in earnings expectations primarily exists for low-ability FiF students at university. We contribute to the literature by showing that around two thirds of the documented gap is attributable to differences in student and field of study characteristics. Using panel data on current labor market participants, we show that the gap we find is similar, but somewhat smaller than the actual gap in earnings between FiF and non-FiF students.[3]

A second strand we contribute to assesses the role of earnings expectations of students. This strand usually focuses on the gender wage gap. Several papers have documented a gender gap in expected earnings that aligns with the realized gender gap in the labor market (see, e.g., for Germany Briel et al., 2022; Leibing et al., 2023; Kiessling et al., 2024). Filippin and Ichino (2005) survey university students in Italy and find a significant gender gap in expected as well as realized earnings one year after graduation. Wiswall and Zafar (2021) survey college students during their studies and six years after graduation. Own earnings, as well as a range of other labor-market related outcomes, are well anticipated by students. Reuben et al. (2017) find that overconfidence and competitiveness relate to higher expected earnings. Many of these studies use samples of university students, often from highly selective institutions.

We contribute to this literature by providing the first evidence on differences in labor market expectations by FiF status at the very beginning of students' university careers. This difference is smaller than the gender gap in earnings expectations in our data. In addition, in contrast to the gender gap, the FiF gap can be explained by student characteristics and study fields to a much larger extent. Nevertheless, FiF students also seem less confident about their ability especially in high-paying fields, in line with discussions about socioeconomic determinants of self-confidence

[1]Other papers study the role of FiF status conditional on having attended university. Edwards et al. (2022) find that in Australia, FiF students enter university with lower cognitive but similar non-cognitive skills. They show that these students have lower grade-point averages and are more likely to drop out after the first year at university, a finding that is driven by female FiF students.

[2]FiF status is also associated with lower earnings in the UK, although this finding only holds for women (Adamecz-Völgyi et al., 2023).

[3]About 60% of survey participants gave us consent to link the survey data to administrative data once the students enter the labor market. This will allow us to compare the gap in earnings expectations to the realized gap for the same students.

(e.g., Friedman and Laurison, 2019). In addition, we study students at a large public university that is arguably more representative of university students at large than the samples from highly selective US institutions.

Third, our paper contributes to the literature on students' expectations and sorting decisions. Expected earnings have been shown to matter for sorting into occupations and study fields. For example, Arcidiacono et al. (2020) study a sample of undergraduate students at Duke University and find that expected earnings positively relate to occupational choices after studying. Earnings expectations especially seem to play a role for sorting into business and law-related occupations. Wiswall and Zafar (2018, 2021) conduct similar surveys and find that earnings expectations play a role in students' enrollment decisions: students seem to choose to major in the field in which they consider their prospective earnings to be highest, all else equal. We contribute to this literature by showing that expectations about earnings seem to matter much less for FiF students. We show that this is due to FiF students' beliefs about their own ability and their expectations about correlated non-wage amenities in high-paying fields. Relative to the prior literature, our sample is very large, covering around half of each of two cohorts of students entering the university from which the data is drawn.

2 Setting and Data

2.1 Institutional Details and Data Collection

Institutional Setting. Our data cover undergraduate students at a large public university in Germany with about 40,000 students in total.[4] The university offers a large variety of undergraduate programs from all academic disciplines, including teacher training, arts and humanities, social sciences, natural sciences, engineering, business and economics, law, and medicine. At undergraduate level, three-year degrees are most common, but there are also programs that take longer to complete (in particular in the fields of teacher training, law, and medicine).

The university is typical for non-elite schools in many OECD countries (Lovenheim and Smith, 2023). It offers tertiary education to students that have earned a university entrance qualification by completing a total of 12 or 13 years (depending on the school type) of primary and secondary school education. Because most larger cities in Germany host a university offering similar undergraduate programs broadly at the same level, the majority of students are from the region.[5] Importantly, students

[4]This section is borrowed from Adler et al. (2025).

[5]More than 80 percent of the university's undergraduate students obtained their university entrance qualification in the federal state where the university is located. Around 70 percent are from places that are in commuting distance to the university.

about to begin their undergraduate studies enroll in a specific field of study, as is typical at universities outside North America and Australia (Lovenheim and Smith, 2023). Credits earned are field-specific and usually cannot be transferred if a student changes to a different field. As a result, we consider a setting where students choose a field of study rather than a university only, and where field choice is a consequential decision to be made right at the beginning of tertiary education. At an individual level, there are few limitations on field choice. Notably, at this university, only degrees in Medicine, Psychology, and Pharmacy have entry barriers (mostly based on the GPA of the university entrance qualification) that materially restrict entry into these fields.

At the university we consider, the academic year comprises two semesters. Most study programs start in the fall semester, which begins in October and ends in February. The summer term begins in April and ends in July.

Survey Data. Our data come from a survey which we conducted among first-year students at the university we study in 2022 and 2023.[6] All eligible students were invited via e-mail to participate in our survey in the second week after they started their university studies. Students who completed the survey were paid €10 in 2022 and €15 in 2023; they could choose to receive their payment either via bank transfer or as a voucher for a large international online retailer. The survey was implemented in oTree (Chen et al., 2016).

The survey comprised several blocks, which were shown to the participants in semi-randomized order. We elicited economic preferences and psychological personality traits in the first part for a companion paper (Adler et al., 2025), followed by study motives and questions related to subjects' parental background in the second part. Finally, we asked students about their earnings expectations, relative ability ranks, and expectations about non-wage amenities in their future job.

"First-in-Family" Status. We elicited the status of being a "first-in-family" student in our survey by asking for the highest degree obtained by the respondent's mother and father. We classify an individual as an FiF student if they report that none of their parents have a university degree. We acknowledge that this classification may also capture differences between FiF and non FiF students that go beyond parents' educational attainment. For instance, our FiF classification might (partly) capture broader differences in socio-economic status.

Expected Earnings, Ability, and Non-Wage Amenities. We elicited earnings expectations in two ways. First, we asked subjects to assume they would be working full time at the ages of 30 and 40, respectively. We then asked them to state their

[6] Note that we have been collecting this survey data since 2020. Earnings expectations, however, were only elicited from 2021 onward, with a change in the elicitation method from 2022 onward, which is why we only use data from the two survey waves 2022 and 2023 in this paper.

expected monthly gross salaries at both points in time.[7] We call these data *unconditional earnings expectations* because they are not explicitly linked to any degree the individual might obtain. In a second step, we showed participants eight study field categories plus a ninth category labeled as dropout or not obtaining any degree. We then again asked for expected gross monthly earnings provided the subjects would work full time but under the additional assumption that they would obtain a degree in each of the listed study fields. We call these *conditional earnings expectations*.

To what extent are the earnings expectations we elicit accurate? Recent literature on this question has found mixed results (e.g., Arcidiacono et al., 2020; Briel et al., 2022; Diaz-Serrano and Nilsson, 2022). To add to this debate, we compare earnings expectations of students in our sample to actual earnings in two fields where we expect a tight link between field of study and occupation and where occupational earnings are largely regulated. First, almost all teachers in Germany work for the specific state their school is in and are paid according to the pay scheme for civil servants. When we restrict our expectations data to students who study to become teachers, we find that their average earnings expectations at age 30 are around 4350 Euros per month. This is fairly comparable to actual earnings for early career teachers in Bavaria at the time, the state in which this university is and where most students will end up working.[8] Interestingly, studying the earnings expectations conditional on completing a teaching degree for students in other fields reveals that students generally have accurate expectations about early career earnings in this occupation, with their expectations being statistically identical to the expectations of teacher students. Conditional on completing medicine, the earnings expectations at age 30 of students are around 6300 Euros, while the actual monthly gross base earnings at the time for an early career medical doctor were around 5000 Euros (in 2022, see `https://oeffentlicher-dienst.info/c/t/rechner/aerzte/kommunal?id=tv-aerzte-vka-2022&matrix=1`). Students who actually study medicine expect somewhat lower earnings, at around 6000 Euros, than students who do not study medicine. Given that medical doctors earn extra pay for night and weekend shifts beyond their base salary, these earnings estimates again seem quite accurate. All in all, this suggests that at least in those fields where earnings information is easily available, students in our sample have rather accurate expectations. Our results also suggest that in these fields, FiF and non-FiF students do not significantly differ in their earnings expectations.

[7] To reduce noise in the data, participants were asked to select one out of 14 boxes corresponding to their expected monthly gross salary, with salary ranging from €2,000 to over €8,000 in €500 intervals. We provided definitions of full time work and gross salary.

[8] For example, according to the pay scheme that was in place in October 2022, early career teachers earned between almost 4000 and 4600 Euros per month, depending on the school type (and thus the pay grade) they worked in. See `https://oeffentlicher-dienst.info/c/t/rechner/beamte/by/a?id=beamte-bayern-2021&matrix=1`.

We additionally asked for subjects' expected relative ability in each of the study fields. We asked subjects to imagine they would major in each of the listed study fields, and asked for their perceived rank (100 being the best and 1 being the worst) among all other students majoring in the same study field. To what extent is it plausible to assume that students have good information about their counterfactual ability across fields? Appendix A.2 shows that when comparing students' actual percentile in the distribution of high-school GPA grades across fields, given the students in the field, to their belief about their ability in this field, the average deviations are relatively small (outside medicine around 6-7 percentiles on average). This likely stems from the fact that the survey takes place when newly enrolled students spend their first days at university. Most of them have finished high school just a few months earlier. Moreover, the survey refers to broad study fields only, not specific study programs. We believe it is plausible that from close interaction with their peer group at high school, most students have at least some idea about the ability distribution among prospective students aiming at one of these broadly defined fields.

Lastly, we elicited students' expectations about three non-wage amenities conditional on graduating in each of the nine potential study fields. Individuals were asked to indicate how likely they considered the respective amenities to be fulfilled conditional on graduating in each of the listed study fields. The non-wage amenities we included are sufficient leisure, the possibility to have children, and enjoying work, all on an 11-point Likert scale.

Study Motives. We additionally asked nine questions regarding students' motivation to choose their study program. These questions were answered on 5-point Likert scales and covered pecuniary as well as non-pecuniary motives. For example, we asked subjects how important it was for their choice of study program to have fun during their studies or to meet people like or unlike themselves. Other questions related to job perspectives and we asked, e.g., how important the perspective of a safe or of a high–paying job was for the choice of study program.[9]

Administrative Data. We matched the survey data with administrative data from the university comprising information on students' high school GPA, nationality, and the study program they enrolled in. We were able to match all students who participated in the survey with the administrative records because students participated in the survey via their campus login, uniquely identifying them in the administrative data.

Participation and Sorting. Of all invited students, 3,566 completed the survey (1,941 in 2022 and 1,625 in 2023), which corresponds to a mean take-up rate of 47.5% (50%

[9]The remaining questions were on the following study motives: interest for the study field, regional attachment, leisure time and enough time for family, having a job that is useful for society.

in 2022 and 45% in 2023). Around 40 percent of students completing the survey are FiF university students (1,428 students).[10] Table A1 in the Online Appendix shows selection into survey participation across waves. Survey participants have slightly better high school GPAs, are somewhat less likely to be of foreign nationality, and have a slightly different field of study composition than non-participants in both waves. All in all, however, the differences are rather small. Our data thus represents the overall student population fairly well.

Sample Characteristics. Table 1 provides an overview of the sample characteristics by displaying differences in means by FiF status for all survey items and student characteristics obtained from the administrative data. Table 1 shows that FiF students on average expect considerably lower monthly earnings both at age 30 and 40. There are also differences in student characteristics. For instance, FiF students in our sample have significantly lower high-school GPA.[11] FiF students are also slightly older, and we observe more female students in the group of FiF students. Finally, there are more students with a migration background and of foreign nationality in the group of non-FiF students.[12] There is also some heterogeneity between FiF- and non-FiF students regarding the motives for study program choice. It seems more important to FiF students to meet people who are "like them." FiF students also have a stronger motive for a future job with more leisure and allowing to have a family.

Turning towards self-perceived ability, students differ by FiF-status in the assessment of their own relative ability across (actual and counterfactual) study fields. FiF students consider themselves less able in all study fields, with the mean gap being significant in law, medicine, natural sciences, engineering, and business &

[10] This is comparable to other data on the share of FiF students at German universities. For example, a survey-based report (Stifterverband, 2021) shows that around 47% of students in Germany are FiF students. According to the report, in the parent generation, the parents of around 72% of children did not attend university. Thus, while the conditional probability of attending university is much lower for FiF students in Germany, the unconditional probability of a random student being an FiF student is almost the same as this student having parents who attended university. Data from the German Centre for Higher Education Research and Science Studies (DZHW, Apolinarski et al., 2019) shows that in 2016, the share of FiF students among first-year students with a German university entry qualification is 41 percent.

[11] In the German system, the GPA runs from 1 to 4 and lower numbers indicate better grades. We have recoded high-school GPA such that higher values indicate better grades, with 4 now being the best and 1 the worst grade.

[12] This could either be driven by sorting into university differing by group or by university degrees in foreign countries covering more occupations, given the large share of Germans in apprenticeships. We define migration background as having a second mother tongue apart from German, and the three most commonly listed second languages in our survey are English, Turkish, and Russian. We see that in the group of non-FiF students, the share of individuals with English or Russian as a second mother tongue is significantly higher than in the FiF sample (5% vs. 3% and 6% vs. 4% for English and Russian, respectively). There are no significant group differences regarding Turkish as a second mother tongue (around 4% of students in both groups list this as their second mother tongue).

economics, all of which are high-paying fields on average.[13] Note, however, that most of these fields are also STEM fields. Thus, it is well possible that the reason for these differences is that FiF students perceive themselves to be weaker in mathematical or technical skills than students from academic backgrounds.

All in all, FiF students have lower earnings expectations on average but also differ in important characteristics from non-FiF students. To which extent these differences in earnings-relevant characteristics explain the differences in earnings expectations by FiF status is what we investigate next.

Comparison to Representative Data. Since our data only covers one university, a relevant question is to what extent it is comparable to the German university landscape more broadly. To provide some information on this, we use the scientific use files from the 21st social survey of the German Centre for Higher Education Research and Science Studies (DZHW) which was conducted in 2016 (Apolinarski et al., 2019). Our data restrictions leave us with 4,543 students in this data set.[14]

Table 2 shows a comparison of some characteristics that are comparably contained in both datasets for these students by FiF status. We first investigate the personal characteristics of students across these two datasets. As the table shows, our data is broadly comparable to data from German universities. In particular, the share of women and the average high-school GPA among students by FiF status in the DZHW data are qualitatively very close to the values in our data.[15] Next, we investigate the field of study distribution, again by FiF status. Our data is very close to the DZHW data in medicine, mathematics and natural sciences, engineering and teacher training.

[13] Using data from Adler et al. (2025) on students' self-esteem, we find that FiF students have lower self-esteem. In line with this, when we compare students' self-perceived ability with their rank in the distribution of high-school GPAs by field, we find that they are more confident in their own field than non-FiF students, but less confident outside their own field. FiF students thus seem to select into fields they feel competitive in and confident about their ability relative to other students. See Appendix A.2 for details.

[14] To make the samples broadly comparable, we use the dataset on students with a German university entrance qualification, restrict to students in Germany, to students at universities, to students in their first year, and to those where all necessary information for a comparison is present. The reason we ignore the data set of students in Germany with a non-German university entrance qualification is that this dataset lacks important student characteristics for our comparison, such as the semester in which students currently are enrolled in or their high-school GPA.

[15] There are two discrepancies, however. First, our data contains more students with a migration background than common in Germany. This can have two sources: First, more recent cohorts in Germany contain more students with a migration background, reflecting trends in German demography. Second, the region around the university we study used to have a strong industrial base, leading to a higher share of migrants in the population. For example, the city the university is located in is one of few municipalities in Germany where the majority of the local population has some migration background. The second discrepancy is that the share of foreigners is somewhat higher in our data. The reason for this likely is that in the DZHW data, we had to focus on the dataset that only contains students with a German university entrance qualification, as explained above.

Table 1: Sample Characteristics

	Non-FiF	FiF	Difference
Expected Earnings			
Unconditional, Age 30	4695.63	4438.19	257.44***
Unconditional, Age 40	5665.24	5321.53	343.71***
Student Characteristics			
Age in years	20.84	21.37	-0.52***
High school GPA	293.44	276.18	17.25***
Migration background	0.33	0.24	0.09***
Foreign nationality	0.06	0.04	0.02***
Female	0.55	0.61	-0.07***
Study Motives			
Interest in subject	3.72	3.73	-0.01
Fun	3.28	3.27	0.02
Regional attachment	2.47	2.53	-0.06
Meet people like me	2.95	2.95	-0.00
Meet people unlike me	2.25	2.19	0.05
Safe job	3.35	3.43	-0.08***
High-paying job	3.16	3.17	-0.01
Family and leisure compatible job	2.84	2.96	-0.12***
Socially relevant job	2.88	2.92	-0.04
Relative Ability			
Teacher Training	65.49	64.90	0.58
Law	54.85	53.04	1.81*
Medicine	53.67	50.70	2.96***
Humanities	55.28	53.97	1.30
Social Sciences	58.20	56.86	1.33
Natural Sciences	57.60	53.42	4.19***
Engineering	53.45	49.70	3.75***
Business & Economics	57.42	54.80	2.62**
Amenity Preferences			
Sufficient leisure	6.65	6.77	-0.13*
Have kids	6.95	7.10	-0.15
Enjoy work	8.73	8.83	-0.11**
Observations	1363	1991	3354

Note: This table displays means of all variables by FiF status, with asterisks denoting significant differences between the means (measured with a two-sided t-test). * $p<0.1$ ** $p<0.05$ *** $p<0.01$. Study motives were elicited on a 1-5 scale, with higher levels indicating stronger importance of the respective motive for study program choice. Preferences for non-wage amenities were elicited on a 1-11 scale, with higher levels indicating stronger preference for the respective amenity. Expected earnings are unconditional on completing a degree in a certain study field and are to be understood as gross monthly earnings. Relative ability was elicited on a 1-100 scale. We modified the variable such that the higher the displayed value, the higher the student's perceived relative rank. We recoded the German high-school GPAs such that higher values indicate better grades.

There are some discrepancies in law, business and economics, and social sciences and in arts and humanities, but our data still covers these fields to a relevant extent.

Table 2: Comparison to Representative Data

	FiF	Non-FiF	Difference
Representative Data			
Migration Background	0.20	0.15	-0.05***
High-school GPA	2.77	3.05	0.28***
Foreign	0.03	0.02	-0.02***
Female	0.62	0.58	-0.04***
Arts & Humanities	0.14	0.10	-0.04***
Law, Business and Economics, and Social Sciences	0.30	0.31	0.01
Mathematics and Nat. Sci.	0.15	0.15	0.01
Medicine	0.05	0.10	0.04***
Engineering	0.15	0.19	0.03***
Teacher Training	0.17	0.12	-0.05***
Our Data			
Migration background	0.24	0.33	0.09***
High school GPA	2.76	2.93	17.25***
Foreign	0.04	0.06	0.02***
Female	0.61	0.55	-0.07***
Arts & Humanities	0.08	0.06	0.02**
Law, Business and Economics, and Social Sciences	0.40	0.39	0.00
Natural Sciences	0.12	0.14	-0.02
Medicine	0.08	0.10	-0.02
Engineering	0.12	0.19	-0.07***
Teacher Training	0.19	0.12	0.08***

Note: This table displays means of selected student characteristics in our own data and in representative data of first-year university students with a German university entrance qualification by the German Centre for Higher Education Research and Science Studies (Apolinarski et al., 2019) by FiF status.

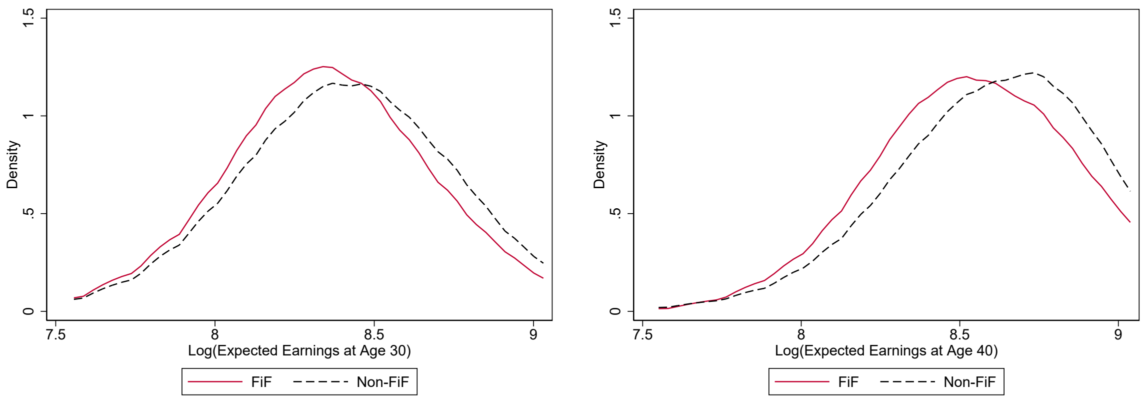
3 Differences in Earnings Expectations by FiF Status

3.1 Distribution of Earnings Expectations by FiF Status

We now examine the gap in (unconditional) expected earnings between FiF and non-FiF students in more detail. As can be seen in Table 1, FiF students in our sample on average expect significantly lower earnings both at age 30 and at age 40. Figure 1 provides descriptive evidence of the distribution of expected earnings at these two points in students' lives. The two kernel density plots, especially the plot for earnings expectations at age 40, show a clear shift of FiF students' earnings expectations to the left, suggesting that FiF students expect to earn considerably less than their non-FiF classmates at age 30 as well as at age 40.[16]

[16]Kolmogorov–Smirnov tests for equality of distributions suggest that the two distributions indeed differ from each other, $p < 0.001$.

Figure 1: Kernel Density Plots of Unconditional Expected Earnings by FiF Status



Note: This figure shows kernel density plots of log unconditional expected monthly earnings by first-in-family status (bandwidth $= 0.15$, $n = 3354$). We display expected earnings at age 30 and 40 in the left and the right graph, respectively. Red solid lines and black dashed lines denote expectations of FiF and non-FiF students, respectively.

3.2 Empirical Approach

To examine the mean gap in earnings expectations between FiF and non-FiF students, we estimate the following equation:

$$\mathbf{y}_{is} = \alpha + \beta \cdot FiF_{is} + \mathbf{D}_s'\sigma + \mathbf{X}_{is}'\gamma + \mathbf{M}_i'\delta + \mathbf{A}_i'\theta + \epsilon_{is}, \tag{1}$$

where $\mathbf{y}_{is}$ denotes the outcome of student i in study field s (i.e., log expected earnings at age 30 and 40) and FiF_{is} denotes FiF status. We then gradually include more covariates when estimating equation (1). $\mathbf{D}_s$ are study field dummies, and $\mathbf{X}_{is}$ is a vector of student characteristics (labeled administrative controls and containing age, high-school GPA, dummy$=1$ if student has a migration background, i.e., a second mother tongue, and dummy$=1$ if student is female). The vectors $\mathbf{M}$ and $\mathbf{A}$ comprise the study motives and preferences for non-wage amenities elicited in our survey, respectively. Note that because FiF status arguably influences field of study choice, high-school GPA, study motives, and amenity preferences, one might view these covariates as "bad controls" in a regression of earnings expectations on FiF status. This partly reflects that FiF students differ from non-FiF students not only in terms of parental educational attainment, but also in other dimensions. We include the covariates to examine to what extent they explain differences in earnings expectations by FiF status. We estimate equation (1) by OLS and report robust standard errors. Finally, note that all survey responses were voluntary, which is why we lose some observations when including more controls.

3.3 Regression Results

Main results. Table 3 shows the results of the OLS regressions, adding covariates in each column. Regressing log unconditional expected earnings on only the FiF dummy suggests that FiF students expect to earn 5.3% (6.4%) less than their non-first-generation classmates when they are 30 (40) years old (column (1)). These differences diminish if we include covariates at age 30. For earnings expectations at age 40, the gap remains significant at the 1%-level even after including all controls. Controlling for all observables in our sample, we estimate a gap between FiF- and non-FiF students in expected earnings at age 30 (40) of 2.0% (2.8%).[17]

Table 3: FiF Gaps in Unconditional Earnings Expectations

	(1)	(2)	(3)	(4)	(5)
I. Outcome: Log EE, Age 30					
FiF	-0.0530***	-0.0334***	-0.0173*	-0.0193**	-0.0196**
	(0.0104)	(0.0100)	(0.0101)	(0.0099)	(0.0098)
Adjusted R^2	0.007	0.115	0.143	0.194	0.197
II. Outcome: Log EE, Age 40					
FiF	-0.0630***	-0.0374***	-0.0245***	-0.0274***	-0.0279***
	(0.0100)	(0.0093)	(0.0093)	(0.0090)	(0.0089)
Adjusted R^2	0.011	0.171	0.204	0.270	0.275
Observations	3354	3354	3354	3354	3354
Study Field FE	No	Yes	Yes	Yes	Yes
Student Characteristics	No	No	Yes	Yes	Yes
Study Motives	No	No	No	Yes	Yes
Amenities	No	No	No	No	Yes

Note: This table reports coefficients of a dummy denoting first-generation student status in regressions of log monthly earnings expectations (EE) at age 30 on study field fixed effects, student characteristics (age, gender, migration background, high-school GPA), study motives (have fun during studies, interest in subject, regional attachment, meet people like me, meet people unlike me, safe job, high-paying job, family and leisure compatible job, socially relevant job; 5-point Likert scales), and preferences for non-wage amenities (preferences for leisure, having children, having fun at the job; 11-point Likert scales). Robust standard errors in parentheses. * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

Heterogeneity by gender. In Table 4, we report differences in expected earnings by FiF status for female and male students. We find suggestive evidence that the observed gap is mainly driven by male students. The raw FiF gap at age 30 is around four to five percent for both genders and widens to six (five) percent for female (male) students at the age of 40. The gap at age 30 becomes insignificant for female students when including controls, whereas it stays borderline significant for male students and amounts to around 2.6 percent after including all controls. We observe a similar

[17] As a consequence of the higher expected group differences at 40 than at 30, we observe that FiF students expect their earnings to grow more slowly than non-FiF students do. This difference becomes insignificant when including control variables and amounts to less than one percentage point when including all covariates. Results are available upon request.

pattern for the FiF gap for both genders at age 40, with a gap of 3.0 percent for male students after including all controls. This somewhat contrasts recent literature that found that, at least in the United Kingdom, only female FiF students face wage penalties in the labor market, even when controlling for many characteristics such as degrees and non-cognitive skills (Adamecz-Völgyi et al., 2023). However, none of the differences by gender we find is statistically significant at conventional levels in interacted models (not shown).

Heterogeneity by student ability. We next examine how the FiF gap in expected earnings varies with students' cognitive ability. We proxy ability by the high-school GPA. We define subjects as "high-ability" ("low-ability") if their high-school GPA is below (above or equal) the median in our sample. Table 5 shows that the overall gap in expected earnings is primarily driven by students of low ability. Controlling only for FiF status, the coefficients suggest that high-ability students expect a smaller gap in earnings at both points in time. When including covariates, the gap in expected earnings reduces significantly and becomes insignificant for expected earnings both at age 30 and at age 40. Among low-ability students, however, FiF students expect to earn 7.2% (7.7%) less than their non-FiF peers when they are 30 (40). This gap shrinks to 3.3% to 4.2% when including all covariates, but remains significantly different from zero. These differences are statistically significant at conventional levels in interacted models (not shown).

Table 4: Mean Differences in Expected Earnings—Heterogeneity by Gender

	(1)	(2)	(3)	(4)	(5)
I. Outcome: Log Expected Earnings, Age 30					
Panel A: Female students					
FiF	-0.0423***	-0.0220*	-0.0133	-0.0173	-0.0171
	(0.0131)	(0.0127)	(0.0130)	(0.0128)	(0.0127)
Observations	1932	1932	1932	1932	1932
Adjusted R^2	0.005	0.089	0.094	0.151	0.152
Panel B: Male students					
FiF	-0.0501***	-0.0455***	-0.0256	-0.0252	-0.0264*
	(0.0165)	(0.0158)	(0.0159)	(0.0156)	(0.0155)
Observations	1422	1422	1422	1422	1422
Adjusted R^2	0.006	0.115	0.142	0.189	0.194
II. Outcome: Log Expected Earnings, Age 40					
Panel A: Female students					
FiF	-0.0576***	-0.0303**	-0.0238*	-0.0281**	-0.0280**
	(0.0128)	(0.0120)	(0.0124)	(0.0120)	(0.0119)
Observations	1932	1932	1932	1932	1932
Adjusted R^2	0.010	0.137	0.141	0.212	0.215
Panel B: Male students					
FiF	-0.0491***	-0.0415***	-0.0280**	-0.0252	-0.0301**
	(0.0151)	(0.0139)	(0.0142)	(0.0156)	(0.0135)
Observations	1422	1422	1422	1422	1422
Adjusted R^2	0.007	0.174	0.186	0.189	0.260
Study Field FE	No	Yes	Yes	Yes	Yes
Student Characteristics	No	No	Yes	Yes	Yes
Study Motives	No	No	No	Yes	Yes
Amenities	No	No	No	No	Yes

Note: This table reports coefficients of a dummy denoting first-generation student status in regressions of log earnings expectations at ages 30 and 40 on study field fixed effects, student characteristics (age, gender, migration background, high-school GPA), study motives (have fun during studies, interest in subject, regional attachment, meet people like me, meet people unlike me, safe job, high-paying job, family and leisure compatible job, socially relevant job; 5-point Likert scales), and preferences for non-wage amenities (preferences for leisure, having children, having fun at the job; 11-point Likert scales). We split the sample by gender. Robust standard errors in parentheses. * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

Comparison to actual earnings gap by FiF status. How does the gap in expected earnings by FiF status in our data compare to actual earnings differences by FiF status in the labor market? To inform this question, we leverage data from the German Socio-Economic Panel (Goebel et al., 2019; SOEP, 2023).[18] Using this data, we focus on respondents (around) age 30 and 40 that have attended university and work full

[18] Once the respondents to our survey enter the labor market, we will be able to compare their earnings expectations with actual realizations for around 60% of the sample, namely those students who gave consent to link their data to administrative labor market data from the IAB Institute for Employment Research.

Table 5: Mean Differences in Expected Earnings—Heterogeneity by Ability

	(1)	(2)	(3)	(4)	(5)
I. Outcome: Log Expected Earnings, Age 30					
Panel A: Low-ability students					
FiF	-0.0722***	-0.0525***	-0.0357**	-0.0336**	-0.0332**
	(0.0152)	(0.0145)	(0.0147)	(0.0145)	(0.0144)
Observations	1637	1637	1637	1637	1637
Adjusted R^2	0.013	0.124	0.149	0.183	0.189
Panel B: High-ability students					
FiF	-0.0300**	-0.0097	0.0037	-0.0033	-0.0035
	(0.0145)	(0.0139)	(0.0137)	(0.0134)	(0.0134)
Observations	1717	1717	1717	1717	1717
Adjusted R^2	0.002	0.110	0.141	0.209	0.209
II. Outcome: Log Expected Earnings, Age 40					
Panel A: Low-ability students					
FiF	-0.0793***	-0.0549***	-0.0441***	-0.0336**	-0.0422***
	(0.0145)	(0.0135)	(0.0137)	(0.0145)	(0.0131)
Observations	1637	1637	1637	1637	1637
Adjusted R^2	0.017	0.185	0.204	0.183	0.262
Panel B: High-ability students					
FiF	-0.0399***	-0.0142	-0.0016	-0.0095	-0.0101
	(0.0138)	(0.0129)	(0.0126)	(0.0121)	(0.0121)
Observations	1717	1717	1717	1717	1717
Adjusted R^2	0.004	0.157	0.206	0.291	0.292
Study Field FE	No	Yes	Yes	Yes	Yes
Student Characteristics	No	No	Yes	Yes	Yes
Study Motives	No	No	No	Yes	Yes
Amenities	No	No	No	No	Yes

Note: This table reports coefficients of a dummy denoting first-generation student status in regressions of log earnings expectations at ages 30 and 40 on study field fixed effects, student characteristics (age, gender, migration background, high-school GPA), study motives (have fun during studies, interest in subject, regional attachment, meet people like me, meet people unlike me, safe job, high-paying job, family and leisure compatible job, socially relevant job; 5-point Likert scales), and preferences for non-wage amenities (preferences for leisure, having children, having fun at the job; 11-point Likert scales). We split the sample at the median of students' high-school GPA. Robust standard errors in parentheses. * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

time, a restriction that makes the sample comparable to the hypothetical situation (expected earnings at age 30 and 40) in our survey. In the SOEP, we leverage responses of respondents on the highest educational degree of their parents to define FiF status. We then regress (log) earnings on a FiF dummy along with controls for gender and migration background, as well as birth cohort fixed effects to account for time trends and age differences.

We show the results from this analysis in Online Appendix Table A5. The table shows that the expected earnings gap by FiF status is remarkably similar to the gap we find in the SOEP data at age 30 without controls and only slightly lower than the gap we find in the SOEP when adding comparable controls.[19] At age 40, students who took our survey expect lower gaps in earnings by FiF status than we find in the SOEP.[20] We can only speculate whether these differences reflect actual misperceptions, unaccounted expectations about other determinants of wages such as non-wage amenities, or expectations about time trends.

Comparison to gender differences in expected earnings. The raw FiF difference in expected earnings of about five (six) percent at age 30 (40) amounts to around half of the gender gap in earnings expectations in the same data (see Table A6). In contrast to the FiF gap, which is more pronounced for expectations at the age of 40, we see that females expect to earn 12 (14) percent less at age 30 (40) than their male peers. These gaps are similar to recent findings on gender gaps in early wage expectations in Germany (Briel et al., 2022; Leibing et al., 2023; Kiessling et al., 2024). Conditional on including all control variables, the FiF gap amounts to around one quarter of the gender gap in earnings expectations both at age 30 and at age 40. Thus, we can explain larger shares of the FiF earnings expectations gap than of the gender earnings expectations gap using our student characteristics.

3.4 Oaxaca–Blinder Decomposition

In a next step, we examine which parts of the observed gap between FiF and non-FiF students can be explained by student characteristics as well as study motives

[19]Note that the estimates we provide in this table based on the survey slightly differ from our main estimates. The reason is that in the SOEP, we do not have any information on high-school GPA or (expectations about) non-wage amenities or study motives. To keep the estimates comparable, we thus do not control for these variables in the table.

[20]The gap we find in the SOEP is larger than comparable estimates from other countries. Recently, for example, Stansbury and Rodriguez (2024) show that PhD graduates in the US that stem from non-academic backgrounds earn around 1.6 log points less in industry and up to 2.7 log points less in academia, conditional on demographic characteristics. There is no earnings gap in government and non-tenure track education. The discrepancy may be due to a lack of control variables in the SOEP on earnings determinants such as high school GPA or productivity, which may differ by FiF status.

and preferences for non-wage amenities. To do so, we conduct Oaxaca–Blinder decompositions of unconditional expected monthly earnings at age 30 and 40 (Oaxaca, 1973; Blinder, 1973; Blau and Kahn, 2017). To this end, we estimate two separate OLS regressions per outcome for FiF and non-FiF students, respectively:

$$\mathbf{y}_{isf} = \mathbf{X}'_{isf}\beta_f + \epsilon_{isf} \tag{2}$$

$$\mathbf{y}_{isn} = \mathbf{X}'_{isn}\beta_n + \epsilon_{isn} \tag{3}$$

Here, $\mathbf{y}_{is}$ again denotes the outcome of interest. For simplicity, we combine all student information in the vector $\mathbf{X}'_{is}$. β_f and β_n then denote coefficients for the FiF and non-FiF subsample, respectively. Denoting OLS estimates with b_n and b_f and indicating group means by bars, we can rewrite the difference in group means as follows (Blau and Kahn, 2017):

$$\bar{y}_{isn} - \bar{y}_{isf} = b_n\bar{X}_{isn} - b_f\bar{X}_{isf} = b_n(\bar{X}_{isn} - \bar{X}_{isf}) + \bar{X}_f(b_n - b_f)$$

With this reformulation, the difference in group means is split into an explained and an unexplained part. $b_n(\bar{X}_{isn} - \bar{X}_{isf})$ denotes the explained part and can be interpreted as the difference in characteristics (or endowments, see, e.g., Blau and Kahn, 2017) evaluated with the FiF coefficient. $\bar{X}_f(b_n - b_f)$ is the unexplained differential which stems from a difference in coefficients.

Figure 2 shows the results of this decomposition. The results show that we can explain around 61% (56%) of the differences through endowments and 35% (44%) through differences in coefficients at age 30 (40). More precisely, if we applied the coefficients of FiF students to non-FiF students' wage expectations, these would be lower by around 1.8 log points at age 30 and around 2.7 log points at age 30. If non-FiF students had FiF students' endowments, their earnings expectations would be lower by 3.3 log points at age 30 and 3.6 at age 40, instead.

A large share of the observed gap can be explained by group differences in a few student characteristics (e.g., high-school GPA, gender, migration background). These differences in observable characteristics account for approximately half of the explained part of the observed difference at age 30 and 40% at age 40. Another factor that contributes significantly to explaining the observed difference between FiF- and non-FiF students is field of study. Around a fourth of the observed gap can be explained by study field choice or, put differently, a third (half) of the explained part of the observed difference at age 30 (40). These results are in line with findings in the literature that FiF students choose less prestigious degrees, on average (e.g., Adamecz-Völgyi et al., 2023; Shure and Zierow, 2023).

Figure 2: Oaxaca–Blinder–Decomposition of Unconditional Earnings Expectations

Note: This figure shows the results of Oaxaca–Blinder decompositions of log unconditional expected earnings at age 30 and 40 in the upper and lower graph, respectively. The upper part in both graphs provides information on the FiF difference and the respective shares that can be explained by differences in endowments, coefficients, or the interaction of the two (the displayed percentages always refer to the overall difference). The lower part in both graphs breaks down the share that can be explained by differences in endowments from the upper part into its components and their contributions to that share of the difference. Student characteristics are age, gender, migration background, and high-school GPA. Study motives are: have fun during studies, interest in subject, regional attachment, meet people like me, meet people unlike me, have a safe job, have a high-paying job, have a family and leisure compatible job, have a socially relevant job; all asked on 5-point Likert scales. Preferences for non-wage amenities comprise preferences for leisure, preferences for a job that enables one to have children, and preferences for having fun at the job; all asked on 11-point Likert scales.

Comparison to the decomposition of the gender gap. For comparison, Appendix Figure A1 reports an Oaxaca–Blinder decomposition of the gap in earnings expectations by gender. The main finding from this exercise is that our data explain only 30% (21%) of the gender gap at age 30 (40), much less than the 63% (56%) in the case of the gap by FiF status. As with the FiF gap, study fields contribute substantially to explaining the gender gap. They explain up to 21% of the observed gender difference or, put differently, contribute up to 80% of the explained part. Note that our results suggest that the gender gap in expectations would be even slightly larger if female and male students were identical in the characteristics we draw from the administrative data and if they had the same preferences for non-wage amenities.

To summarize this section, our data documents a pronounced FiF gap in students' expected earnings at age 30 and 40. This gap seems similar, albeit somewhat smaller than the actual FiF earnings gap among university graduates in the labor market. The gap seems to be larger for low-ability and for male students. It decreases somewhat but remains significant when including a wide set of covariates, at least at age 40. Our data explains approximately two thirds of the observed FiF difference, with around 37% (44%) of the observed gap at age 30 (40) remaining unexplained. Most importantly, personal characteristics and field of study choice explain over 90% of this explained part. While we naturally cannot say anything about the origins of the unexplained part, a possible candidate is lower confidence of FiF students (e.g., Friedman and Laurison, 2019).

4 Sorting Conditional on Earnings Expectations

In the previous sections, we documented a distinct FiF gap in expected earnings which is robust to the inclusion of. In a next step of our analysis, we want to examine how labor market expectations relate to study field choice.

4.1 Empirical Approach

Similar in spirit to Arcidiacono et al. (2020), we make use of the counterfactual earnings expectations we elicited in our survey. Although in our setting the sorting into fields of study has already taken place, we think it is a useful exercise to gauge the consequences of differences in earnings expectations for field of study choice. This is because the survey took place at the very beginning of the students' time at university (i.e., in the first week of their first semester). As a result, it seems justified to assume that the students' survey responses were affected by their field of study environment.

In the survey, we asked students about their expected earnings at age 30 under the condition of obtaining a degree in eight different study fields. Additionally, we elicited the students' perceived relative ability in comparison to their peers in each of the study fields. This ability measure ranges between 1 and 100, with 1 (100) indicating the lowest (highest) rank (i.e., higher numbers indicating higher perceived relative ability). Similarly, we elicited students' expectations regarding non-wage amenities conditional on graduating in the different fields. For each of three amenities (much leisure, having the chance to start a family, and enjoying work), we asked subjects to indicate how likely (on a scale from 0 to 100) they consider these amenities to be present conditional on graduating in each of the hypothetical study fields. From this, we create a panel (individual by study field). The counterfactual expectation and ability data as well as the students' actual study field of choice allow the estimation of an individual fixed–effects model:

$$y_{is} = \alpha_i + \beta \cdot Ln(Earnings)_{is} + \gamma \cdot Ability_{is} + \eta \cdot Amenities_{is} + \epsilon_{is}, \tag{4}$$

where y_{is} is an indicator taking on value one for student i's actual field of study and zero for all counterfactual fields s. We regress this indicator on log expected earnings by field, relative ability by field, and amenity expectations by field and cluster all error terms at the student level. Given our estimation equation, we can interpret our coefficients of interest as the increase in likelihood of a given field being the student's actual field given a one percent increase in expected earnings, an increase in expected relative ability by one (i.e., an upward shift in rank by one), or an increase in the probability of a certain non-wage amenity to be present by one percentage point, relative to students' counterfactual expectations for the other fields.

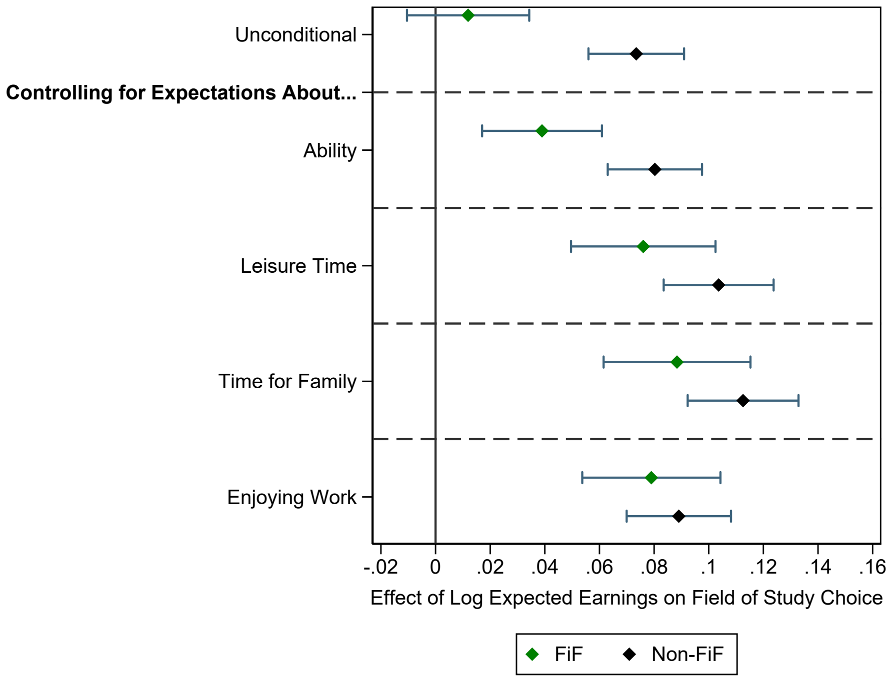
Our main coefficient of interest is β, the impact of one percent higher expected earnings on the likelihood that this is student i's field of study. The reason is that we want to investigate how differences in expected earnings relate to field of study choice across FiF and non-FiF students. We subsequently add controls for students' ability perceptions and expectations about non-wage amenities to investigate whether beliefs about the correlation between own ability, non-wage amenities, and earnings across fields affects field of study choice differentially by FiF status.

4.2 Results

Figure 3 reports the main results from this analysis. It shows how FiF and non-FiF students' earnings expectations across fields relate to their actual choice of field of study. The first coefficient (green diamond) shows that if we do not condition on perceived ability and amenities, FiF students' choices of a field of study are not systematically related to their earnings expectations. By contrast, the second coefficient (black diamond) shows that non-FiF students tend to choose fields with systematically higher expected earnings. We next control for students' perceived ability across fields. Whereas this leaves the relationship between expected earnings and field of study more or less unchanged for non-FiF students, the coefficient for FiF students increases and becomes statistically significant. This suggests that perceived own ability and expected field-specific earnings are negatively correlated among FiF students, in line with differences in major and university choice across FiF and non-FiF students conditional on observables found in the literature. Thus, perceptions of own ability seem to matter substantially when explaining differences between FiF and non-FiF students. Interestingly, when analyzing how this perception relates to actual ability of students by field (using students' high-school GPA rank as ability proxy), FiF students seem to be underconfident about their ability (see Appendix A.2).[21]

[21] This Appendix also shows that these gaps are neither present in perceived own ability nor in self-confidence between FiF and non-FiF students in students' own field of study. Thus, beliefs about own ability in counterfactual fields are driving this result.

Figure 3: Sorting Into Study Fields



Note: This figure reports coefficients of log conditional expected earnings in regressions of study field choice on expected earnings, relative ability, and non-wage amenities. We estimate an individual fixed-effects model with the artificial panel data created from our survey data set. We cluster standard errors at the student level and display 95% confidence bands. Each panel reports two coefficients, one for the FiF sample (in green) and one for the non-FiF sample (in black). The panels display coefficients of regressions with staggered inclusion of controls, i.e., the uppermost coefficients ("Unconditional") are from regressions of study field choice only on expected earnings. The second coefficient stems from a regression that adds perceived own relative ability, the third regression controls for perceived own relative ability and expectations about relative leisure time, etc.

The difference between FiF and non-FiF students in how strongly earnings expectations predict field-of-study choice decreases even further when controlling for students' expectations about the amount of leisure and expectations about the likelihood of having a family. While both of these measures are negatively correlated with expected earnings for non-FiF students (in line with a standard leisure-earnings trade-off, but also with models of compensating differentials like Rosen, 1986), this correlation seems substantially stronger for FiF students. Finally, both FiF- and non-FiF students seem to expect higher-paying fields to lead to work they would enjoy more, lowering the point estimate of the relation between expected earnings and field of study and making FiF and non-FiF students even more similar than before.

In summary, non-FiF students do seem to select into study fields conditional on expected earnings, while for of FiF students earnings expectations seem to matter much less. This pattern changes once we control for perceived own ability. Including expectations about non-wage amenities further closes the gap between FiF and non-FiF students in how their filed-of-study choice responds to earnings expectations. Including all counterfactual expectations, FiF and non-FiF students seem to select similarly on expected earnings.

5 Conclusion

In this paper, we shed light on FiF differences in expected earnings of first-year university students in a large public German university. In a survey, we elicited earnings expectations of students at specific ages as well as counterfactual earnings expectations in other fields, students' perceived relative ability, expectations about non-wage amenities, and study motives. Combining this unique survey data with administrative university records allows us to not only study FiF differences in expected earnings. It also allows us to gauge the consequences of differences in earnings expectations, namely how students select into study fields conditional on their expectations.

We document a pronounced FiF gap in expected earnings at various points in life. FiF students expect to earn less than their non-FiF peers at the age of 30 as well as the age of 40. This gap is driven by low-ability and by male FiF students. A decomposition of this gap into an explained and an unexplained part reveals that our data, most importantly personal characteristics and field of study choice, can explain approximately two thirds of the observed gap. Comparing these results to the gender gap in expected earnings, we see that the FiF gap is only half of the gender gap in size. The contribution of observable student characteristics is larger for the FiF than for the gender expectations gap.

Finally, we study selection decisions in an individual fixed-effects model and provide evidence that FiF students are less responsive to their earnings expectations when sorting into fields of study. Our findings suggest that this stems from FiF students' lower perceived ability in well-earning fields as well as FiF students' more negative expectations about non-wage amenities in such fields, like leisure and having the option to start a family. Our results suggest that FiF students' lower beliefs about own ability are not reflected in proxies of ability, suggesting lower self-confidence as a potential explanation.

Taken together, our findings suggest that policies targeting FiF students may help in closing existing gaps in perceived own ability and labor market-related expectations between FiF and non-FiF students. Our data imply that closing these gaps may help in shifting field of study choices among FiF students towards better-paying fields, with the potential of these changes leading to smaller earnings gaps between non-FiF and FiF graduates in the future. A major policy tool to induce such changes could be interventions that aim at correcting FiF students' perceived own ability in high-paying study fields. For instance, universities could survey students before they choose a study field and inform FiF students about their ability rank across fields or about the fact that people with a similar background tend to have lower expectations about own ability and later earnings. However, such interventions are paternalistic in nature and should be used cautiously, mainly because one cannot preclude that individual students are worse off if they adjust their perceptions in the intended direction. Therefore, future research should explore to what extent such interventions are effective in changing field-of-study choices among FiF students and assess the risk of leading students to make worse decisions.

References

Adamecz, A., M. Henderson, and N. Shure (2024): "Intergenerational educational mobility – The role of non-cognitive skills," *Education Economics*, 32, 59–78.

Adamecz-Völgyi, A., M. Henderson, and N. Shure (2020): "Is 'First in Family' a Good Indicator for Widening University Participation?" *Economics of Education Review*, 78, 102038.

——— (2023): "The Labor Market Returns to "First-in-Family" University Graduates," *Journal of Population Economics*, 36, 1395–1429.

Adler, K., F. Kosse, M. Nagler, and J. Rincke (2025): "Sorting into Career Paths on Personality and Preferences," *Mimeo*.

Apolinarski, B., K. Becker, P. Bornkessel, T. Brandt, S. Heissenberg, E. Middendorff, H. Naumann, and J. Poskowsky (2019): "21st Social Survey," (2016). Data Collection: 2016. Version: 2.0.0. Data Package Access Way: SUF: Download.

Arcidiacono, P., V. J. Hotz, A. Maurel, and T. Romano (2020): "Ex Ante Returns and Occupational Choice," *Journal of Political Economy*, 128, 4475–4522.

Blanden, J., M. Doepke, and J. Stuhler (2023): "Educational Inequality," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, S. Machin, and L. Woessmann, Elsevier, vol. 6, chap. 6, 405–497.

Blau, F. D. and L. M. Kahn (2017): "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature*, 55, 789–865.

Blinder, A. S. (1973): "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources*, 8, 436–455.

Boneva, T. and C. Rauh (2018): "Socio-Economic Gaps in University Enrollment: The Role of Perceived Pecuniary and Non-Pecuniary Returns," *CESifo Working Paper No. 6756*.

Briel, S., A. Osikominu, G. Pfeifer, M. Reutter, and S. Satlukal (2022): "Gender Differences in Wage Expectations: The Role of Biased Beliefs," *Empirical Economics*, 62, 187–212.

Chen, D. L., M. Schonger, and C. Wickens (2016): "oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments," *Journal of Behavioral and Experimental Finance*, 9, 88–97.

Diaz-Serrano, L. and W. Nilsson (2022): "The reliability of students' earnings expectations," *Labour Economics*, 76, 102182.

Edwards, R., R. Gibson, C. Harmon, and S. Schurer (2022): "First-in-Their-Family Students at University: Can Non-Cognitive Skills Compensate for Social Origin?" *Economics of Education Review*, 91, 102318.

Filippin, A. and A. Ichino (2005): "Gender Wage Gap in Expectations and Realizations," *Labour Economics*, 12, 125–145.

Friedman, S. and D. Laurison (2019): *The class ceiling: Why It Pays to Be Privileged*, Bristol University Press.

Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2019): "The German Socio-Economic Panel (SOEP)," *Jahrbücher für Nationalökonomie und Statistik*, 239, 345–360.

Henderson, M., N. Shure, and A. Adamecz-Völgyi (2020): "Moving On Up: 'First in Family' University Graduates in England," *Oxford Review of Education*, 46, 734–751.

Kiessling, L., P. Pinger, P. Seegers, and J. Bergerhoff (2024): "Gender Differences in Wage Expectations and Negotiation," *Labour Economics*, 87, 102505.

Leibing, A., F. Peter, S. Waights, and C. K. Spiess (2023): "Gender Gaps in Early Wage Expectations," *Economics of Education Review*, 92, 102398.

Lovenheim, M. and J. Smith (2023): "Returns to Different Postsecondary Investments: Institution Type, Academic Programs, and Credentials," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, S. Machin, and L. Woessmann, Elsevier North Holland, vol. 6, chap. 4, 187–318.

Oaxaca, R. (1973): "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14, 693–709.

Reuben, E., M. Wiswall, and B. Zafar (2017): "Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender," *Economic Journal*, 127, 2153–2186.

Rosen, S. (1986): "The Theory of Equalizing Differences," in *Handbook of Labor Economics*, ed. by O. C. Ashenfelter and R. Layard, North Holland, vol. 1, chap. 12, 641–692.

Schwerdt, G. and L. Woessmann (2017): "The information value of central school exams," *Economics of Education Review*, 56, 65–79.

SHURE, N. AND L. ZIEROW (2023): "High Achieving First-Generation University Students," *CESifo Working Paper No. 10832*.

SOEP (2023): "Socio–Economic Panel (SOEP), Version 87, Data for Years 1984–2021 (SOEP-Core v38.1, Remote Edition – Update)," Doi:10.5684/soep.core.v38.1r.

STANSBURY, A. AND K. RODRIGUEZ (2024): "The Class Gap in Career Progression: Evidence from US academia," Mimeo.

STIFTERVERBAND (2021): "Vom Arbeiterkind zum Doktor," *Diskussionspapier 2*.

WISWALL, M. AND B. ZAFAR (2018): "Preference for the Workplace, Investment in Human Capital, and Gender," *Quarterly Journal of Economics*, 133, 457–507.

——— (2021): "Human Capital Investments and Expectations About Career and Family," *Journal of Political Economy*, 129, 1361–1424.

A Appendix

A.1 Additional Information on Sample

Table A1: Selection into Survey Participation

	Non-Participants	Participants	Difference
Female	0.54	0.57	-0.03**
High-School GPA	268.83	285.93	-17.11***
Foreign	0.09	0.06	0.03***
Age	19.65	19.51	0.13***
Teacher Training	0.15	0.15	-0.00
Humanities	0.15	0.07	0.08***
Social Sciences	0.07	0.09	-0.03***
Natural Sciences	0.16	0.13	0.03***
Engineering	0.18	0.16	0.02**
Business & Econ.	0.15	0.22	-0.06***
Law	0.08	0.08	-0.00
Medicine	0.06	0.09	-0.03***
Observations	3,854	3,566	

Note: This table shows differences between survey participants and non-participants. All data come from the administrative university data set. High-school GPA is recoded such that higher values indicate better grades. "Foreign" refers to nationality which is given in the administrative data.

A.2 Student Confidence by FiF Status

To gauge to what extent confidence about own ability could differ between FiF and non-FiF students, we computed a confidence measure that considers each student's rank in the distribution of believed ability in their own field as well as the respective "true" ability distribution by field using all students' high-school GPA as a proxy of ability. In the German context, this is a strong measure of student ability, especially in the university's state, where the GPA largely reflects the outcomes of a centralized school exit exam (e.g., Schwerdt and Woessmann, 2017). The difference between these two measures thus gives a proxy for confidence, with positive (negative) values implying overconfidence (underconfidence).

Interestingly, we see that in their own fields, FiF students are significantly more overconfident than non-FiF students on average, as shown in Table A2. This is driven by fields in which FiF students are underrepresented, especially Engineering and Law. While this may at first seem counterintuitive, it is actually well in line with FiF students being less self-confident than non-FiF students. To show this, we also analyzed students' beliefs outside their own field.

We thus computed a confidence measure for all students for the respective fields they are *not* enrolled in. We did so by again first computing students' percentile rank in the "true" overall ability distribution, namely the distribution of high school GPA. Next, we again used the counterfactual ability expectations to compute each student's perceived percentile rank in the ability distribution for all fields but their own. We then again took the difference of the objective rank and the perceived rank for the respective field as a confidence measure, with positive (negative) values implying overconfidence (underconfidence).

Table A3 shows the differences in confidence by FiF status. FiF students are always significantly less confident about their ability in the respective field than non-FiF students.

All in all, FiF students seem to select into fields they feel competitive in and confident about their ability relative to other students.

Table A2: (Over)Confidence by FiF Status in Own Field

	FiF	Non-FiF	Difference
Teacher Training	0.6	-2.0	2.6
Humanities	-1.9	-0.9	-1.0
Engineering	8.5	-4.8	13.3***
Natural Sciences	1.4	-2.2	3.6
Law	5.7	-4.5	10.2**
Business & Economics	-0.6	0.1	-0.7
Medicine	6.3	-0.4	6.7
Social Sciences	1.8	0.1	1.7

Note: This table displays means of student confidence of own ability by own field of study by FiF status. Confidence is defined by the difference between the student's belief about their ability rank, conditional on completion and their high-school GPA rank in the distribution of all students in the field. Asterisks denoting significant differences between the means (measured with a two-sided t-test). * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

Table A3: (Over)Confidence by FiF Status Outside Own Field

	FiF	Non-FiF	Difference
Teacher Training	8.6	14.5	-5.9***
Humanities	2.4	9.2	-6.8***
Engineering	-2.4	1.4	-3.8**
Natural Sciences	-10.7	-7.1	-3.6**
Law	4.3	9.6	-5.2***
Business & Economics	8.4	14.2	-5.8***
Medicine	-31.7	-29.1	-2.6**
Social Sciences	-9.8	-4.1	-5.7***

Note: This table displays means of student confidence of own ability by field of study other than students' own field by FiF status. Confidence is defined by the difference between the student's belief about their ability rank in the specific field, conditional on completion and their high-school GPA rank in the distribution of all students in the same field. Asterisks denoting significant differences between the means (measured with a two-sided t-test). * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

Table A4: Additional Sample Characteristics—Relative Ability in Own Study Field

	FiF	Non-FiF	Difference
Teacher Training	72.99	74.91	1.92
Humanities	62.03	64.76	2.73
Engineering	71.35	69.51	-1.84
Natural Sciences	68.92	70.77	1.85
Law	76.77	73.20	-3.57
Business & Economics	71.62	74.37	2.75
Medicine	66.56	67.42	0.86
Social Sciences	64.25	68.86	4.61*

Note: This table reports summary statistics for various subsamples of our sample of undergraduate students. We display means of relative ability by first-in-family student status, with asterisks denoting significant differences between the means (measured with a two-sided t-test). Higher values indicate a higher (i.e., better) relative rank. In each line, we restricted the sample to students of the respective field, such that the means refer to relative ability of a student of field i in field i. * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

A.3 Comparison to SOEP Data

To compare the FiF gap in expectations to (differences in) realized wages, we look at data from the German Socioeconomic Panel Study (SOEP) (Goebel et al., 2019; SOEP, 2023). We select a sample of SOEP participants for which there is information on their parents' highest degree, general sociodemographic information, and information on earnings at age 30 and 40 in the data. As in our main specification, we define individuals as first-in-family if none of their parents attended college. To mimic our survey sample as closely as possible, we restrict the SOEP sample to people who work full time at age 30 and 40 and to people who went to college themselves. To increase sample size, we select an age window of two years for each points in time, i.e., Panel I (II) of Table A5 uses a sample of individuals aged 29-31 (39-41).[22] Note that because of the panel structure of the data, selecting an age window implies individuals repeatedly occurring in the data. For this reason, we cluster all standard errors at the individual level.

Table A5: FiF Gaps in Gross Monthly Earnings—Comparison to SOEP Data

	(1) Survey	(2) SOEP	(3) Survey	(4) SOEP
I. Outcome: Log (Expected / Actual) Earnings, Age 30				
FiF	-0.0452***	-0.0494***	-0.0329***	-0.0529***
	(0.0103)	(0.0149)	(0.0102)	(0.0148)
Observations	3566	4900	3566	4900
Adjusted R^2	0.017	0.195	0.057	0.219
II. Outcome: Log (Expected / Actual) Earnings, Age 40				
FiF	-0.0604***	-0.0980***	-0.0476***	-0.103***
	(0.0099)	(0.0163)	(0.0097)	(0.0153)
Observations	3566	6251	3566	6251
Adjusted R^2	0.013	0.214	0.069	0.300
Controls	No	No	Yes	Yes

Note: This table reports coefficients of a dummy denoting first-generation student status in regressions of expected (columns (1) and (3)) and realized (columns (2) and (4)) log gross monthly earnings at age 29–31 (39–41) in Panel I (Panel II). Columns (1) and (3) use our survey data on students, whereas columns (2) and (4) use data from the German Socioeconomic Panel (SOEP). Columns (1) and (2) show the unconditional FiF gap in expected and realized earnings (we include fixed effects for year of birth in Column (2)). In columns (3) and (4), we include fixed effects for year of birth (for survey participants) as well as controls for gender and migration background. In columns (2) and (4), the sample is restricted to SOEP participants who work full time and who went to college. In columns (2) and (4) of Panel I (II), we selected an age window of 29-31 (39-41) for our sample. Standard errors (clustered at the individual level) in parentheses. * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

[22] The results are robust to more narrow or wider age windows as well as shifts of the age windows to ages above 30 and 40, respectively.

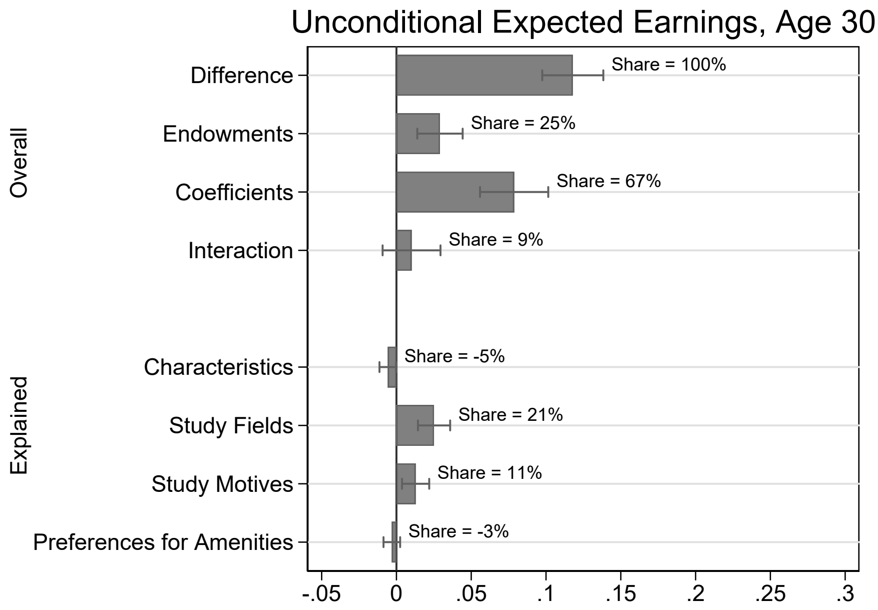
A.4 Comparison to Gender Gap in Earnings Expectations

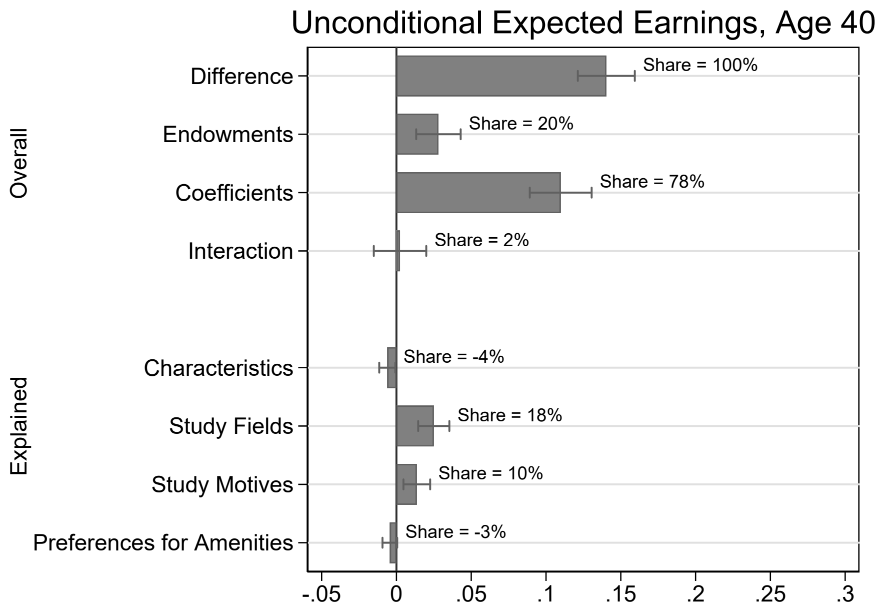
Table A6: Gender Gaps in Unconditional Earnings Expectations

	(1)	(2)	(3)	(4)	(5)
I. Outcome: Log EE, Age 30					
Female	-0.118***	-0.0797***	-0.0872***	-0.0840***	-0.0831***
	(0.0104)	(0.0105)	(0.0105)	(0.0104)	(0.0105)
Adjusted R^2	0.037	0.127	0.143	0.194	0.197
II. Outcome: Log EE, Age 40					
Female	-0.1403***	-0.1039***	-0.1101***	-0.1097***	-0.1105***
	(0.0097)	(0.0096)	(0.0096)	(0.0094)	(0.0094)
Adjusted R^2	0.058	0.195	0.204	0.270	0.275
Observations	3354	3354	3354	3354	3354
Study Field FE	No	Yes	Yes	Yes	Yes
Student Characteristics	No	No	Yes	Yes	Yes
Study Motives	No	No	No	Yes	Yes
Amenities	No	No	No	No	Yes

Note: This table reports coefficients of a dummy denoting being female in regressions of log monthly earnings expectations (EE) at age 30 on study field fixed effects, student characteristics (age, FiF status, migration background, high-school GPA), study motives (have fun during studies, interest in subject, regional attachment, meet people like me, meet people unlike me, safe job, high-paying job, family and leisure compatible job, socially relevant job; 5-point Likert scales), and preferences for non-wage amenities (preferences for leisure, having children, having fun at the job; 11-point Likert scales). Robust standard errors in parentheses. * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

Figure A1: Oaxaca–Blinder–Decomposition of the Gender Difference in Unconditional Earnings Expectations





Note: This figure shows the results of Oaxaca–Blinder decompositions of log unconditional expected earnings at age 30 and 40 in the upper and lower graph, respectively. The upper part in both graphs provides information on the gender difference and the respective shares that can be explained by differences in endowments, coefficients, or the interaction of the two (the displayed percentages always refer to the overall difference). The lower part in both graphs breaks down the share that can be explained by differences in endowments from the upper part into its components and their contributions to that share of the difference. Student characteristics are age, FiF status, migration background, and high-school GPA. Study motives are: have fun during studies, interest in subject, regional attachment, meet people like me, meet people unlike me, have a safe job, have a high-paying job, have a family and leisure compatible job, have a socially relevant job; all asked on 5-point Likert scales. Preferences for non-wage amenities comprise preferences for leisure, preferences for a job that enables one to have children, and preferences for having fun at the job; all asked on 11-point Likert scales.