Gender Gaps in Early Wage Expectations*

Andreas Leibing[†] Frauke Peter[‡] Sevrin Waights[§]
C. Katharina Spiess[¶]

February 6, 2023

Abstract

Using detailed data from a unique survey of high school graduates in Germany, we document a gender gap in expected full-time earnings of more than 15%. We decompose this early gender gap and find that especially differences in coefficients help explain different expectations. In particular, the effects of having time for family as career motive and being first-generation college student are associated with large penalties in female wage expectations exclusively. This is especially true for higher expected career paths. Resulting expected returns to education are associated with college enrollment of women and could thus entrench subsequent gaps in realized earnings.

*We thank Teodora Boneva, Gabriella Conti, Thomas Dohmen, Bernd Fitzenberger, Henning Hermes, Hannah Schildberg-Hörisch, Jonas Jessen, Victor Lavy, Anna Maria Mayda, Markus Nagler, Jürg Schweri, Felix Weinhardt and all participants of the annual meeting of the German Economic Association 2021, the meeting of its standing field committee in Education Economics 2022, the 1st Berlin Workshop on Empirical Public Economics, the 21st journées LAGV 2022, the 2022 meeting of the SEHO and the CIDER-LERN conference 2022 for valuable comments and suggestions. We especially thank Vaishali Zambre for her initial work on wage expectations in the *Berliner-Studienberechtigten-Panel*. Frauke Peter and C. Katharina Spiess gratefully acknowledge funding from the Einstein Foundation Berlin (A-2010-025 (FU)). Andreas Leibing, Frauke Peter, Sevrin Waights and C. Katharina Spiess further acknowledge funding from the German Science Foundation (SP 1091/2-1). Sevrin Waights also acknowledges funding from an SAW project on regional inequality (Leibniz Association grant number: K258/2019). The usual disclaimer applies.

[†]Freie Universität Berlin and DIW Berlin
[‡]DZHW Hannover
[§]Humboldt-Universität zu Berlin and DIW Berlin
[¶]Federal Institute for Population Research (BiB), Johannes Gutenberg University Mainz and IZA

1 Introduction

The literature on gender pay gaps continues to grow rapidly in parallel with a lively public debate (see Blau and Kahn, 2017, for an overview). Many studies show that gender gaps are not only prevalent in actual earnings, but that already in college, women hold lower wage expectations than men (see e.g. Blau and Ferber, 1991, Reuben et al., 2017). However, less is known about the determinants of gender gaps in wage expectations before postsecondary enrollment. Such *early* wage expectations can influence later human capital investment and career decisions. Thus, a better understanding of early gender gaps in expected wages can also provide insights into the formation of later disparities.

Gender gaps in expected wages prior to labor market entry may become self-fulfilling as they can entrench actual wage inequalities through at least two channels. First, lower expected earnings reduce the incentives to invest in higher education (see e.g. Ferber and McMahon, 1979, McMahon and Wagner, 1981). Recent evidence shows that expected wages are a significant predictor for the choice which level of education to pursue, as well as for college major choice.[1] Second, wage expectations might affect starting wages through the formation of reservation wages (Brown and Taylor, 2013). If, in turn, lower reservation wages result in lower starting salaries, they are likely to have a persistent effect on actual wage trajectories.[2] Caliendo et al. (2017) show that gender differences in reservation wages can indeed account for a large share of the subsequent gender gap in realized wages. This is exacerbated by gender differences in negotiation styles and outcomes, which are larger for inexperienced negotiators (see Mazei et al., 2015, for an overview) and thus matter especially for starting wages.

Our study asks: what factors drive the gender gap in expected wages among high school students? Our analysis draws on data from a unique survey among high school graduates in Germany, in which we asked individual students for their expected full-time earnings range at age 35 years in three different scenarios: (i) if they enrolled for a vocational degree, (ii) if they enrolled for a bachelor's degree or (iii) if they enrolled for a master's degree. We examine the determinants of the expected gender gap using a regression-compatible Oaxaca-Blinder (OB) decomposition (Oaxaca, 1973, Blinder, 1973, Fortin, 2008) and place a particular focus on differences in coefficients (*unexplained part*) as opposed to differences in endowments (*explained part*), giving a detailed overview of components attributable to socio-demographic factors, intended college major choice, career motives and both cognitive and noncognitive abilities.

[1] See e.g. Arcidiacono et al. (2012), Attanasio and Kaufmann (2014, 2017), Belfield et al. (2020), Favara et al. (2021) for high school and college enrollment and Boneva et al. (2021) for postgraduate enrollment. For college major choice see Zafar (2013), Ruder and Noy (2017), Wiswall and Zafar (2018) and Arcidiacono et al. (2020).

[2] Evidence on the adverse effect of lower starting wages is provided by Oreopoulos et al. (2012). The authors show that entering the labor market during a recession has potentially long lasting *scarring effects* on wages.

We find that the gender gap in average wage expectations after high school graduation amounts to over 15%. Our decomposition results indicate that endowments explain some of the difference in future wages: around a quarter of the total gap. In particular, intentions to choose a major in Science, Technology, Engineering, and Mathematics (STEM) and a high academic self-efficacy[3] are more prevalent in men and on average associated with higher expected wages. However, it is the differences in coefficients rather than differences in endowments that play a bigger role. Being potential first generation at college, intending to study business or management and expressing *time for family* as a career motive all have significantly more negative impacts on female expected wages than on male expected wages. Some further factors, such as having *good health and safety conditions* as a career motive, also have relatively more positive impacts on female expectations. The fact that these coefficient effects offset each other, explains the relatively low share of the gender gap in wage expectations explained by differences in coefficients.

We investigate the decomposition not just for mean expected wages pooled over educational scenarios, but also for minimum and maximum expected wages, as well as bachelor and master scenarios separately, to shed light on underlying heterogeneity. Certain factors, e.g. *time for family*, are associated with the largest contribution via the unexplained part for the maximum expected wages and the master scenario, implying the differential impacts are exacerbated for the highest career tracks. This result suggests that women anticipate having to give up higher career paths and leadership positions in order to have more flexible work arrangements.[4] Recent evidence by Wiswall and Zafar (2021) stresses the close link between human capital investments, including major choice, and expectations about career and family.

In our findings, having a very high preference for *time for family* as a career motive is the largest single and most consistent driver of the gender gap in wage expectations. Although time for family could cover many family related issues, as caring for parents, partners, or siblings in need, it is very often related to childcare. Our analysis suggests that *family penalties*, one of the most central factors in explaining inequalities in earnings, are already reflected in the expectations of high school students. Finally, to assess possible pathways of how the gender gap in wage expectations might translate into actual earnings, we examine the association between expected returns to college and college enrollment. We find strong heterogeneities across gender and socioeconomic background.

The majority of existing studies on earnings expectations focus on averages, asking students for point estimates of their future wages (for overviews, see e.g. Brunello et al., 2004, Manski, 2004, Giustinelli, 2022). By eliciting the minimum and maximum expected

[3] Academic self-efficacy refers to individuals convictions that they can successfully perform given academic tasks at designated levels (Bong and Skaalvik, 2003). See Section 2 for details.

[4] The phenomenon of such *glass ceilings* is well reflected in the relatively low share of women in leading positions in firms (see e.g. Bertrand and Hallock (2001) for the U.S. and Kirsch and Wrohlich (2020) for Germany) but also in academia (Ceci and Williams, 2011).

wages of high school students, we are able to give a more comprehensive picture and are able to assess drivers of wage expectations at different margins. Many studies in this strand of the literature compare expected wages of students to actual wages of different reference groups and focus predominantly on college students (e.g. Manski, 1993, Betts, 1996, Wolter, 2000, Carvajal et al., 2000, Wolter and Zbinden, 2002, Huntington-Klein, 2015, Alonso-Borrego and Romero-Medina, 2016). By eliciting wage expectations for different degree types, we can compare wage expectations to actual earnings of a wider range of reference groups. Other studies investigate the *accuracy* of college students' expectations and compare them to later realized wages by following students into the labor market (e.g. Webbink and Hartog, 2004, Filippin and Ichino, 2005, Jerrim, 2011, 2015).

Our study is not the first to elicit the wage expectations of pre-college students. However, it is one of only a few studies that aims to explain the gender gap in expectations at this early age. Many studies of the wage expectations of pre-college students do not specifically estimate gender gaps (e.g. Dominitz and Manski, 1996, Wolter, 2000, Attanasio and Kaufmann, 2014, Hastings et al., 2015, 2016, Schweri and Hartog, 2017) whereas several others estimate the gender gap only in passing while summarising the data (e.g. Mazza and Hartog, 2011, Boneva and Rauh, 2020, Belfield et al., 2020). Only three previous studies place a specific focus on estimating and explaining the gender gap in wage expectations, Eliophotou-Menon (1997a,b) for secondary school students in Cyprus, Attanasio and Kaufmann (2017) for high school graduates in Mexico, and Boneva et al. (2022) for secondary school students Germany. All three studies regress explanatory factors on wage expectations for the girls and boys groups separately, but none carries out a decomposition analysis to assess the determinants of gender gaps in early wage expectations.

Eliophotou-Menon (1997a,b) elicits wage expectations of secondary school students in Cyprus finding a gap in wage expectations and finding that factors such as family background and perceived ability affected the wage expectations of girls and boys differently. Attanasio and Kaufmann (2017) document a gap in wage expectations for high school graduates in Mexico also findings differences between boys and girls in the relationship between expectations and personal characteristics. Their study also analyzes later college enrollment, but places a focus on marriage market returns. Boneva et al. (2022) elicit the gender gap in competitiveness of adolescents from lower socio-economic backgrounds in Germany by looking at earnings expectations. Their focus is on the development of gender differences in competitiveness and the role of the social environment in this process. Additionally, the authors show that also earnings expectations of girls and boys already differ significantly at around fourteen years of age. Their findings suggest that the gender gap in wage expectations emerges already at the beginning of adolescence and is larger for children from lower socio-economic backgrounds.

A small number of studies decomposes the gender gap in wage expectations, but only for college students, i.e. when the decision to invest in higher education is already made.[5] Delaney et al. (2010) decompose the gender gap in expected earnings of college students, using data from the Irish Universities Study. As one of only a few studies, they also consider career motives and noncognitive abilities as determinants of wage expectations. Their findings attribute a large share of the gender gap to risk preferences, but also indicate that family considerations can play a significant role. Kiessling et al. (2019) use data from the online survey *Fachkraft 2030* that elicits wage expectations of college students in Germany.[6] Their findings attribute a larger role in explaining the gender gap in expected wages to occupational sorting and negotiation styles. Surprisingly, prospective child-related labor force interruptions have a relatively small effect in their framework, which focuses primarily on endowment effects. Fernandes et al. (2021) use data on wage expectations of students majoring in business at two Swiss colleges. Their findings confirm the importance of both fertility considerations and career preferences in explaining the gender gap in wage expectations, but focus mostly on endowment effects. Consistent with our results, they also find that both males and females overestimate their wages compared to actual ones. Furthermore, males respond in an overconfident manner to information about realized wages. Briel et al. (2022) decompose the gender gap in wage expectations for German university applicants. With an average age of 21 year these prospective students are older than those in our study and have already decided whether to invest in higher education. The authors carry out quantile regressions finding larger gaps at the bottom of the distribution. They find an important role of *biased beliefs*: men tend to overestimate both the average salary in their fields as well as their own likelihood to earn above this average. In a decomposition, these factors all help explain the gender gaps in expectations, leaving most other factors seemingly irrelevant.

Lastly, related studies by e.g. Wiswall and Zafar (2018) and Reuben et al. (2017) investigate the relationship between career motives, major choice and wage expectations among college students, using data from high-ability undergraduate students at New York University. Wiswall and Zafar (2018) confronts these students with multiple hypothetical job choice scenarios that vary in expected wages and other job characteristics. Their results show that, among others, women have a significantly higher willingness to pay (in terms of expected wages) for flexible working hours and more secure jobs than men. Since *time for family* can be interpreted as a form of higher flexibility in working hours, our findings are roughly in line with these results. However, we cannot confirm women's

[5] A further related literature decomposes the gender gap for realized earnings in Germany (see Francesconi and Parey (2018) for college graduates and Collischon (2019) for the working population).

[6] The data was collected among participants of largest job board *jobmensa.de* in Germany. Ehrmantraut et al. (2020) use the same data to study the expected signalling value of completing higher education, using college dropout as counterfactual.

high willingness to pay for secure jobs.[7] Reuben et al. (2017) also investigates the role of preferences in explaining gender differences in wage expectations. Based on the same survey among New York University undergraduates they document a large gender gap in expected wages. While part of the gap is due to gender differences in college major choice, the gap in wage expectations within a college major still amounts to around 20%. They further show that gender differences in preferences such as overconfidence, competitiveness and risk aversion, explain 18% of the gender gap in expectations.

Our study stands out from the related literature on wage expectations as it combines central aspects of the before-mentioned studies. To the best of our knowledge, this paper is the first to analyze determinants of the gender gap in earnings expectations of individuals before their decision to invest in higher education by explicitly carrying out detailed decompositions at different margins. Moreover, our panel data allows to track students over time and assess the association of wage expectations with college enrollment. Lastly, and unlike most other studies, we consider different noncognitive abilities and career motives, such as *time for family*, as potential drives of the gender gap. Our study therefore contributes to a better understanding of gender gaps in wage expectations and the interplay between educational investments and family and career considerations that take place already at a relatively early and crucial age.

The paper is organized as follows. Section 2 introduces the data, provides detailed information on the measurement of wage expectations and presents descriptive statistics. Section 3 introduces the OB decomposition method. Section 4 presents the decomposition results, examining the role of different factors in explaining the gender gap in wage expectations. Section 5 shows associations of wage expectations with subsequent college enrollment. Section 6 concludes.

2 Data and descriptive statistics

The data that we use in our empirical analysis are based on a survey of high school students at different institutions in the German capital city of Berlin. They include detailed information on student characteristics, educational aspirations and eventual educational choices. All of these aspects were surveyed as part of a larger research project, the ***Be****rliner-****Stu****dienberechtigten-****P****anel* (Best Up), which aims to study educational paths of high school students (see Ehlert et al., 2017b, for details).

The survey aimed at obtaining a sample of students who were predominantly from lower socioeconomic backgrounds. Thus, the 27 selected schools are located in districts with a high share of individuals without college degree and cover 20% of all upper-secondary schools in Berlin. The survey followed all students from the end of their penultimate year in high school through two years after graduating from high school. All

[7] A likely explanation is the fact that the German labor market is much more regulated than the American labor market. For example, *at-will employment* is not possible in Germany.

students were surveyed five times over that period. Except for the first paper and pencil survey conducted in schools, the subsequent surveys were administered online.[8] Of the 1,578 students surveyed in the first wave, 1,105 participated in the second and 1,033 in the third wave, where wage expectations were surveyed (see Appendix Figure F.1).

Wage expectations in this context have first been studied by Zambre (2018).Other elements of the Best Up project consisted in separate randomized controlled trials. First: an information intervention, in the form of a workshop, on the returns to tertiary education to study effects on college enrollment intentions (Peter and Zambre, 2017), college applications (Ehlert et al., 2017a) and actual enrollment (Peter et al., 2021). Second: a financial intervention in form of a temporary monthly subsidy for students without enrollment intentions to study the effect on college applications, which remained ineffective (Peter et al., 2017). Even though these interventions are not the focus of our analysis, they might still affect wage expectations differently for both genders and thus also the gender gap in expectations (see Appendix Figure F.2 for the information treatment). Therefore, we include indicator variables accounting for school-level assignment to each intervention throughout this study. Those participants who decided to enroll at college were surveyed in subsequent waves to study the transition into postgraduate education. This so called PostGrad-Best Up project thus focused only on students, who were enrolled in college in 2017 (see Berkes et al. (2022) and Appendix Figure F.1).

2.1 Wage expectations

In the expected earnings module of the survey in summer 2014, students were asked to state the minimum and maximum net wage that they might expect to earn at the age of 35 conditional on working full time.[9] Following Guiso et al. (2002) and Attanasio and Kaufmann (2014, 2017), students were then asked what they think is the probability they will earn more than the midpoint of the range between their stated minimum and maximum and we assume a single triangular distribution between the minimum and maximum expected wage, scaling each half of the triangle such the area of the right half matches the reported probability p to earn above the midpoint (see Appendix Figure B.1). As such the average expected wage is then computed as $E(y) = \frac{1-p}{3}(2 \cdot y_{min} + y_{max}) +$

[8] In the first wave, students were invited to fill out a paper based questionnaire at school during their classes and were possibly more motivated to participate in the survey. This could explain the large attrition rates in comparison to later waves, where the survey was administered online and there were e.g. less peer-effects at work.

[9] Given that we condition on full time employment, biases arising from different labor supply expectations are ruled out by construction. The original question suggested a regular earned income arising from an employment agreement and we therefore use the terms *wage* and *earnings* interchangeably. It was formulated as follows: *"Now assume that you earned a vocational or university degree and work full-time, meaning that you are economically active for about 39 hours per week. What do you expect your monthly minimum net earnings and monthly maximum net earnings (i.e. the minimum/maximum amount of money that is transferred directly on your bank account) at the age of 35 to be if you ... [have earned a vocational/Bachelor/Master degree]".*

$\frac{p}{3}(y_{min} + 2 \cdot y_{max})$.

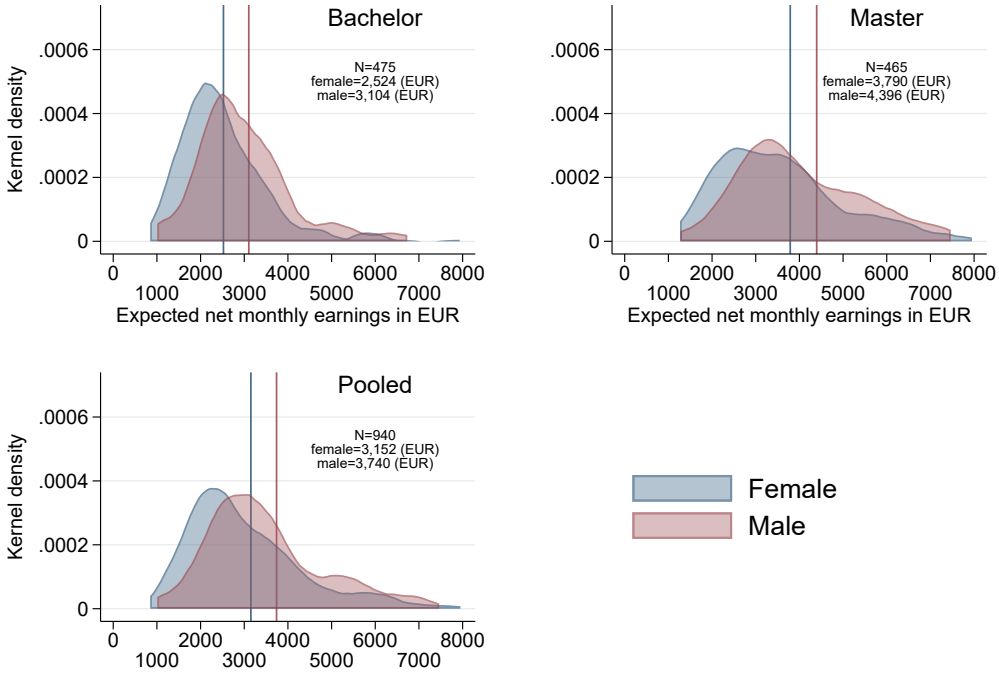
We asked each individual for three pieces of information for three different hypothetical educational scenarios, in which they have earned: i) a vocational degree, ii) a bachelor's degree, or iii) a master's degree. Respondents could either state their expectations or leave open the question. In our main analysis, we pool individual expectations on bachelor's and master's degrees and discard information on wage expectations with a vocational degree.[10] Hence, in our analysis we can only consider individuals who gave full information (i.e. *min*, *max* and p) for at least one degree (bachelor's or master's). Of the 1033 individuals who participated in wave three, 376 (36.4%) did not give any information on either their minimum or maximum wage expectations on any of the two degrees. Another 36 individuals (3.5%) gave only either their minimum or maximum wage expectations for both degrees. We also exclude students whose responses were not logically consistent: 13 individuals (1.3%) stated equal minimum and maximum expectations and another 8 individuals (1%) gave no, or invalid, information (i.e. $p = 1$ or $p = 0$) on their expected probability to earn above the midpoint. We further exclude 9 students (1%) with average wage expectations either above the highest or below the lowest percentile of the cross-sectional distribution of expected wages for each education degree. Further, we drop 78 students (8%) who did not provide complete information on all covariates and could thus not be used in the decomposition. The final sample for the decomposition analysis consists of 513 students, of whom 205 are male and 308 are female.[11]

Individual average expected wages, pooled over degrees are presented in Figure 1. In line with previous findings in the literature we find a large gender gap. Even before entering the labor market women expect to earn considerably less than their male counterparts. While men expect to earn on average around 3,192 EUR per month, women expect to earn 2,691 EUR. This difference implies a gender gap in wage expectations of 15.7%. Several other observations are noteworthy: First, average expected wages increase with the level of the education scenario, indicating that students are aware of the monetary returns to higher levels of education. Second, the higher the educational degree, the more dispersed the distribution. This shows that expected wage risk is increasing for higher levels of education, which is in line with the findings on actual labor market data (Koerselman and Uusitalo, 2014). In Appendix Section C we also present descriptive statistics by gender for the minimum and maximum expected wages. Differences in the average are driven by both male students reporting higher maximums and female students reporting lower minimums, although this depends somewhat on the hypothet-

[10] We use expectations conditional on vocational degrees as reference when computing the expected returns to a college degree that we use to analyze college enrollment in Section 5.

[11] We explore item non-response for the wage expectation question and the full set of covariates in Appendix Section A. While non-response is related to certain characteristics such as academic performance, the selection is similar between men and women, which is central to our study.

Figure 1: Expected wages by education scenario and gender.



Notes: This figure shows the cross-sectional distribution of expected average wages with different educational degrees for men (red) and women (blue). Observations above the highest and below the lowest percentile of the respective distributions are excluded. For illustration purposes wage expectations exceeding 8,000 EUR per month are not depicted. *Source: Berliner-Studienberechtigten-Panel* wave 3.

ical degree. Appendix Table D.3 shows detailed decomposition results for the implied individual range in wage expectations as a measure for wage risk.

Third, both male and female wage distributions are right-skewed, just as is actual earnings distributions. Appendix Table C.2 further compares average wage expectations to actual earnings by gender. Interestingly, both genders overestimate the returns to higher degrees, but females overestimate more often and to a higher degree than men. Finally, in all education scenarios the distribution of men is shifted to the right and exhibits a thicker right tail, implying that men expect higher wages than women on average and are more likely to expect exceptionally high wages. A Kolmogorov-Smirnov test confirms that the distributions of cross-sectional average expected wages differ significantly by gender in each education scenario. More detailed descriptive statistics can be found in Appendix Table C.1.

2.2 Sample characteristics

The data include detailed information on characteristics potentially related to wage expectations. We outline those variables below and report means by gender in Table 1.

(I) **Baseline characteristics** comprise socio-demographic factors such as migration background, parental educational background, and attended high school type (academic

high school, integrated high school, and vocational high school). Men and women do not differ significantly in these characteristics, with the exception of attended high school type. We also report whether students attended an *information intervention school* or a *financial intervention school* of the Best Up project.

(II) **Cognitive abilities** comprise students' final high school GPA as well as test scores on a verbal and figural cognition test.[12] We would expect higher performing students to anticipate that their higher ability (or signal thereof) is rewarded in the labor market. Surprisingly, women score lower on the verbal cognition test by around one point. Following Fortin (2008), we use standardized values for all cognitive abilities in the decomposition.

(III) **Intended college major** accounts for the well-documented wage differences between college majors as well as differences in college major choice by gender. The intended college major should proxy the type of occupation and/or industry individuals aspire to work in and should thus reflect differences in wage expectations due to sorting (Montmarquette et al., 2002, Arcidiacono et al., 2020). Based on the classification of the German Statistical Office (Destatis, 2012), the different majors are grouped into ten fields of study, as listed in Table 1.[13] Women are significantly more likely to express intentions to enroll in medical studies and teaching and are less likely to report intentions for a major in STEM.

(IV) **Career motives** capture the importance of different job attributes that students assign to their future job choice. Similar to the intended college major, wage expectations are likely to be affected by the career plans that students hold, which in turn are likely to differ by gender (Daymont and Andrisani, 1984). In particular, one might expect that women anticipate future career breaks, e.g. due to care duties related to children or other relatives, which may explain their lower wage expectations compared to men (Chevalier, 2007). Although we do not have direct information on e.g. child bearing plans, the survey includes eleven items that capture the importance of different career aspects for students' future job choice, on a four-point Likert scale that ranges from one *"not important at all"* to four *"very important"*. This includes the preference to have a job that leaves sufficient *time for family* commitments. In the decomposition, following Fortin (2008), we use binary indicators that signals if an individual considers a factor very important.

Career motives differ significantly between genders. Overall, women have higher pref-

[12] Note that final high school GPA traditionally ranges from one (best) to four (worst) in Germany. However, we reverse this score such that a higher GPA corresponds to higher performance. The cognition tests were conducted in the first wave 1 of the *Berliner-Studienberechtigten-Panel*. Higher scores on the cognitive tests indicate higher skills as well.

[13] The information on students intended major is derived from different waves of the survey. Firstly, if students already applied to university or reported to plan on applying in the third wave, we have information on which majors they applied for. If students apply for more than one major, we use the major that students rank as their first choice. Secondly, students who reported during high school that they intend to enroll in university, were also asked about the major that they would like to enroll in.

Table 1: Sample characteristics.

	Male	Female		Male	Female
Baseline characteristics:			**Career motives:**		
Information intervention school	0.327	0.357	High income	3.229	3.175
Financial intervention school	0.405	0.344	Promotion possibilities	3.176	3.162
Migration background	0.420	0.477	Recognition	2.795	2.942*
First generation at college	0.595	0.597	Interesting job	3.498	3.630**
Academic high school	0.254	0.331*	Independent working	2.946	3.062
Integrated high school	0.405	0.341	Social interaction	2.688	3.055***
Vocational high school	0.341	0.328	Important for society	2.615	2.747*
Fast track option	0.034	0.052	Help others	2.561	2.981***
			Spare time	2.756	2.805
Cognitive abilities:			Health/safety conditions	3.239	3.545***
Final high school GPA	2.508	2.567	Time for family	3.093	3.351***
Verbal cognitive skills	10.917	9.808***			
Figural cognitive skills	11.024	11.244	**Noncognitive abilities:**		
			Extraversion	4.849	4.880
Intended college major:			Openness	4.970	5.136*
Arts & Humanities	0.063	0.091	Conscientiousness	4.698	5.028***
Social Sciences & Economics	0.024	0.036	Neuroticism	3.611	4.633***
Business & Management	0.093	0.130	Agreeableness	5.029	5.389***
STEM	0.332	0.153***	Locus of control (int.)	5.027	4.879**
Teaching	0.054	0.107**	Academic self-efficacy	3.220	3.062***
Law	0.039	0.032	Self-confidence	5.200	4.727***
Health & Medicine	0.107	0.192**			
Other	0.063	0.032*			
No college aspiration	0.220	0.195			
Missing information	0.005	0.032**			
N	205	308		205	308

Notes: This table shows means of individual characteristics for men and women. Significance stars signal mean differences based on a two-sided t-test. *Source: Berliner-Studienberechtigten-Panel* waves 1-2. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

erences for all intrinsically, socially, and work-life-balance motivated factors. Although there are no significant differences for extrinsically motivated (or monetary) factors, such as preferences for a high income and good promotion possibilities, it is unclear what men and women assume to be good promotion possibilities and a high income. Gender differences in wage expectations suggest that men and women indeed have a fundamentally different understanding of the latter.

(V) **Noncognitive abilities** cover personality traits, locus of control and confidence measures. A large strand of literature emphasizes the importance of such attributes in explaining educational choices and labor market outcomes (e.g. Heckman et al., 2006) and document gender differences with respect to these non-cognitive skills (see Bertrand, 2011, for an overview).[14] In Section 3 we investigate how these differences are related to gender differences in expectations.

For personality traits, we use an adjusted version of the Five Factor Model that covers openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (Big Five) (McCrae and Costa, 1996). Each dimension is represented by three statements that are answered on a seven-point Likert scale ranging from one *"does not apply at all"* to seven *"fully applies"* (Dehne and Schupp, 2007). Unlike when measuring adults, measuring youth' openness to experience is based on four questions as defined in Weinhardt and Schupp (2011). Based on this information, we generate summation scores for each personality dimension. Comparing personality traits, women are more open, more conscientious, more neurotic, and more agreeable than men. The extent of extraversion is the only dimension that does not differ significantly between gender.

The (internal) locus of control indicates how strongly an individual believes that what happens is a consequence of her own actions as opposed to external factors, e.g. luck or fate (Rotter, 1966).[15] Women show a lower internal locus of control, indicating that they perceive their life to be more affected by circumstances outside their control than men.

Self-confidence accounts for gender differences in the assurance to succeed in the labor market that could result in higher expected wages. This (general) self-confidence is approximated by the extent students agree with the following statement: *"I am a person who has a positive attitude toward herself."* The extent of agreement is measured on a seven-point Likert scale. Additionally, students were asked how likely they think it is that they could successfully graduate from university, indicating their academic self-efficacy. Answers are given on a four-point Likert scale ranging from *very low* to *very high*. Table 1 shows that women have a considerably lower general self-confidence and lower academic

[14] For example, Grove et al. (2011) find that the inclusion of measures on non-cognitive skills and work preferences significantly increases the explained part of the gender pay gap for a sample of individuals with a master's degree in the U.S.

[15] This measure is based on eight different items, capturing the extent to which individuals agree (on a seven-point Likert scale) with statements such as *"The possibilities we have in life are dependent on social circumstances."*

self-efficacy when compared to men. Following (Fortin, 2008), we use standardized values for all noncognitive abilities in the decomposition.

3 Methodology

In order to investigate the determinants of the gender gap in wage expectations, we rely on the regression-compatible adaption of the standard Oaxaca-Blinder (OB) decomposition (Oaxaca, 1973, Blinder, 1973) as proposed by Fortin (2008).[16] We start by regressing expected log wages y of each individual i associated with each degree d separately for females f (males m) in a simple multivariate setting:

$$ln(y_{id}^{j}) = \beta_{0f} + \boldsymbol{X_i'\beta_f} + \varepsilon_{id}, \tag{1}$$

$$ln(y_{id}^{j}) = \beta_{0m} + \boldsymbol{X_i'\beta_m} + \varepsilon_{id}, \tag{2}$$

which gives estimates of the expected wage structure $\boldsymbol{\beta}$ of each group (male and female) multiplied with the set of corresponding control variables in $\boldsymbol{X}$. These comprise baseline characteristics, intended college major choice, career motives, as well as cognitive and noncognitive abilities as described in Section 2. Superscript j indicates the different outcomes y_{min}, y_{max} and $E(y)$. In a second step, we estimate the corresponding reference wage structure $\boldsymbol{\gamma}$ in a similar regression, pooled over males and females:

$$\begin{aligned} ln(y_{id}^{j}) = \gamma_0 + \gamma_{0f} \cdot F_i + \gamma_{0m} \cdot M_i + \boldsymbol{X_i'\gamma} + \boldsymbol{\nu_{id}}, \\ \text{subject to } \gamma_{0f} + \gamma_{0m} = 0. \end{aligned} \tag{3}$$

To account for group-membership effects, the pooled regression includes separate gender intercepts for males and females. This mitigates the problem that pooled coefficients overstate the effects of variables with large gender differences (Fortin, 2008). In each specification we account for clustering at the individual level.

After obtaining male, female and reference wage structure, gender gaps in expected wages can be decomposed into:

$$\overline{ln(y_m^j)} - \overline{ln(y_f^j)} = \triangle \boldsymbol{X}\hat{\boldsymbol{\gamma}} + [\bar{\boldsymbol{X}}_{\boldsymbol{m}}'(\hat{\boldsymbol{\beta}}_{\boldsymbol{m}} - \hat{\boldsymbol{\gamma}}) + (\hat{\beta}_{0m} - \hat{\gamma}_0)] - [\bar{\boldsymbol{X}}_{\boldsymbol{f}}'(\hat{\boldsymbol{\beta}}_{\boldsymbol{f}} - \hat{\boldsymbol{\gamma}}) + (\hat{\beta}_{0f} - \hat{\gamma}_0)], \tag{4}$$

which is a reformulation of the traditional triple-difference OB decomposition (Cotton, 1988, Neumark, 1988, Oaxaca and Ransom, 1994). In this framework, gender gaps can be decomposed into an explained part ($\triangle \boldsymbol{X}\hat{\boldsymbol{\gamma}}$) that captures *differences in characteristics* and an unexplained part that captures *differences in coefficients*. The unexplained part

[16] For a detailed discussion of different decomposition methods see Fortin et al. (2011).

itself can be decomposed into a part that reflects the male advantage ($\bar{\boldsymbol{X}}'_{\boldsymbol{m}}(\hat{\boldsymbol{\beta}}_{\boldsymbol{m}} - \hat{\boldsymbol{\gamma}}) + (\hat{\beta}_{0m} - \hat{\gamma}_0)$) and a part that reflects the female disadvantage ($\bar{\boldsymbol{X}}'_{\boldsymbol{f}}(\hat{\boldsymbol{\beta}}_{\boldsymbol{f}} - \hat{\boldsymbol{\gamma}}) + (\hat{\beta}_{0f} - \hat{\gamma}_0)$).

The explained part gives the share of the overall gender gap that can be explained by women having characteristics that are associated with higher expected wages on average (in both groups). The unexplained part, in contrast, gives the share of the remaining gap that can be explained by women having a different smaller association between certain characteristics and higher expected wages. Traditionally, the explained part reflects differences in *endowments*, i.e. differences in characteristics that capture an individuals productivity. The unexplained part could then be interpreted as discrimination, under the assumption that covariates fully capture productivity differences. In a decomposition of expected wages, the unexplained part reflects subjective beliefs on how different career choices or motives will affect later earnings. These subjective beliefs may themselves be subject to social norms, for example regarding care-giving roles during parenthood, or different role models of young men and women. If women implicitly expect having to give up higher career paths or earnings in order to undertake family duties, then this could also be interpreted as implicitly expected discrimination. Therefore, as both the explained and the unexplained part of the gap provide important information as to whether differential characteristics or differential subjective beliefs are the source of gender gaps in expectations, both have their own policy implications.

4 Decomposition Results

4.1 Average wage expectations

Table 2 presents results of the OB decomposition for average expected wages pooled over master's and bachelor's degree. Column (1) shows the contribution of each covariate via the explained part and Column (2) gives the respective contribution via the unexplained part. Coefficients behind a covariate category (in bold) give the joint contribution of a set of factors, i.e. the sum of all individual coefficients in that category. The gender gap in expected wages (0.173 log points) equals the sum of the the explained and unexplained part plus the regression constant, which signals the share attributable to unobserved factors. The share attributable to differences in endowments (*explained part*) is positive and significant with 0.045 log points - about a quarter of the overall gap. The share attributable to coefficients (*unexplained part*) is smaller and sums up to 0.016 log points. However, coefficients of the unexplained part are considerably larger on average, but offset each other. Together, the differences in endowments and coefficients of observable characteristics account for about 60% of the overall gap. The remaining constant (0.111 log points) collects all unobserved factors that affect wage expectations.

The two most important categories of endowment (explained part) are intended college major and noncognitive abilities. Differences in intended college major explain about

Table 2: Detailed decomposition for average expected earnings (Pooled).

	(1) Explained		(2) Unexplained	
Baseline characteristics:	0.005	(0.008)	0.084	(0.082)
Information intervention school	-0.000	(0.001)	0.039	(0.032)
Financial intervention school	0.003	(0.003)	0.021	(0.033)
Migration background	0.001	(0.002)	-0.025	(0.034)
First generation at college	-0.000	(0.001)	0.069*	(0.039)
Academic high school	-0.001	(0.002)	-0.002	(0.015)
Integrated high school	-0.001	(0.002)	-0.008	(0.019)
Vocational high school	0.000	(0.000)	0.011	(0.016)
Fast track to vocational degree	0.005	(0.006)	0.001	(0.006)
Master	-0.001	(0.002)	-0.023	(0.014)
Cognitive abilities:	0.004	(0.010)	-0.003	(0.007)
Final high school GPA	0.000	(0.003)	-0.000	(0.003)
Verbal cognitive skills	0.002	(0.009)	-0.003	(0.005)
Figural cognitive skills	0.002	(0.003)	0.000	(0.001)
Intended college major:	0.017	(0.011)	-0.008	(0.029)
Arts & Humanities	-0.000	(0.001)	-0.016*	(0.008)
Social Sciences & Economics	0.001	(0.004)	-0.000	(0.004)
Business & Management	-0.001	(0.002)	0.031***	(0.012)
STEM	0.019**	(0.008)	-0.013	(0.018)
Teaching	-0.005	(0.004)	-0.001	(0.008)
Law	0.001	(0.002)	-0.012*	(0.007)
Health & Medicine	-0.001	(0.004)	0.008	(0.013)
Other	-0.000	(0.003)	0.003	(0.008)
No college aspiration	0.000	(0.002)	-0.011	(0.018)
Missing information	0.003	(0.002)	0.002	(0.002)
Career motives:	0.004	(0.015)	-0.060	(0.068)
High income	0.002	(0.003)	-0.036	(0.029)
Promotion possibilities	-0.001	(0.003)	-0.005	(0.032)
Recognition	-0.004	(0.004)	0.010	(0.021)
Interesting job	0.004	(0.004)	0.055	(0.046)
Job security	-0.005	(0.004)	-0.012	(0.037)
Independent working	-0.000	(0.004)	0.006	(0.022)
Social interaction	-0.004	(0.007)	-0.034	(0.022)
Important for society	-0.000	(0.001)	-0.013	(0.020)
Help Others	0.004	(0.008)	-0.018	(0.026)
Spare time	0.001	(0.002)	-0.001	(0.015)
Health/safety conditions	0.002	(0.006)	-0.081**	(0.039)
Time for family	0.006	(0.005)	0.068**	(0.030)
Noncognitive abilities:	0.015	(0.009)	0.004	(0.008)
Openness	0.000	(0.002)	0.000	(0.001)
Extraversion	0.000	(0.002)	0.000	(0.002)
Conscientiousness	-0.000	(0.003)	-0.000	(0.001)
Neuroticism	-0.000	(0.000)	0.001	(0.006)
Agreeableness	-0.000	(0.000)	0.000	(0.001)
Locus of control (int.)	0.003	(0.004)	0.003	(0.005)
Academic self-efficacy	0.012*	(0.007)	0.001	(0.003)
Self-confidence	0.000	(0.005)	-0.000	(0.003)
Subtotal	0.045*	(0.024)	0.016	(0.106)
Constant			0.111	(0.111)
Total gap	0.173***	(0.035)		
N	940			

Notes: This table presents estimates of a detailed Oaxaca-Blinder decomposition using pooled coefficients as weighting scheme. The outcome variable is average expected wages of high school graduates pooled for education scenarios with a bachelor's and master's degree. Joint contribution of factors in each category given by coefficients behind the categories name (in bold). Standard errors allow for clustering at the individual level and presented in parentheses. Source: *Berliner-Studienberechtigten-Panel* waves 1-3. $p<0.1$, ** $p<0.05$, *** $p<0.01$.

0.017 log points (around 10% of the overall gender gap in expected wages) and operate mainly through STEM enrollment.[17] The relatively small share explained by other fields of study suggests that much of the gender difference in expected average wages may occur within (intended) college majors, in line with findings by Reuben et al. (2017). An on average higher academic self-efficacy of men explains 0.012 log points, i.e. about 7% of the total gap. This coefficient reflects a higher confidence of male high school graduates in their ability to successfully graduate from college. General self-confidence plays virtually no role in explaining gender differences in expected earnings.[18] The remaining coefficients are insignificant.

In terms of contributions via the unexplained part, several factors are statically significant. Being a first generation potential college students contributes 0.069 log points to the gender gap in expected wages. Figure 2 shows that this factor has a marginally significant average effect on women, but not on men. This difference alone can explain almost 40% of the overall gap. Of similar sizes in magnitude (although in opposite directions) are the career motives *good health and safety conditions* (-0.081 log points) and *time for family* (0.068 log points). Finally, several intended college majors have differential expected earnings associations for men and women. Most notably, Business & Management is associated with 0.031 log points lower earnings for women than for men. Arts & Humanities and Law, however, are associated with somewhat higher earnings by women compared with men. The negative contribution of the covariate *Master* via the unexplained part shows, that the gender gap in wage expectations for a master's degree is 0.023 log points lower than the average.[19]

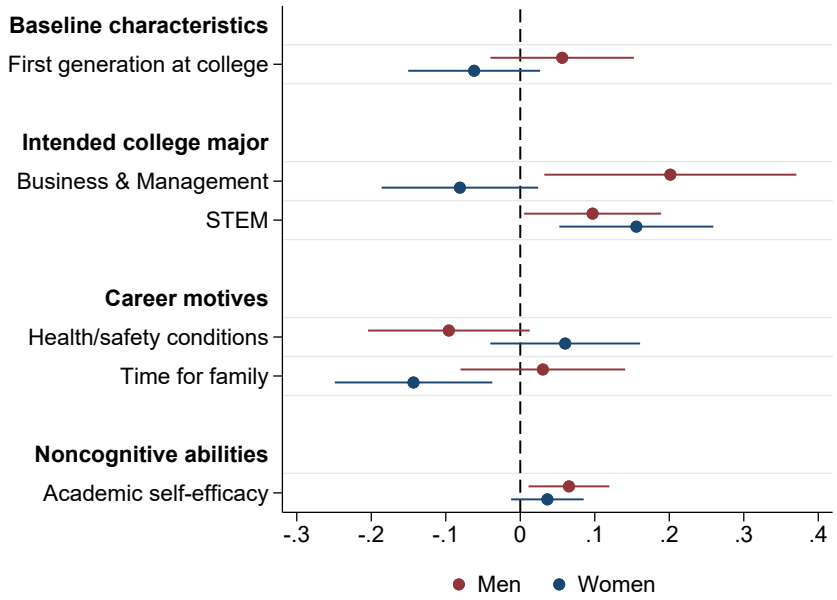
To provide further understanding of contributions via the explained and unexplained part, Figure 2 plots selected coefficients from the separate regressions for men (Equation (1)) and women (Equation (2)). We see that intending to study a STEM field and a positive academic self-efficacy is associated with higher earnings for both men and women. Table 1 shows that these factors are more prevalent for men, resulting in positive endowment effects. On the other hand, very high preferences for *time for family* and being potential first generation at college have differential effects by gender; they are associated with lower earnings for women but not for men. The former association is large and significant for women. Having a very high preference for *time for family* reduces female wage expectations by 14.8%, while men are virtually unaffected. Similarly, studying Business & Management is associated with higher earnings for men but not for women. These are the differences that are captured by the unexplained part in Table 1.

[17] Table 1 shows that women are less likely to enroll in STEM, which is on average associated with higher earnings.

[18] While there are large gender differences for both measures of confidence (see Table 1), self-confidence is associated with higher expected earnings only for women.

[19] Since all observations in the pooled sample are either conditional on a hypothetical bachelor's or master's degree, the gender gap for wage expectations associated with a bachelor's degree would be 0.023 log points higher than the average (see Appendix Table D.2 for details).

Figure 2: Effect of preferences on average expected earnings (Pooled).



Notes: This figure shows coefficient plots (95% confidence interval) for the effect of selected variables on average expected earnings pooled over degrees for men (red) and women (blue). Source: *Berliner-Studienberechtigten-Panel* waves 1-3.

As discussed earlier, the unexplained part may reflect expected discrimination. The constant (0.111 log points) could reflect a fixed level of discrimination due to unobservables while the other unexplained part could capture discrimination associated with observable characteristics. Women who place an importance on family time or who are potential first generation at college may plausibly expect to face greater discrimination on the labor market, and therefore hold lower wage expectations than comparable male counterparts. Women may also anticipate facing a more discriminating environment in Business & Management field, where wages are potentially less regulated and subject to negotiation.[20] A further possibility, however, is that the unexplained part reflects sorting into lower paying occupations or industries conditional on intended major and other characteristics, rather than discrimination. Such sorting can reflect the cultural norm in Germany that women tend to be the primary caregivers. As such they may expect to earn less even working full-time e.g. due to the requirement of more flexible working hours. The large contribution of *time for family* as a career motive via the unexplained part is consistent with this.

[20] Gender differences in negotiation styles can especially help in explaining differences in advancement rates (see e.g. Bertrand, 2018). Kiessling et al. (2019) argue that gender differences in negotiation styles are important determinants for both expected wages and starting wages.

4.2 Other statistical moments

So far, we have presented decomposition results for the mean of average expected wages, pooling across education scenarios (master's and bachelor's degree). We now turn to a decomposition of statistical moments of both the individual distribution of wage expectations (i.e. minimum and maximum expectations for different degrees), as well as a decomposition of the aggregate distribution of average wage expectations at different quantiles, in order to gain further insight into the formation of wage expectations.

4.2.1 Individual distribution

Figure 3 plots contributions via the explained and unexplained part for selected variables using decompositions for minimum and maximum expected wages as well as for master and bachelor scenarios, separately. The Figure also reports the sum of each factors contribution via the explained and unexplained part as a percentage of the overall gender gap to give an idea of overall importance of each factor for each expectation scenario. Underlying detailed decomposition results and coefficients are presented in Appendix Section D.

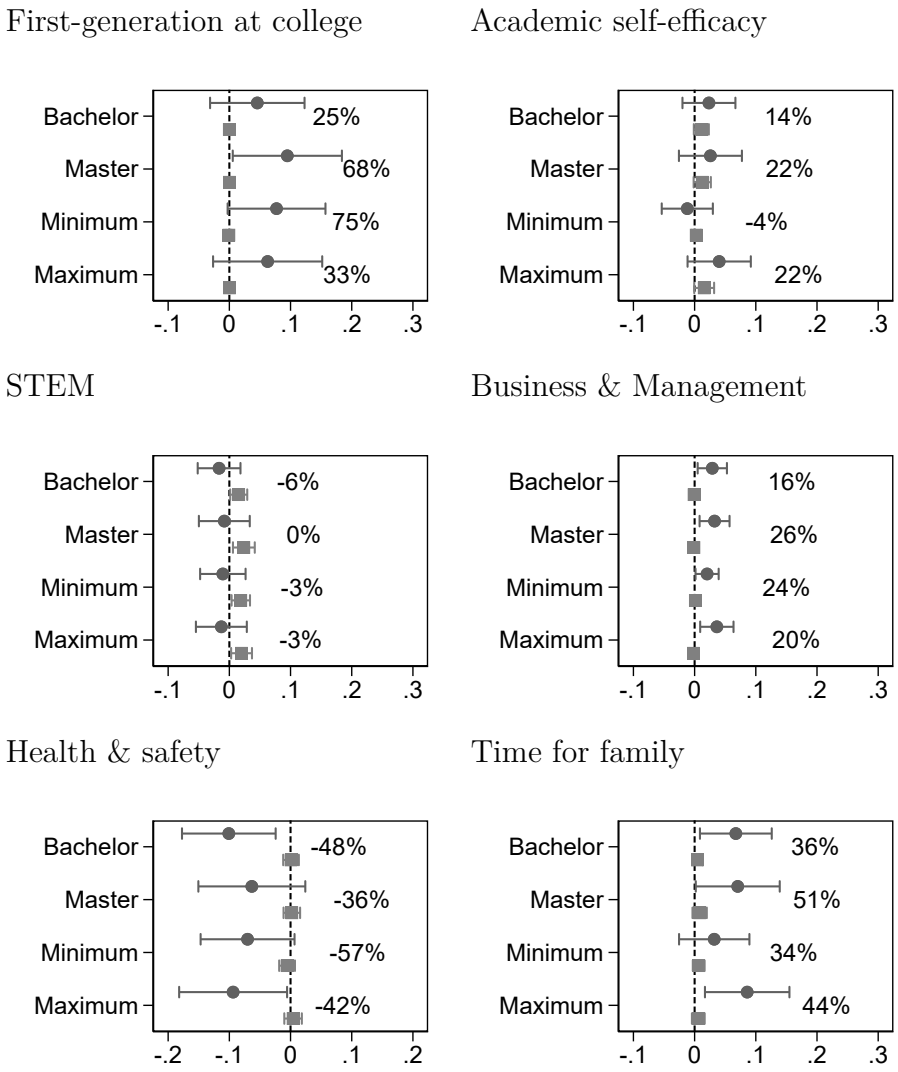
Overall, the share explained by differences in coefficients (*unexplained part*) is considerably larger for the selected coefficients than the share of the gap explained by differences in endowments (*explained part*). This is especially true for potential first generations at college and career motives such as a very high preference for *good health and safety conditions* or *time for family*. For some components, these factors explain over half of the gender gap in wage expectations alone.[21] For other factors that might indicate occupational sorting, such as intended STEM enrollment, contributions via the explained part are slightly more comparable in size to contributions via the unexplained part.

Splitting out the expectations in this way demonstrates that certain factors have larger impacts on maximum expected wages rather than the minimum expected wages. This difference is especially clear for preferences for time with family, and for academic self-efficacy. Conversely, other factors, such as first-generation at college, appear to have a larger impact through minimum expected wages. There are also differences in the coefficients depending on the education scenario, i.e. wage expectations associated with a bachelor's and a master's degree. Certain factors such as first-generation and *time for family* are associated with larger penalties in expected wages and hence with a larger contribution under the master track, and other such as health and safety, having a larger contribution under the bachelor track.

If certain factors contribute mainly to expected gender gaps in maximum earnings and scenarios with higher educational attainment, we interpret them as affecting mainly higher career paths. If e.g. women expect certain career motives to have a detrimental

[21] Note that even though these two factors would jointly overexplain the gap, other factors explain negative shares and therefore counteract.

Figure 3: Contributions via explained and unexplained part for selected variables.



Notes: This figure shows absolute contributions of different factors to the gender gap in wage expectations via the unexplained part (dark grey circles) and the explained (light grey squares) in log points and for different components: Bachelor's and master's degree, expected minimum and maximum wages). Joint relative contributions via explained and unexplained part are reported in percent. Source: *Berliner-Studienberechtigten-Panel* waves 1-3.

effect on the upper end of their expected earnings distribution (i.e. they expect a lower maximum wage), this might imply that they expect (having to) forgo higher career paths. A good example is the preference for *time for family*. Appendix Section E is dedicated to the role of this preference and shows that expected earnings losses are comparable in size to actual motherhood penalties in Germany.

The only other job preference that seems to have a large and significant (negative) impact on the gender gap in wage expectations is the one for good health and safety condition. Contributions via the unexplained part are large for both expected minimum and maximum wages and both college degrees. Appendix Figure D.2 shows that men who

hold this preference expect about 10% lower wages than their male counterparts. For female high school graduates, the association is rather positive. One possible interpretation of the differences in male and female coefficients is that *good health and safety conditions* are, on average, associated with very different types of jobs by men and women. Another notion is that men expect wage premiums for *risky* jobs as a form of compensating differential (see e.g. Biddle and Zarkin, 1988). Sorting of men into occupations with high risk premia could be one way of how this factor translates into realized earnings gaps.[22]

The factor with the largest single contributions (via the unexplained part for master's degrees and minimum expected wages) is being potential first generation at college. Especially female high school graduates without college educated father or mother often hold lower wage expectations than their female counterparts, while for men the association is rather positive (Appendix Figures D.1 and D.2). However, coming from a family with lower educational attainment might not only increase the gap in wage expectations, but could also be associated with a lower probability of college enrollment. In turn, also wage expectations for a college degree and thus potential returns to college could affect college enrollment. Another factor that is also likely to affect college enrollment, namely the individuals academic self-efficacy, seems to contribute mostly to the gender gap in expectations for a master's degree and maximum earnings.

4.2.2 Overall distribution

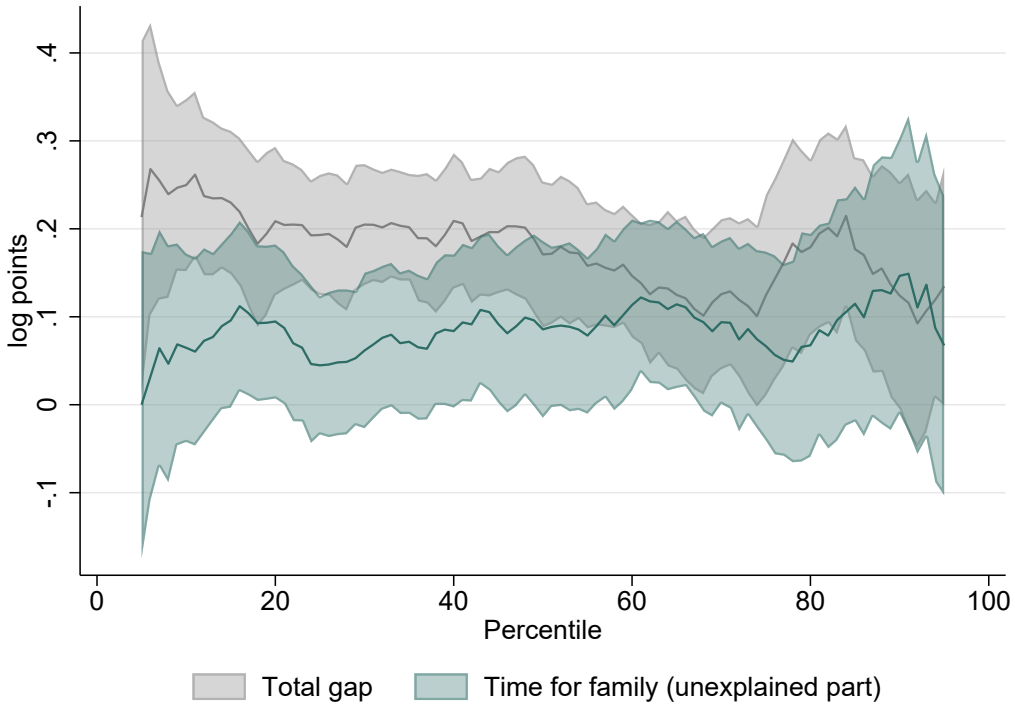
An extension of the OB decomposition based on recentered influence functions (RIF) allows to assess the gender gap at other statistical moments than the mean, using quantile regressions (see Firpo et al., 2018, for details). Here, we use this RIF decomposition to decompose the gender gap at different quantiles. As in the main analysis, we again use pooled average wage expectations of all individuals pooled as outcome. Before, we showed that e.g. *time for family* as a career motive has a higher relative importance for individual maximum vs. minimum wage expectations. If this career motive also explains a larger share of the gender gap at higher quantiles of the overall distribution of average expectations, this would confirm our interpretation, that already after high school graduation, many women expect having to give up higher career paths in order to take on family responsibilities.

Figure 4 shows decomposition results along the distribution (5^{th}–95^{th} percentile). In line with e.g. Briel et al. (2022) and Kiessling et al. (2019), the gender gap is largest at lower quantiles and smaller at higher quantiles. In contrast, the share attributable differences in coefficients (unexplained part) of preferences for *time for family* is relatively stable and tends to increases along the distribution. Hence, the relative importance of *time for family* is highest for individuals that expect to be top earners. Around the 90^{th}

[22] See Table D.3 for a decomposition of the implicitly expected range (y_{max} - y_{min}) as a measure of earnings risk. Overall, the same factors that affect average wage expectations also affect the range with similar relative importance.

quantile, differences in subjective beliefs about the returns to *time for family* commitments even overexplain the gap. Effectively, other factors in the decomposition counteract this.[23] The increasing relative importance of *time for family* at higher quantiles of the distribution of wage expectations supports our interpretation, that, at early age and in contrast to men, women expect having to trade off career opportunities against family commitments.

Figure 4: RIF decomposition for different percentiles (pooled).



Notes: This figure shows results of a Oaxaca-Blinder type RIF decomposition (Firpo et al., 2018) at each percentile (5 to 95). Men serve as reference group. The outcome variable is average expected wages of high school graduates pooled over education scenarios (bachelor's and master's degree). Total gap in gray and contribution of *time for family* via unexplained part in green. The decomposition controls for all baseline characteristics, cognitive and noncognitive ability measures, intended college major and career motives. Bootstrapped 95% confidence intervals (400 replications). N=948. Source: *Berliner-Studienberechtigten-Panel* waves 1-3.

5 College Enrollment

Initially, we motivated the relevance of wage expectations as determinant of educational investment and thus future realized gender wage gaps. Lower wage expectations of female high school graduates may affect the actual gender wage gap through lower expected returns to education and thus lower incentives to invest in education. Addi-

[23] Appendix Table D.4 shows more detailed RIF decomposition results at the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} quantile for the factors of interest from Figure 3. Overall, the results are similar to the decomposition of average wage expectations.

tionally, these lower expectations may also impact through the formation of reservation wages which are likely to influence the decision to accept wage offers. While we cannot address the latter (see e.g. Kiessling et al., 2019, on this topic), we can complement a strand of literature that empirically investigates whether wage expectations are related to college enrollment (Schweri and Hartog, 2017) and resulting socioeconomic gaps (Boneva and Rauh, 2020) and gender gaps (Belfield et al., 2020) in enrollment.[24]

In addition, we can corroborate findings regarding the relevance of gender-specific wages for additional educational investment. Studies have shown that gender-specific differences in work values (Berkes et al., 2022) and the anticipated time commitment to the work or family domain (Wiswall and Zafar, 2021) influence gender differences in educational choices. Given this evidence, a closer look at how expected returns to education could influence gender differences in a key educational choice: enrolling in college or not, goes beyond existing evidence, examining choice models (Boneva et al., 2022, 2021), by investigating how expectations influence actual behavior.[25] Studies have shown that young people's decision processes are context dependent, whereby they often rely on observed educational choices or occupations, which encompass gender-specific stereotypes (Wiswall and Zafar, 2015, 2021, Mead, 2022). Therefore, considering this evidence on the prevailing gender differences in choices, we examine how wage expectations could influence enrollment in college education by also looking at differences between female and male students, as such choices explain a significant portion of the gender wage gap (Reuben et al., 2017, Wiswall and Zafar, 2018). Information on college enrollment and expected returns to college is available for 445 individuals. Our results give supporting evidence that expected returns to college education and enrollment are indeed linked.

Table 3 shows descriptive statistics for college enrollment and expected returns to college for the complete subsample as well as by gender and socioeconomic status (SES).[26] College enrollment (C) measures the percentage of high school graduates in our sample that enrolls for college within two years of graduation, i.e. between winter term 2014 and summer term 2016. Overall enrollment rates and enrollment rates by gender are representative of actual population shares. Surprisingly however, graduates with and without at least one parent with a tertiary degree (i.e. high and low SES graduates) show similar enrollment rates. Enrollment rates after two years are only 2.2 percentage points (pp) lower for graduates that are the potential first generation at college (68.5% vs 66.3%). This difference is not statistically significant. Expected returns to college (R) are defined as the difference between expected earnings with a bachelor's and a vocational degree (in

[24] The 2015 cohort of high school graduates in Germany shows a gender gap of eight percentage points in the enrollment rates of high school graduates after two years (male: 71%, female: 63%, see Destatis (2020)).

[25] See e.g. Favara et al. (2021) on how wage expectations are associated with actual college enrollment in developing countries.

[26] We define high SES as having at least one parent with college education and low SES as being potential first generation at college

Table 3: Expected returns and college enrollment by subgroup.

	Main	Gender		SES	
		Male	Female	High	Low
	(1)	(2)	(3)	(4)	(5)
Enrollment (C)	67.2	71.2	64.4	68.5	66.3
Expected Returns (R)	0.340	0.345	0.336	0.311	0.359*
	(0.270)	(0.268)	(0.272)	(0.260)	(0.276)
N	445	184	261	181	264

Notes: This table presents descriptive statistics on college enrollment rates (in percent) and expected returns to college (in log points) for the complete college enrollment analysis sample and by gender and SES. Standard deviations in parentheses are reported for expected returns to college. Significance stars indicate differences based on a two-sided t-test within the respective subgroups. Source: *Berliner-Studienberechtigten-Panel* waves 1-5. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

log points). Expected returns to college are large[27] for the subsample at hand and vary extensively[28], but there is virtually no difference between genders. Expected returns to college held by low SES students (0.359 log points, about 48.7%) are significantly higher than those held by high SES graduates (0.311 log points, about 41.2%). Table C.2 shows that the actual returns to college in 2014 range from 17.6% for females to 20.2% for males, meaning that high SES students are better informed about the returns to college.

To estimate the association between college enrollment and the expected returns to college we rely on a simple linear probability model:

$$C_i = \beta_0 + \beta_1 R_i + \beta_2 F_i + \beta_3 (R_i \times F_i) + \boldsymbol{X_i'\gamma} + \mu_s + \varepsilon_i, \tag{5}$$

where $C_i \in \{0, 1\}$ signals college enrollment within two years of graduation, $F_i \in \{0, 1\}$ is a dummy that signals if a student is female, $\boldsymbol{X_i}$ is the vector of covariates from the decomposition and μ_s are school (s) fixed effects. Expected returns to college R_i simply measure the difference between expected earnings with a bachelor's and a vocational degree ($ln(y_{i,d=\text{BA}}) - ln(y_{i,d=\text{voc}})$). Table 4 shows the corresponding estimation results. Results in Panel A are based on a bivariate model that includes only expected returns as explanatory variables. We also check for robustness of these results to the inclusion of school fixed effects (Panel B) and school fixed effects plus the set of control variables used in the decomposition (Panel C). In addition to the baseline coefficient in the respective first column, each Panel shows results from a fully interacted model with interactions by gender and by socioeconomic status. In all specifications we account for clustering at the

[27] Mean: $45.7\% \approx 100\% \times (e^{(2\times 0.340 + 0.270^2)/2} - 1)$

[28] SD: $40\text{pp} \approx 100\text{pp} \times \sqrt{(e^{0.270^2} - 1) \times e^{2\times 0.340 + 0.270^2}}$

school level, reflecting e.g. possible peer effects. As the number of clusters is small (27 schools), we also report p-values from a wild cluster bootstrap (WCB, Roodman et al., 2019). Additionally, Figure 5 shows the underlying subgroup coefficients for both genders as well as low and high SES graduates.

Panel A of Table 4 shows that in the baseline specification without control variables and fixed effects, there is a positive and significant association between the expected returns to college and college enrollment. Column 1 indicates that a 1% increase in the expected returns to college is associated with an on average 0.147 pp higher probability of subsequent college enrollment.[29] Column 2 additionally shows the interaction for potential first generation students at college. Higher expected returns are associated with a significantly higher enrollment rates for students who have at least one parent with a college degree. For potential first generation college students, the resulting subgroup coefficient is insignificant and close to zero (see Figure 5).[30] Lastly, Column 3 shows that while the association between expected returns and enrollment intentions for men is close to zero, the average association between the two variables is significantly higher for female high school graduates. The resulting subgroup coefficient for women is about twice as large (0.289) as the average effect and statistically different from zero at the 5% level, confirmed by a WCB. This result suggests, that monetary concerns play a larger role for women than for men when deciding whether to enroll at college or not.

Table 4: Expected returns to further education and college enrollment.

	Panel A			Panel B			Panel C		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Expected returns to college	0.147**	0.342***	-0.067	0.096	0.322***	-0.063	0.094	0.320***	-0.061
	(0.061)	(0.085)	(0.116)	(0.063)	(0.087)	(0.132)	(0.055)	(0.097)	(0.126)
	[0.037]	[0.002]	[0.588]	[0.155]	[0.003]	[0.638]	[0.091]	[0.002]	[0.626]
× First generation		-0.305**			-0.354**			-0.368**	
		(0.148)			(0.152)			(0.156)	
		[0.044]			[0.022]			[0.020]	
× Female			0.357*			0.274			0.269
			(0.174)			(0.200)			(0.195)
			[0.064]			[0.183]			[0.180]
School FE				✓	✓	✓	✓	✓	✓
Controls							✓	✓	✓
N	445	445	445	445	445	445	445	445	445

Notes: This table presents estimates from a linear probability model with college enrollment as dependent variable. Expected returns to college are calculated as the difference between (log) expected wages with a bachelor's degree and a vocational degree. Control variables comprise the variables included in the decomposition. Standard errors in parentheses allow for clustering at the school level. P-values in brackets are based on a wild cluster bootstrap with 1999 replications. Source: *Berliner-Studienberechtigten-Panel* waves 1-5. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

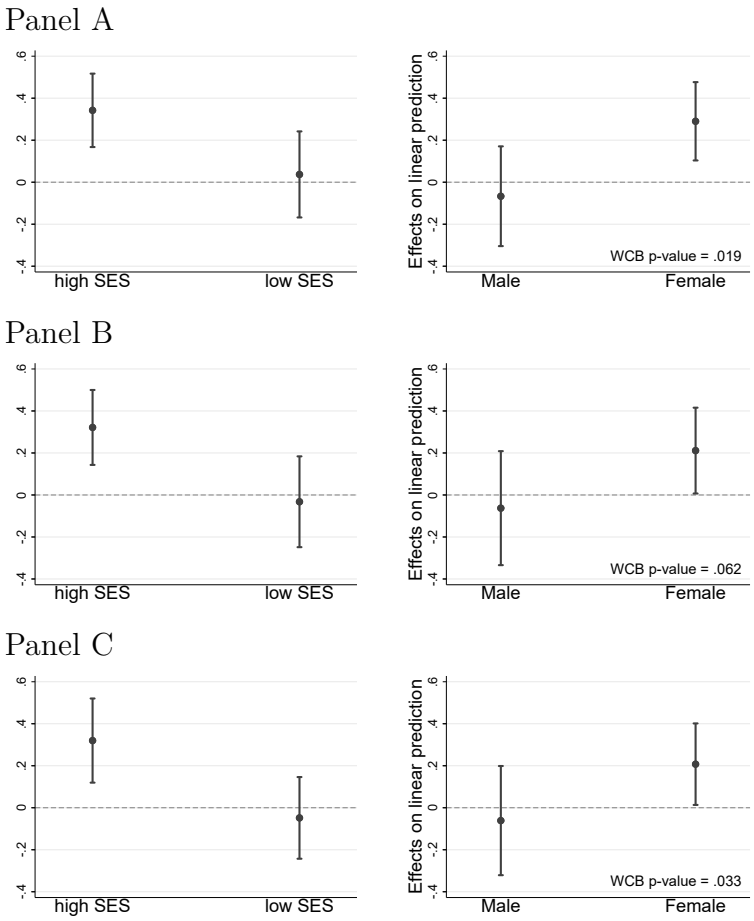
Lastly, Panels B and C show results from specifications that include school fixed effects and the set of control variables from the decomposition in addition to school fixed

[29] This associations might seem small, but Table 3 reveals that expected returns are large (about 45.7% for the subsample at hand) and vary extensively (standard deviation about 40 pp). I.e. a one standard deviation change in expected returns is associated with an about 5.9 pp higher enrollment probability.

[30] One might expect that this association is explained by high SES graduates simultaneously showing higher wage expectations and college enrollment rates. However, Table 3 shows that this is not the case.

effects, respectively. Overall, the results follow a similar pattern as in Panel A. Again, large and significant interactions for different socioeconomic backgrounds exist, while the average association between expected returns to college and college enrolment is smaller and not significant. The subgroup coefficient for men is again small and negative, while the difference to the association for women is large. Even though this interaction is not statistically significant, Figure 5 shows that the underlying subgroup coefficients for women are large and positive (0.211 in Panel B and 0.208 in Panel C) and statistically significant at the 10% and 5%-level, again confirmed by a wild cluster bootstrap. Overall, wage expectations seem to matter for the college enrollment of women, while men seem to be less affected in their enrollment decisions by such considerations. This association is robust to an inclusion of school fixed effects and a large range of control variables.

Figure 5: Average marginal associations by gender and SES.



Notes: This figure shows subgroup coefficient plots (95% confidence intervals) by gender and socioeconomic background for each Panel in Table 4. Standard errors allow for clustering at the school level. P-values from a wild cluster bootstrap with 1999 replications are reported for females. Source: *Berliner-Studienberechtigten-Panel* waves 1-5.

The results above suggest that expected returns to college play a larger role for college enrollment of high school graduates from more educated families and for women. Both

associations cannot be explained by women or high SES graduates simultaneously having higher expected returns to college and enrollment rates. We therefore provide supporting evidence for the hypothesis that wage expectations matter for human capital investment and thus predictive of actual wage differences between genders. Previous work with Best Up found positive effects of information on the returns to college on both enrollment intentions (Peter and Zambre, 2017) and actual college enrollment (Peter et al., 2021), supporting the notion that a gender gap in wage expectations might affect educational choices and thus entrench subsequent earnings inequalities.

6 Conclusion

Based on a unique survey in which we elicit the range of high school graduates' wage expectations for different degree types, this study investigates respective gender differences and how different factors, such as preferences for certain job attributes and college majors, as well as cognitive and noncognitive abilities, shape them. The results of our analysis can be summarized in four key findings.

First, already at high school graduation female students expect to earn over 15% less than their male counterparts. This observation cannot be explained by women being better informed about wages that are currently paid on the labor market. Second, female high school graduates expect large family penalties, even conditional on full-time employment. These family penalties are especially large for individual maximum expectations, higher degree types and higher quantiles of the overall distribution of wage expectations. This indicates that, early on, women expect having to trade off advancements opportunities and thus leadership positions in order to take on family responsibilities. However, female high school graduates still underestimate the size of family penalties when compared to contemporaneous child penalties, a subset of family penalties, observed in Germany.

Third, including expected family penalties, most factors affect the gender gap in wage expectations mainly through differences in coefficients rather than differences in endowments. Related studies often neglect the contribution of expected family penalties and other factors via the unexplained part therefore underestimate its relative importance. Finally, differences in the expected returns to higher education are likely to impact female human capital investment. Thereby, gender gaps in wage expectations can also affect the formation of gender gaps in realized earnings.

The findings of this study provide some evidence for a selection of women into lower paying occupations or industries in exchange for certain job characteristics. However, certain job characteristics and majors are associated with fundamentally different expected returns by men and women, which signals expected gender-based discrimination. As such expected discrimination might itself be simultaneously associated with the formation of preferences, expected discrimination and sorting can go hand in hand.

Future policies that affect early wage expectations and encourage women to pursue

higher career paths can be roughly divided into two groups. The first set of policies primarily includes measures that aim at improving the compatibility of family and career for women. For instance, such measures might aim to increase incentives for men to undertake a larger share of care work: An adjustment of income splitting, the expansion of parental allowances, especially for fathers, a reduction of statutory working hours, or family leave regulations also for those who take care of older parents or disabled relatives might be measures which help to combine work and family time. Other measures might aim at the expansion of publicly funded daycare for children and disabled elderlies or encourage *top sharing*, i.e. shared leadership positions between men and women, associated with a reduction in working hours. If young women observe such policies and update their beliefs about the compatibility of family and career accordingly, also expected family penalties and hence the gender gap in wage expectations might be reduced.

The second set of policies includes measures that directly aim to adjust wage expectations and career plans. However, such measures face a dilemma: On the one hand, information campaigns about the costs of family commitments can be beneficial as they could enable more informed career and family care decisions of women and thus might give them more bargaining power when it comes to the intra-household division of labor. On the other hand, we present evidence that young women underestimate family penalties when compared to contemporaneous labor market data. If this can be interpreted as optimism or motivation to arrange family and career, it might be counterproductive to lower their wage expectations by informing them about family penalties. Thus, possible information campaigns should also include information on family allowances and career counseling for women, to avoid discouragement.

While the results of this study are descriptive in nature, they provide evidence on why women (expect having to) give up on higher paying career paths and postulate a fruitful perspective for future research on the causal mechanisms that shape wage expectations and their importance for students' future career and human capital investment.

References

Alonso-Borrego, C. and A. Romero-Medina (2016): "Wage Expectations for Higher Education Students in Spain," *LABOUR*, 30, 1–17.

Arcidiacono, P., V. J. Hotz, and S. Kang (2012): "Modeling college major choices using elicited measures of expectations and counterfactuals," *Journal of Econometrics*, 166, 3–16.

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

Attanasio, O. P. and K. M. Kaufmann (2014): "Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender," *Journal of Development Economics*, 109, 203–216.

——— (2017): "Education choices and returns on the labor and marriage markets: Evidence from data on subjective expectations," *Journal of Economic Behavior & Organization*, 140, 35–55.

Belfield, C., T. Boneva, C. Rauh, and J. Shaw (2020): "What Drives Enrolment Gaps in Further Education? The Role of Beliefs in Sequential Schooling Decisions," *Economica*, 87, 490–529.

Berkes, J., F. Peter, C. K. Spiess, and F. Weinhardt (2022): "Information Provision and Postgraduate Studies," *Economica*, 89, 627–646.

Bertrand, M. (2011): "New perspectives on gender," in *Handbook of Labor Economics*, ed. by D. Card and O. Ashenfelter, Elsevier, vol. 4, part B, chap. 17, 1543–1590.

——— (2018): "Coase lecture–the glass ceiling," *Economica*, 85, 205–231.

Bertrand, M. and K. F. Hallock (2001): "The Gender Gap in Top Corporate Jobs," *ILR Review*, 55, 3–21.

Betts, J. R. (1996): "What do students know about wages? Evidence from a survey of undergraduates," *Journal of Human Resources*, 31, 27–56.

Biddle, J. E. and G. A. Zarkin (1988): "Worker preference and market compensation for job risk," *The Review of Economics and Statistics*, 660–667.

Blau, F. D. and M. A. Ferber (1991): "Career plans and expectations of young women and men: The earnings gap and labor force participation," *Journal of Human Resources*, 26, 581–607.

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., T. Buser, A. Falk, and F. Kosse (2022): "The Origins of Gender Differences in Competitiveness and Earnings Expectations: Causal Evidence from a Mentoring Intervention," *CEPR Discussion Paper No. 17008*.

Boneva, T., M. Golin, and C. Rauh (2021): "Can Perceived Returns Explain Enrollment Gaps in Postgraduate Education?" *Labour Economics*, 101998.

Boneva, T. and C. Rauh (2020): "Socio-Economic Gaps in University Enrollment: The Role of Perceived Pecuniary and Non-Pecuniary Returns," *Working paper*.

Bong, M. and E. M. Skaalvik (2003): "Academic self-concept and self-efficacy: How different are they really?" *Educational psychology review*, 15, 1–40.

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.

Bronson, M. A. (2015): "Degrees are forever: Marriage, educational investment, and lifecycle labor decisions of men and women," *Unpublished manuscript*.

Brown, S. and K. Taylor (2013): "Reservation wages, expected wages and unemployment," *Economics Letters*, 119, 276–279.

Brunello, G., C. Lucifora, and R. Winter-Ebmer (2004): "The wage expectations of European business and economics students," *Journal of Human Resources*, 39, 1116–1142.

Caliendo, M., W.-S. Lee, and R. Mahlstedt (2017): "The gender wage gap and the role of reservation wages: New evidence for unemployed workers," *Journal of Economic Behavior & Organization*, 136, 161– 73.

Carvajal, M. J., D. Bendana, A. Bozorgmanesh, M. A. Castillo, K. Pourmasiha, P. Rao, and J. A. Torres (2000): "Inter-gender differentials between college students' earnings expectations and the experience of recent graduates," *Economics of Education Review*, 19, 229–243.

Ceci, S. J. and W. M. Williams (2011): "Understanding current causes of women's underrepresentation in science," *Proceedings of the National Academy of Sciences*, 108, 3157–3162.

Chevalier, A. (2007): "Education, occupation and career expectations: Determinants of the gender pay gap for UK graduates," *Oxford Bulletin of Economics and Statistics*, 69, 819–842.

Collischon, M. (2019): "Is there a glass ceiling over Germany?" *German Economic Review*, 20, e329–e359.

Cotton, J. (1988): "On The Decomposition of Wage Differentials," *The Review of Economics and Statistics*, 236–243.

Daymont, T. N. and P. J. Andrisani (1984): "Job preferences, college major, and the gender gap in earnings," *Journal of Human Resources*, 408–428.

Dehne, M. and J. Schupp (2007): "Persönlichkeitsmerkmale im Sozio-oekonomischen Panel (SOEP): Konzept, Umsetzung und empirische Eigenschaften," *DIW Research Notes No. 26*, German Institute for Economic Research.

Delaney, L., C. Harmon, and C. Remond (2010): "Decomposing gender differences in college student earnings expectations," *UCD Geary Institute Discussion Paper Series; WP 10 38*.

Destatis (2012): "Bildung und Kultur. Studierende an Hochschulen, WS 2011/2012," *Fachserie 11, Reihe 4.1*, Statistisches Bundesamt, Wiesbaden.

——— (2020): "Nichtmonetäre hochschulstatistische Kennzahlen," *Fachserie 11 Reihe 4.3.1*, Statistisches Bundesamt, Wiesbaden.

Dominitz, J. and C. F. Manski (1996): "Eliciting student expectations of the returns to schooling," *Journal of Human Resources*, 31, 1–26.

Ehlert, M., C. Finger, A. Rusconi, and H. Solga (2017a): "Applying to college: Do information deficits lower the likelihood of college-eligible students from less-privileged families to pursue their college intentions?: Evidence from a field experiment," *Social Science Research*, 67, 193–212.

Ehlert, M., F. Peter, C. Finger, A. Rusconi, H. Solga, K. C. Spiess, and V. Zambre (2017b): "Berliner-Studienberechtigten-Panel (Best Up) – Data Description and Documentation," *DIW Data Documentation*, 90, German Institute for Economic Research.

Ehrmantraut, L., P. Pinger, and R. Stans (2020): "The Expected (Signaling) Value of Higher Education," *CRC TR 224 Discussion Paper Series*.

Eliophotou-Menon, M. (1997a): "Perceived economic benefits of higher education: the case of Cyprus," *Education Economics*, 5, 53–61.

——— (1997b): "Perceived rates of return to higher education in Cyprus," *Economics of Education Review*, 16, 425–430.

Favara, M., P. Glewwe, C. Porter, and A. Sanchez (2021): "Expecting better? How young people form their earnings expectations," *IZA Discussion Paper No. 14289*.

Ferber, M. A. and W. W. McMahon (1979): "Women's Expected Earnings and Their Investment in Higher Education," *The Journal of Human Resources*, 14, 405–420.

Fernandes, A., M. Huber, and G. Vaccaro (2021): "Gender differences in wage expectations," *Plos one*, 16, e0250892.

Filippin, A. and A. Ichino (2005): "Gender wage gap in expectations and realizations," *Labour Economics*, 12, 125–145.

Firpo, S. P., N. M. Fortin, and T. Lemieux (2018): "Decomposing Wage Distributions Using Recentered Influence Function Regressions," *Econometrics*, 6.

Flabbi, L. and A. Moro (2012): "The effect of job flexibility on female labor market outcomes: Estimates from a search and bargaining model," *Journal of Econometrics*, 168, 81–95.

Fortin, N., T. Lemieux, and S. Firpo (2011): "Decomposition Methods in Economics," in *Handbook of Labor Economics*, Elsevier, vol. 4, 1–102.

Fortin, N. M. (2008): "The gender wage gap among young adults in the united states the importance of money versus people," *Journal of Human Resources*, 43, 884–918.

Francesconi, M. and M. Parey (2018): "Early Gender Gaps Among University graduates," *European Economic Review*, 109, 63–82.

Giustinelli, P. (2022): "Expectations in Education: Framework, Elicitation, and Evidence," *Permanent HCEO Working Paper*.

Goldin, C. (2014): "A grand gender convergence: Its last chapter," *American Economic Review*, 104, 1091–1119.

Goldin, C. and L. F. Katz (2011): "The cost of workplace flexibility for high-powered professionals," *The Annals of the American Academy of Political and Social Science*, 638, 45–67.

Grove, W. A., A. Hussey, and M. Jetter (2011): "The gender pay gap beyond human capital," *Journal of Human Resources*, 46, 827–874.

Guiso, L., T. Jappelli, and L. Pistaferri (2002): "An empirical analysis of earnings and employment risk," *Journal of Business & Economic Statistics*, 20, 241–253.

Hastings, J., C. A. Neilson, and S. D. Zimmerman (2015): "The effects of earnings disclosure on college enrollment decisions," Tech. rep., National Bureau of Economic Research.

Hastings, J. S., C. A. Neilson, A. Ramirez, and S. D. Zimmerman (2016): "(Un) informed college and major choice: Evidence from linked survey and administrative data," *Economics of Education Review*, 51, 136–151.

Heckman, J. J., J. Stixrud, and S. Urzua (2006): "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior," *Journal of Labor Economics*, 24, 411–482.

Huntington-Klein, N. (2015): "Subjective and projected returns to education," *Journal of Economic Behavior & Organization*, 117, 10–25.

Jerrim, J. (2011): "Do UK higher education students overestimate their starting salary?" *Fiscal Studies*, 32, 483–509.

——— (2015): "Do college students make better predictions of their future income than young adults in the labor force?" *Education Economics*, 23, 162–179.

Jessen, J. (2022): "Culture, children and couple gender inequality," *European Economic Review*, 150, 104310.

Jessen, J., C. K. Spiess, S. Waights, and K. Wrohlich (2021): "Sharing the caring? The gender division of care work during the Covid-19 pandemic in Germany," *IZA Discussion Paper No. 14457*.

Kiessling, L., P. Pinger, P. K. Seegers, and J. Bergerhoff (2019): "Gender Differences in Wage Expectations: Sorting, Children, and Negotiation Styles," *IZA Discussion Paper No. 12522*.

Kirsch, A. and K. Wrohlich (2020): "Frauenanteile in Spitzengremien großer Unternehmen steigen - abgesehen von Aufsichtsräten im Finanzsektor," *DIW Wochenbericht*, 87, 38–49.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimuller (2019): "Child Penalties across Countries: Evidence and Explanations," *AEA Papers and Proceedings*, 109, 122–26.

Koerselman, K. and R. Uusitalo (2014): "The risk and return of human capital investments," *Labour Economics*, 30, 154–163.

Kuziemko, I., J. Pan, J. Shen, and E. Washington (2018): "The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?" *NBER Working Paper No. 2470*.

Manski, C. F. (1993): "Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?" in *Studies of Supply and Demand in Higher Education*, ed. by C. T. Clotfelter and M. Rothschild, University of Chicago Press, chap. 2, 43–60.

——— (2004): "Measuring expectations," *Econometrica*, 72, 1329–1376.

Mas, A. and A. Pallais (2017): "Valuing alternative work arrangements," *American Economic Review*, 107, 3722–59.

Mazei, J., J. Hüffmeier, P. A. Freund, A. F. Stuhlmacher, L. Bilke, and G. Hertel (2015): "A meta-analysis on gender differences in negotiation outcomes and their moderators." *Psychological bulletin*, 141, 85.

Mazza, J. and J. Hartog (2011): "Do they understand the benefits from education? Evidence on Dutch high school students' earnings expectations," *IZA Discussion Paper No. 5714*.

McCrae, R. and P. Costa (1996): "Toward a new generation of personality theories: Theoretical contexts for the five-factor model," in *The five factor model of personality: Theoretical perspectives*, ed. by J. Wiggins, New York: Guilford, chap. 3, 51–87.

McMahon, W. W. and A. P. Wagner (1981): "Expected Returns to Investment in Higher Education," *The Journal of Human Resources*, 16, 274–285.

Mead, D. (2022): "The gender gap in university enrolment: evidence from subjective expectations," *Education Economics*, 1–23.

Montmarquette, C., K. Cannings, and S. Mahseredjian (2002): "How do young people choose college majors?" *Economics of Education Review*, 21, 543–556.

Neumark, D. (1988): "Employers' discriminatory behavior and the estimation of wage discrimination," *Journal of Human Resources*, 23, 279–295.

Oaxaca, R. L. (1973): "Male-Female wage differentials in urban labor markets," *International Economic Review*, 14, 693–709.

Oaxaca, R. L. and M. R. Ransom (1994): "On discrimination and the decomposition of wage differentials," *Journal of Econometrics*, 61, 5–21.

OECD (2018): *OECD Employment Outlook 2018*, Organisation for Economic Co-operation and Development.

Oreopoulos, P., T. von Wachter, and A. Heisz (2012): "The Short- and Long-Term Career Effects of Graduating in a Recession," *American Economic Journal: Applied Economics*, 4, 1–29.

Peter, F., A. Rusconi, H. Solga, and C. K. Spiess (2017): "Money alone is not enough: a new study looks at what may encourage high school graduates to choose college." *WZB Report 2017*, 69–71.

Peter, F., C. K. Spiess, and V. Zambre (2021): "Informing students about college: Increasing enrollment using a behavioral intervention?" *Journal of Economic Behavior & Organization*, 190, 524–549.

Peter, F. and V. Zambre (2017): "Intended college enrollment and educational inequality: Do students lack information?" *Economics of Education Review*, 60, 125–141.

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

Roodman, D., M. Ø. Nielsen, J. G. MacKinnon, and M. D. Webb (2019): "Fast and wild: Bootstrap inference in Stata using boottest," *The Stata Journal*, 19, 4–60.

Rotter, J. B. (1966): "Generalized expectancies for internal versus external control of reinforcement." *Psychological monographs: General and applied*, 80, 1.

Ruder, A. I. and M. V. Noy (2017): "Knowledge of earnings risk and major choice: Evidence from an information experiment," *Economics of Education Review*, 57, 80–90.

Schweri, J. and J. Hartog (2017): "Do wage expectations predict college enrollment? Evidence from healthcare," *Journal of Economic Behavior & Organization*, 141, 135–150.

Wasserman, M. (2019): "Hours constraints, occupational choice, and gender: Evidence from medical residents," *Working Paper*.

Webbink, D. and J. Hartog (2004): "Can students predict starting salaries? Yes!" *Economics of Education Review*, 23, 103–113.

Weinhardt, M. and J. Schupp (2011): "Multi-Itemskalen im SOEP Jugendfragebogen," *DIW Data Documentation No. 60*, German Institute for Economic Research.

Wiswall, M. and B. Zafar (2015): "How Do College Students Respond to Public Information about Earnings?" *Journal of Human Capital*, 9, 117–169.

——— (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.

Wolter, S. C. (2000): "Wage expectations: a comparison of Swiss and US students," *Kyklos*, 53, 51–69.

Wolter, S. C. and A. Zbinden (2002): "Labour market expectations of Swiss university students," *International Journal of Manpower*.

Zafar, B. (2013): "College major choice and the gender gap," *Journal of Human Resources*, 48, 545–595.

Zambre, V. (2018): "The Gender Gap in Wage Expectations: Do Young Women Trade off Higher Wages for Lower Wage Risk?" DIW Discussion Paper 1742, German Institute for Economic Research.

Zweck, B., G. Daniel, and A. Glemser (2019): "PostGrad Berliner-Studienberechtigten-Panel (Best Up) 2017 bis 2019," Tech. rep., Kantar Public Division Deutschland.

Appendix

A Item nonresponse

Table A.1: Partial non-response versus final sample.

	All		Women		Men	
	Non-Response	Sample (Difference)	Non-Response	Sample (Difference)	Non-Response	Sample (Difference)
Baseline characteristics:						
Female	0.571	0.029	1.000	0.000	0.000	0.000
Information intervention school	0.321	0.024	0.300	0.057	0.350	-0.023
Financial intervention school	0.363	0.005	0.364	-0.019	0.363	0.042
Migration background	0.548	-0.094***	0.557	-0.080*	0.535	-0.115**
First generation at college	0.636	-0.040	0.654	-0.057	0.612	-0.017
Academic high school	0.292	0.008	0.306	0.025	0.274	-0.020
Integrated high school	0.358	0.009	0.367	-0.026	0.345	0.060
Vocational high school	0.350	-0.017	0.327	0.001	0.381	-0.040
Fast track to vocational degree	0.040	0.004	0.044	0.008	0.036	-0.002
Cognitive abilities:						
Final high school GPA	2.419	0.125***	2.438	0.128***	2.388	0.120**
Verbal cognitive skills	9.478	0.774***	8.672	1.136***	10.557	0.361
Figural cognitive skills	10.727	0.429**	10.766	0.477**	10.674	0.350
Intended college major:						
Arts & Humanities	0.075	0.005	0.094	-0.003	0.049	0.014
Social Sciences & Economics	0.042	-0.011	0.054	-0.018	0.027	-0.003
Business & Management	0.106	0.009	0.101	0.029	0.112	-0.019
STEM	0.179	0.045*	0.091	0.062**	0.296	0.036
Teaching	0.042	0.043***	0.061	0.047**	0.018	0.036**
Law	0.048	-0.013	0.064	-0.032*	0.027	0.012
Health & Medicine	0.133	0.025	0.172	0.020	0.081	0.027
Other	0.025	0.020*	0.017	0.016	0.036	0.028
Missing information	0.040	-0.019*	0.037	-0.005	0.045	-0.040***
Career motives:						
High income	0.327	0.010	0.307	-0.002	0.355	0.031
Promotion possibilities	0.377	0.007	0.412	-0.022	0.332	0.044
Recognition	0.234	0.020	0.258	0.018	0.202	0.018
Interesting job	0.554	0.075**	0.590	0.072*	0.507	0.074
Independent working	0.274	0.016	0.308	-0.022	0.229	0.068
Social interaction	0.308	-0.019	0.377	-0.029	0.216	-0.016
Important for society	0.226	-0.041	0.242	-0.067**	0.205	-0.005
Help Others	0.282	-0.025	0.332	-0.017	0.216	-0.045
Spare time	0.132	0.026	0.138	0.005	0.123	0.057
Health/safety conditions	0.560	-0.010	0.654	-0.037	0.433	0.016
Time for family	0.424	-0.018	0.442	-0.003	0.399	-0.043
Noncognitive abilities:						
Openness	5.000	0.070	5.049	0.088	4.932	0.039
Extraversion	4.820	0.048	4.907	-0.027	4.699	0.149
Conscientiousness	4.868	0.028	5.085	-0.057	4.567	0.130
Neuroticism	4.314	-0.089	4.596	0.037	3.921	-0.309***
Agreeableness	5.307	-0.062	5.405	-0.016	5.173	-0.144
Locus of control (int.)	4.971	-0.033	4.972	-0.093*	4.968	0.058
Academic self-efficacy	3.015	0.110***	2.959	0.102*	3.088	0.132**
Self-confidence	4.832	0.084	4.625	0.102	5.126	0.074
Educational aspiration:						
Intended college enrollment	0.688	0.107***	0.687	0.118***	0.691	0.090**
N	520	513	297	308	223	205
N (Total)	1033		605		428	

Notes: This table presents differences in individual characteristics between students who answered the module on wage expectations for at least one education scenario and provided full information on covariates used in the decomposition and those who did not. Means and mean differences are based on a two-sided t-test. Source: *Berliner-Studienberechtigten-Panel* waves 1-3. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

The module on earnings expectations was included in the third wave (spring 2014) of the *Berliner-Studienberechtigten-Panel*, right after students graduated from high school and were about to decide on their post-secondary education. Table A.1 compares individual characteristics across students in our final sample (i.e. those who answered the module and gave full information on the covariates used in the decomposition), with other participants (i.e. those who show some non-response). Overall, the results suggest that item non-response in our questionnaire does not occur randomly. Students with higher cognitive abilities, students with a good academic self-efficacy, students aiming for majors in STEM and teaching, students with a very high preference for an interesting job, as well as students who intend to enroll in college are significantly more likely to provide information on their expected wages and covariates.

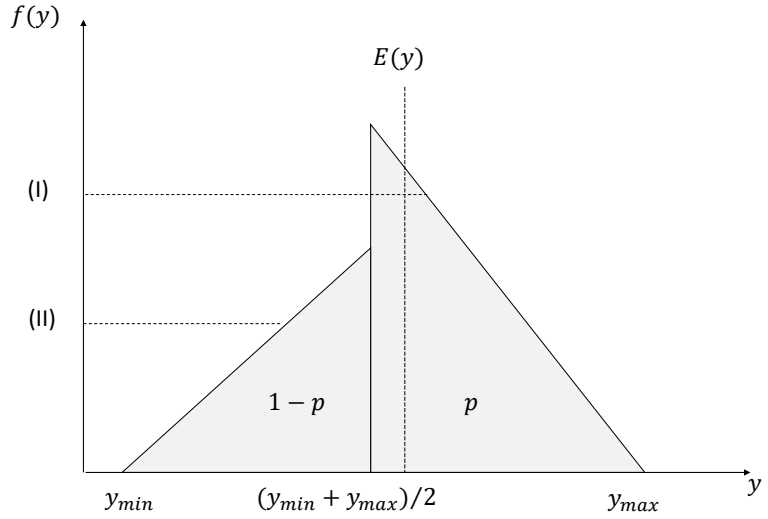
For the current analysis, however, it is more important whether response behavior differs across genders. As seen in Table A.1, women who provide information on their wage expectations are more likely to have higher cognitive abilities, are more likely to intend to enroll in college, to prefer a STEM or teaching major, and have higher confidence and a better academic self-efficacy than women who did not answer the module. Among male students the pattern is similar. Those who provided information on their wage expectations have a better final high school GPA, are more likely to intend to enroll in college, and more likely to have a better academic self-efficacy than men who did not answer the module. However, men who end up in our final sample are much less neurotic than men who show non-response. A pattern which is absent for women. Overall, it seems that students who provided information on their wage expectations and all covariates are somewhat positively selected in terms of their cognitive and noncognitive abilities, as well as educational aspirations. Nevertheless, selection patterns across genders appear similar overall.

B Expected earnings distribution

In order to calculate moments of the individual wage distribution, it is necessary to determine how expected wages are distributed over the two intervals (from the minimum (y_{min}) to the midpoint (y_{mid}) and from the midpoint to the maximum (y_{max})). In this study, we follow Guiso et al. (2002) and Attanasio and Kaufmann (2014, 2017), assuming a triangular distribution, which gives expected wages closer to the midpoint more weight than expected wages further away from that point. Based on these three pieces of information on the individual wage distribution (y_{min}, y_{max}, p) and the distributional assumption we calculate average expected wages $E(y)$ for each student i and education scenario $d = 1, 2, 3$ (i.e. vocational, bachelor's or master's degree). Figure B.1 shows the underlying triangular distribution.

The underlying probability distribution function is given by:

Figure B.1: Triangular distribution of expected wages.



Notes: The triangular distribution of earnings, adapted from Guiso et al. (2002).

$$f(x) = \begin{cases} \frac{8p(y_{max}-y)}{(y_{max}-y)}, & \text{if } y \geq y_{mid} \\ \frac{8(1-p)(y-y_{min})}{(y_{max}-y)}, & \text{otherwise.} \end{cases} \tag{6}$$

Expected average earnings are thus given by:

$$E(y) = \frac{1-p}{3}(2 \cdot y_{min} + y_{max}) + \frac{p}{3}(y_{min} + 2 \cdot y_{max}). \tag{7}$$

C Gender gaps in expected earnings: Components and comparison with realized earnings

Table C.1 shows detailed descriptive statistics for the underlying components of our main outcome variable across different hypothetical education scenarios. We now also present descriptive statistics for wage expectations for a vocational degree, which we use as benchmark to compute expected returns to college education. As expected, average expected wages are higher for more advanced educational degrees for both men and women. In absolute measures, the gender gap is increasing in educational attainment and highest for expectations associated with a master's degree. In relative terms however, the gender gap is highest for bachelor's degrees.

Another finding is that the gender gap in average expected earnings is mostly driven by significantly higher expected maximum wages by men. This is especially the case for wage expectations associated with a master's degree. However, for vocational and bachelor's degrees, women also expect significantly lower minimum earnings. Roughly in line with this finding is that the gender gap in the coefficient of variation (as a unit free measure of variation) is highest for wage expectations with a master's degree. This

might reflect the fact that variation in actual wages is also largest among individuals with higher degrees and therefore also harder to predict for high school graduates.

Table C.1: Gender differences in wage expectations.

	Female	Male	Difference	... in %	SE
Expectations pooled over degrees:					
Expected earnings ($E(y)$)	3152.96	3740.45	-587.49***	15.7***	(127.72)
Minimum (y_{min})	2132.56	2331.57	-199.01***	8.5***	(61.07)
Maximum (y_{max})	4106.61	5106.46	-999.85***	19.6***	(223.59)
Standard deviation	379.15	538.21	-159.06***	29.6***	(40.51)
N	552	388			
N (Total)	940				
Expectations with a Bachelor's degree:					
Expected earnings ($E(y)$)	2524.75	3104.40	-579.65***	18.7***	(119.623)
Minimum (y_{min})	1805.77	2062.44	-256.67***	12.4***	(68.337)
Maximum (y_{max})	3241.91	4180.46	-938.55***	22.5***	(222.403)
Standard deviation	278.27	409.00	-130.73***	32.0***	(38.450)
N	278	197			
N (Total)	475				
Expectations with a Master's degree:					
Expected earnings ($E(y)$)	3790.34	4396.48	-606.14***	13.8***	(211.127)
Minimum (y_{min})	2464.12	2609.16	-145.04	5.6	(93.289)
Maximum (y_{max})	4983.94	6061.55	-1077.61***	17.8***	(372.352)
Standard deviation	481.50	671.48	-189.98***	28.3***	(70.330)
N	274	191			
N (Total)	465				
Expectations with a vocational degree:					
Expected earnings ($E(y)$)	1812.29	2082.66	-270.38***	13.0***	(61.829)
Minimum (y_{min})	1263.36	1413.96	-150.60***	10.7***	(47.347)
Maximum (y_{max})	2391.96	2766.67	-374.70***	13.5***	(105.690)
Standard deviation	216.04	262.81	-46.77**	17.8**	(18.564)
N	290	192			
N (Total)	482				

Notes: This table presents differences in wage expectations between men and women. Means and mean differences are based on a two-sided t-test. Due to item non-response the number of observations may vary. Note that pooled results have three times as many observations. Source: *Berliner-Studienberechtigten-Panel* wave 3. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

To calculate actual wages, we use earnings data from the German Microcensus (2010-2012) for full time employed individuals with 33 to 37 years of age to show how expected wages relate to realized earnings. Since at the time, the majority of degrees were single-tier degree types (e.g. Diploma), we use these degrees as alternative measure for earnings with a comparable master's degree. Appendix Table C.2 shows that wage expectations are very high compared to realized earnings at the time in Germany. Both genders consistently overestimate their earnings over different degrees. Surprisingly, women overestimate their future earnings more often and more strongly than men. It is thus unlikely that better

information of women about current earnings distributions can explain the gender gap in expected earnings.

Table C.2: Are females better informed?

	Females	Males	Difference	N
Panel A: Wage expectations (*Best Up*)				
Vocational	1,812	2,083	-271***	482
Bachelor	2,529	3,104	-580***	475
Master equivalent	3,790	4,397	-606***	465
Panel B: Population earnings (*Microcensus*)				
Vocational	1,760	2,278	-518	6,459
Bachelor	2,070	2,739	-669	472
Master	2,323	3,092	-769	498
Master equivalent	2,325	3,002	-677	11,787
Panel C: Share overestimating				
Vocational	44.8	32.3	12.5***	6,941
Bachelor	65.1	54.3	10.8**	947
Master	80.3	73.8	6.5*	963
Master equivalent	80.3	75.9	4.4	12,252
Panel D: Percentage deviations				
Vocational	28.3	23.2	5.1**	6,941
Bachelor	37.0	32.8	4.1	947
Master	69.6	52.0	17.6**	963
Master equivalent	69.5	55.0	14.5*	12,252

Notes: This table presents differences in actual and expected earnings (in EUR) between males and females. Differences in means are based on two-sided t-tests. Percentage deviations in Panel D are calculated as: $\Delta_i = (|y^e_{id} - \bar{y}^p_d|)/\bar{y}^p_d$, where y^e_{id} represent graduate i expected wage with education degree d and $\bar{y}^p_d$ gives actual average population wages with education degree d. Sources: *Berliner-Studienberechtigten-Panel* wave 3, German Microcensus (2010-2012). * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

D Full results: min, max, bachelor, master, quantiles and range

In the main analysis, we presented detailed decomposition results for expected average wages pooled over degrees. To better understand drivers of wage expectations, we now run an OB decomposition at the endpoints of the individual expectation's distribution and one for each degree type, separately. Table D.1 shows decomposition results for minimum and maximum expectations and Figure D.1 reports selected underlying coefficients. Table D.2 and Figure D.2 report the same for expectations with a bachelor's and a master's degree.

Table D.1: Detailed decomposition for expected minimum and maximum earnings (Pooled).

	Minimum earnings				Maximum earnings			
	(1) Explained		(2) Unexplained		(3) Explained		(4) Unexplained	
Intended college major:	0.016	(0.011)	-0.054*	(0.030)	0.020	(0.013)	0.013	(0.033)
Arts & Humanities	-0.001	(0.002)	-0.006	(0.010)	0.000	(0.002)	-0.018*	(0.010)
Social Sciences & Economics	0.001	(0.002)	-0.007	(0.005)	0.001	(0.004)	0.001	(0.005)
Business & Management	0.001	(0.002)	0.020**	(0.010)	-0.002	(0.003)	0.036***	(0.014)
STEM	0.019**	(0.008)	-0.011	(0.019)	0.020**	(0.009)	-0.013	(0.021)
Teaching	-0.006	(0.004)	0.000	(0.007)	-0.004	(0.004)	-0.000	(0.009)
Law	-0.000	(0.001)	0.005	(0.007)	0.001	(0.003)	-0.016**	(0.008)
Health & Medicine	0.000	(0.004)	-0.017	(0.013)	-0.001	(0.004)	0.016	(0.015)
Other	0.002	(0.003)	-0.002	(0.007)	-0.000	(0.003)	0.006	(0.010)
No college aspiration	-0.002	(0.003)	-0.041**	(0.018)	0.001	(0.002)	0.001	(0.022)
Missing information	0.002	(0.003)	0.003	(0.003)	0.004	(0.003)	0.001	(0.002)
Career motives:	0.007	(0.014)	-0.029	(0.062)	-0.000	(0.017)	-0.076	(0.080)
High income	0.003	(0.003)	-0.030	(0.028)	0.002	(0.003)	-0.039	(0.034)
Promotion possibilities	-0.000	(0.001)	0.015	(0.030)	-0.002	(0.004)	-0.015	(0.037)
Recognition	-0.004	(0.004)	-0.002	(0.022)	-0.005	(0.005)	0.014	(0.024)
Interesting job	0.008	(0.005)	0.065	(0.045)	0.002	(0.004)	0.052	(0.053)
Job security	-0.004	(0.004)	0.026	(0.038)	-0.005	(0.005)	-0.021	(0.043)
Independent working	-0.000	(0.003)	0.009	(0.021)	-0.000	(0.004)	0.007	(0.026)
Social interaction	-0.008	(0.007)	-0.016	(0.022)	-0.002	(0.008)	-0.040	(0.025)
Important for society	0.000	(0.001)	-0.005	(0.019)	-0.001	(0.002)	-0.016	(0.024)
Help Others	0.011	(0.008)	-0.035	(0.025)	-0.000	(0.009)	-0.012	(0.031)
Spare time	0.001	(0.002)	-0.019	(0.016)	0.001	(0.002)	0.004	(0.018)
Health/safety conditions	-0.006	(0.007)	-0.070*	(0.039)	0.004	(0.007)	-0.094**	(0.045)
Time for family	0.007	(0.005)	0.032	(0.029)	0.006	(0.005)	0.086**	(0.035)
Baseline characteristics	0.002	(0.008)	0.079	(0.084)	0.007	(0.009)	0.101	(0.092)
Cognitive abilities	0.007	(0.011)	0.004	(0.007)	0.003	(0.011)	-0.007	(0.008)
Noncognitive abilities	0.006	(0.010)	0.001	(0.006)	0.017	(0.011)	0.006	(0.010)
Subtotal	0.037	(0.024)	0.000	(0.102)	0.047*	(0.027)	0.037	(0.122)
Constant			0.074	(0.110)			0.121	(0.127)
Total gap	0.111***	(0.035)			0.206***	(0.041)		
N	940				940			

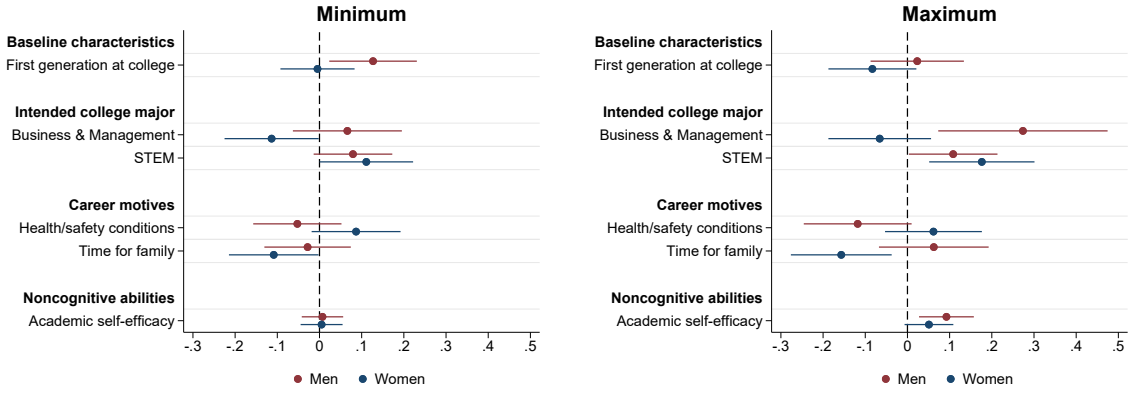
Notes: This table presents estimates of a detailed Oaxaca-Blinder decomposition using pooled coefficients as weighting scheme. Average expected minimum and maximum wages of high school graduates pooled over degrees serve as outcome variable. Columns (1) and (3) [(2) and (4)] show the contribution of each factor via the explained [unexplained] gap in log points. Joint contribution of factors in each category given by coefficients behind the categories name (in bold). Standard errors in parentheses allow for clustering at the individual level. Source: *Berliner-Studienberechtigten-Panel* waves 1-3. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

As a main result of Table D.1, we see that the relative size of the gender gap for maximum expected wages is with 0.206 log points much larger than for minimum expectations (0.111 log points), indicating that women anticipate discrimination especially in higher

career paths. Observed factors explain 33% of the gender gap in minimum expected earnings (0.037 of overall 0.111 log points). For maximum expected earnings, observed characteristics explain 41% of the total gap. Especially for minimum expected earnings, it is differences in endowments rather than differences in coefficients that explain the largest part of the gender gap in wage expectations. Turning to single contributions however, contributions via the unexplained part play a much larger role, but offset each other.

Overall, the role of most factors follows the same patterns as for average earnings. Looking at intended college majors, planned enrollment in STEM and Business & Management majors positively contributes to the gender gap for both minimum and maximum earnings via the explained and unexplained part, respectively. Also, the role of noncognitive abilities and different career motives is very similar for minimum and maximum expected earnings. However, there is one major exception. Expected family penalties have a smaller impact on minimum expected wages and are strongly associated with lower expected maximum wages. Having a very high preference for *time for family* contributes significantly via the unexplained part and is thus responsible for 41.7% of the gap in expected maximum earnings.[31] Given the fact that preferences for *time for family* "only" contribute 28.8% of the unexplained part for minimum expected earnings (not significant), indicates that women anticipated to give up especially higher career paths in order to take on family responsibilities.

Figure D.1: Selected coefficients for expected minimum and maximum earnings (Pooled).



Notes: This figure shows coefficient plots (95% confidence interval) for the effect of selected variables on the minimum expected earnings (y_{min}) and maximum expected earnings (y_{max}) pooled over degrees for men (red) and women (blue). Source: *Berliner-Studienberechtigten-Panel* waves 1-3.

Figure D.1 presents selected underlying subgroup coefficients. Here too, the patterns are broadly similar for minimum and maximum expected earnings. Overall however, differences in coefficients are more distinct for maximum expected earnings. Interestingly, academic self-efficacy which is an important driver of differences in average wage expecta-

[31] This corresponds to 0.417×815.85 EUR ≈ 340 EUR in absolute terms.

tions seems to operate mainly through maximum expected wages. Here, it has a positive and significant effect for both men and women.

The distinction between minimum and maximum expectations shows that determinants of the gender gap vary considerably across the distribution of wage expectations. So far however, we have only looked at wage expectations pooled over degrees. Focusing on single degree types helps to shed light on expectations for different career paths and occupations.

Table D.2: Detailed decomposition for expected average Bachelor and Master earnings.

	Bachelor earnings				Master earnings			
	(1) Explained		(2) Unexplained		(3) Explained		(4) Unexplained	
Intended college major:	0.012	(0.010)	-0.021	(0.029)	0.023	(0.014)	0.004	(0.034)
Arts & Humanities	-0.001	(0.002)	-0.009	(0.008)	0.001	(0.002)	-0.023**	(0.010)
Social Sciences & Economics	-0.000	(0.003)	-0.002	(0.005)	0.003	(0.005)	0.000	(0.005)
Business & Management	-0.001	(0.002)	0.029**	(0.012)	-0.001	(0.002)	0.033***	(0.012)
STEM	0.015**	(0.007)	-0.017	(0.018)	0.024***	(0.009)	-0.008	(0.021)
Teaching	-0.001	(0.003)	-0.002	(0.007)	-0.009*	(0.006)	0.000	(0.009)
Law	-0.000	(0.001)	-0.009	(0.007)	0.002	(0.005)	-0.014*	(0.008)
Health & Medicine	0.000	(0.004)	-0.003	(0.012)	-0.002	(0.005)	0.019	(0.017)
Other	0.000	(0.003)	0.004	(0.008)	0.000	(0.003)	0.003	(0.009)
No college aspiration	-0.000	(0.002)	-0.015	(0.020)	0.000	(0.002)	-0.006	(0.019)
Missing information	-0.000	(0.002)	0.002	(0.002)	0.006*	(0.004)	0.002	(0.002)
Career motives:	-0.002	(0.015)	-0.058	(0.072)	0.009	(0.017)	-0.059	(0.076)
High income	0.002	(0.003)	-0.034	(0.029)	0.003	(0.004)	-0.038	(0.034)
Promotion possibilities	-0.001	(0.002)	-0.020	(0.031)	-0.002	(0.004)	0.012	(0.037)
Recognition	-0.003	(0.003)	0.002	(0.022)	-0.005	(0.005)	0.018	(0.024)
Interesting job	0.001	(0.003)	0.060	(0.046)	0.007	(0.005)	0.053	(0.054)
Job security	-0.006	(0.005)	0.006	(0.038)	-0.004	(0.005)	-0.030	(0.043)
Independent working	-0.000	(0.004)	0.032	(0.023)	-0.001	(0.003)	-0.019	(0.027)
Social interaction	-0.005	(0.007)	-0.032	(0.022)	-0.002	(0.009)	-0.035	(0.025)
Important for society	-0.001	(0.002)	-0.008	(0.021)	-0.000	(0.001)	-0.019	(0.023)
Help Others	0.006	(0.008)	-0.023	(0.027)	0.001	(0.010)	-0.014	(0.030)
Spare time	0.001	(0.002)	-0.008	(0.015)	0.001	(0.002)	0.006	(0.017)
Health/safety conditions	0.001	(0.007)	-0.101***	(0.039)	0.002	(0.007)	-0.063	(0.045)
Time for family	0.005	(0.004)	0.067**	(0.030)	0.008	(0.006)	0.070**	(0.035)
Baseline characteristics	0.008	(0.008)	0.087	(0.080)	0.005	(0.007)	0.129	(0.092)
Cognitive abilities	0.002	(0.010)	-0.005	(0.007)	0.006	(0.011)	-0.003	(0.008)
Noncognitive abilities	0.014	(0.009)	0.006	(0.008)	0.016	(0.011)	0.002	(0.009)
Subtotal	0.035	(0.024)	0.008	(0.107)	0.059**	(0.027)	0.074	(0.120)
Constant			0.154	(0.112)			0.019	(0.129)
Total gap	0.197***	(0.036)			0.151***	(0.041)		
N	475				465			

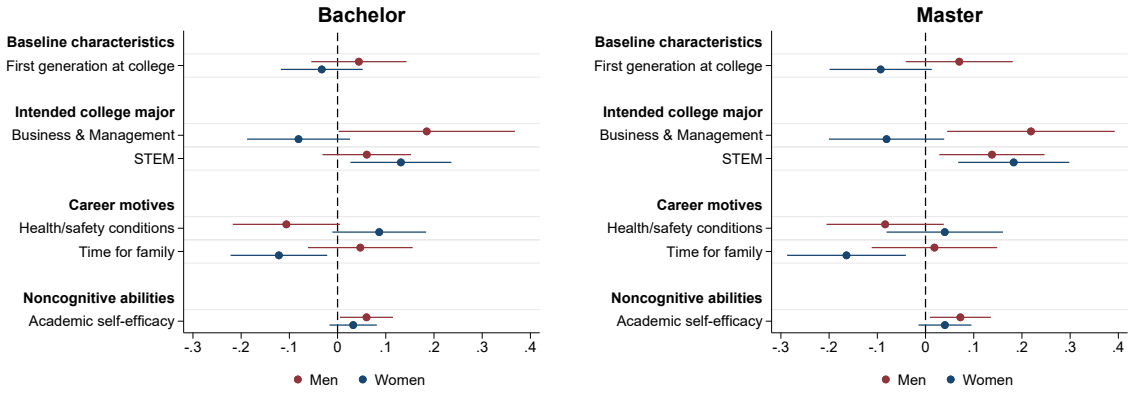
Notes: This table presents estimates of a detailed Oaxaca-Blinder decomposition using pooled coefficients as weighting scheme. Average expected minimum and maximum wages of high school graduates pooled over degrees serve as outcome variable. Columns (1) and (3) [(2) and (4)] show the contribution of each factor via the explained [unexplained] gap in log points. Joint contribution of factors in each category given by coefficients behind the categories name (in bold). Standard errors in parentheses allow for clustering at the individual level. Source: *Berliner-Studienberechtigten-Panel* waves 1-3. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table D.2 presents expectations for a bachelor's and a master's degree. Surprisingly, the relative size of the gap is larger for a bachelor's than for master's degrees (0.197 and 0.151 log points respectively). Overall, expectations for both degree types follow a similar pattern. For both jobs associated with a bachelor's and master's degree, different

intended college majors help to explain a significant part of the gender gap in expected wages. While more men enroll in on average better paying fields such as STEM, women hold clearly lower wage expectations for e.g. Business & Management fields. For a master's degree, also the effects of other fields of study, such as Arts & Humanities and Law, have significant (negative) effects that mitigate the overall contribution of field of study choice via the unexplained part.

The only career motive that significantly contributes to the gender gap in both wage expectations associated with both degree types is having a very high preference for *time for family*. However, even though coefficient effects of *time for family* are of similar magnitude (bachelor: 0.067, master: 0.070 log points) they are of larger relative importance for expectations for a master's degree (bachelor: 34%, master: 46%), supporting the notion that women anticipate family penalties especially via jobs in higher positions that a master's degree typically helps to qualify for.

Figure D.2: Selected coefficients for average expected earnings.



Notes: This figure shows coefficient plots (95% confidence interval) for the effect of selected variables on average expected earnings associated with a bachelor's and a master's degree for men (red) and women (blue). Source: *Berliner-Studienberechtigten-Panel* waves 1-3.

Of all other factors, only noncognitive ability measures have a joint effect on the size of the gender gap (in expected earnings with a bachelor's degree). Figure D.2 shows that this effect operates through differences in the academic self-efficacy as a single factor, which is associated with higher wage expectations by both men and women, but more prevalent for the latter (see Table 1). Overall, the subgroup coefficients are very similar for the two degree types. Negative coefficients of women who are potential first generation at college and have a high preference for *time for family* seem to be slightly more pronounced for wage expectations with a master's degree, but not statistically different from another across degrees.

Table D.3: Detailed decomposition for range of expected earnings (Pooled).

	(1) Explained		(2) Unexplained	
Baseline characteristics:	0.017	(0.016)	0.144	(0.159)
Information intervention school	-0.002	(0.004)	0.122*	(0.065)
Financial intervention school	0.002	(0.005)	0.034	(0.069)
Migration background	0.006	(0.007)	-0.063	(0.071)
First generation at college	0.001	(0.003)	0.029	(0.082)
Academic high school	-0.000	(0.004)	0.028	(0.031)
Integrated high school	0.002	(0.004)	-0.041	(0.040)
Vocational high school	-0.000	(0.002)	0.005	(0.034)
Fast track to vocational degree	0.010	(0.011)	0.014	(0.016)
Master	-0.002	(0.002)	0.016	(0.028)
Cognitive abilities:	0.005	(0.019)	-0.019	(0.014)
Final high school GPA	-0.001	(0.005)	-0.003	(0.003)
Verbal cognitive skills	0.005	(0.017)	-0.016	(0.013)
Figural cognitive skills	0.002	(0.004)	0.000	(0.003)
Intended college major:	0.021	(0.025)	0.101*	(0.060)
Arts & Humanities	0.002	(0.004)	-0.034*	(0.019)
Social Sciences & Economics	0.001	(0.005)	0.008	(0.008)
Business & Management	-0.008	(0.007)	0.056**	(0.025)
STEM	0.023	(0.015)	-0.000	(0.042)
Teaching	-0.001	(0.006)	0.005	(0.017)
Law	0.003	(0.007)	-0.032**	(0.013)
Health & Medicine	-0.006	(0.009)	0.044*	(0.025)
Other	-0.005	(0.008)	0.029	(0.021)
No college aspiration	0.003	(0.005)	0.031	(0.037)
Missing information	0.009	(0.007)	-0.005	(0.005)
Career motives:	-0.014	(0.029)	-0.197	(0.140)
High income	0.003	(0.006)	-0.041	(0.062)
Promotion possibilities	-0.003	(0.006)	-0.050	(0.064)
Recognition	-0.006	(0.007)	0.013	(0.042)
Interesting job	-0.007	(0.008)	-0.001	(0.093)
Job security	-0.003	(0.008)	-0.070	(0.075)
Independent working	-0.000	(0.004)	-0.016	(0.048)
Social interaction	0.003	(0.015)	-0.043	(0.043)
Important for society	-0.002	(0.007)	-0.020	(0.042)
Help Others	-0.012	(0.017)	-0.013	(0.051)
Spare time	-0.000	(0.003)	0.034	(0.035)
Health/safety conditions	0.009	(0.014)	-0.128	(0.082)
Time for family	0.003	(0.008)	0.138**	(0.061)
Noncognitive abilities:	0.033	(0.023)	0.010	(0.017)
Openness	-0.000	(0.002)	0.000	(0.001)
Extraversion	0.000	(0.004)	0.000	(0.003)
Conscientiousness	-0.001	(0.010)	-0.000	(0.001)
Neuroticism	0.000	(0.003)	0.002	(0.015)
Agreeableness	-0.000	(0.002)	0.000	(0.003)
Locus of control (int.)	0.013	(0.010)	0.003	(0.009)
Academic self-efficacy	0.033**	(0.016)	0.006	(0.007)
Self-confidence	-0.013	(0.012)	-0.001	(0.007)
Subtotal	0.062	(0.047)	0.038	(0.209)
Constant			0.245	(0.224)
Total gap	0.346***	(0.073)		
N	940			

Notes: This table presents estimates of a detailed Oaxaca-Blinder decomposition using pooled coefficients as weighting scheme. The outcome variable is individual expected range of wage expectations (minimum vs. maximum) of high school graduates pooled for education scenarios with a bachelor's and master's degree. Joint contribution of factors in each category given by coefficients behind the categories name (in bold). Standard errors allow for clustering at the individual level and presented in parentheses. Source: *Berliner-Studienberechtigten-Panel* waves 1-3. $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table D.4: Firpo decomposition of gender gap (pooled)

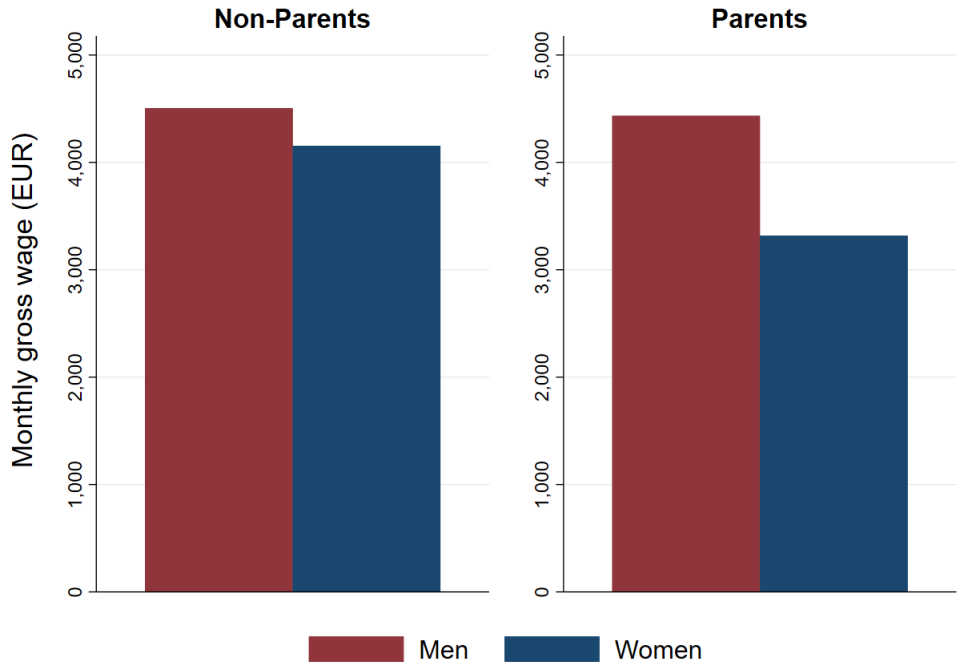
	q10	q25	q50	q75	q90
Male	7.650***	7.845***	8.091***	8.355***	8.675***
	(0.038)	(0.022)	(0.026)	(0.053)	(0.045)
Female	7.400***	7.652***	7.920***	8.230***	8.551***
	(0.029)	(0.025)	(0.031)	(0.024)	(0.051)
Gender Gap	0.250***	0.193***	0.171***	0.124**	0.124*
	(0.047)	(0.032)	(0.042)	(0.059)	(0.068)
Explained	0.041	0.063	0.046	0.120*	-0.007
	(0.056)	(0.047)	(0.047)	(0.062)	(0.083)
First generation at college	-0.001	-0.002	-0.000	0.000	0.001
	(0.006)	(0.007)	(0.003)	(0.004)	(0.005)
Academic self-efficacy	0.001	0.012	0.012	0.026	0.020
	(0.011)	(0.009)	(0.010)	(0.017)	(0.018)
STEM	0.006	0.007	0.024	0.033*	0.015
	(0.015)	(0.012)	(0.015)	(0.018)	(0.019)
Business & Management	-0.003	-0.005	-0.011	-0.010	-0.011
	(0.007)	(0.007)	(0.010)	(0.011)	(0.017)
Health/safety conditions	0.024	0.013	0.010	0.012	0.021
	(0.016)	(0.012)	(0.013)	(0.016)	(0.021)
Unexplained	0.208***	0.130**	0.125**	0.004	0.132
	(0.072)	(0.054)	(0.060)	(0.075)	(0.097)
First generation at college	0.071	0.133**	0.052	0.010	0.026
	(0.070)	(0.054)	(0.054)	(0.063)	(0.081)
Academic self-efficacy	0.000	-0.002	-0.000	-0.003	-0.002
	(0.006)	(0.005)	(0.005)	(0.007)	(0.008)
STEM	-0.011	-0.033*	-0.006	0.003	-0.005
	(0.022)	(0.017)	(0.018)	(0.022)	(0.028)
Business & Management	0.012	0.026	0.046**	0.042*	0.045
	(0.022)	(0.017)	(0.020)	(0.025)	(0.038)
Health/safety conditions	-0.098	-0.092	-0.075	-0.073	-0.087
	(0.072)	(0.058)	(0.063)	(0.074)	(0.099)
Time for family	0.065	0.045	0.086*	0.067	0.147*
	(0.059)	(0.042)	(0.047)	(0.056)	(0.079)

Notes: This table presents aggregate results of a Oaxaca-Blinder type RIF decomposition (Firpo et al., 2018). Men serve as reference group. The outcome variable is average expected wages of high school graduates pooled over education scenarios (bachelor's and master's degree). The decomposition additionally controls for all baseline characteristics, cognitive and noncognitive ability measures, intended college major and career motives. Coefficients behind *Explained* and *Unexplained* give joint contribution of all control variables through via the explained and unexplained part of the gap, respectively. Bootstrapped standard errors (400 replications) in parentheses. N=948. Source: *Berliner-Studienberechtigten-Panel* waves 1-3. p<0.1, ** p<0.05, *** p<0.01.

E Expected family penalties and actual child penalties

Preferences for *time for family* are the largest and most consistent factor that increases the gender gap (Figure 3). In case of childbirth, and under the assumption that *time for family* is invested primarily in raising children, the major share of this expected *family penalty* could be interpreted as an expected child penalty. The aim of this Section is to place these expected family penalties in the context of actual child penalty observed in the labor market around the age 35 conditional on working full time. Figure E.1 depicts monthly gross wages for full-time employed non-single men and women that hold at least a bachelor's degree equivalent (ISCED level 6 or higher) at ages 33-37 separately for parents (at least one child) and non-parents based on data from the German Socioeconomic Panel (SOEP).[32] Men in both scenarios (w/o and w/ children) experience no differences in earnings by parenthood (4,434 vs 4,505 EUR), while for women there is a large decrease under in the sample of parents compared with non-parents (4,155 vs 3,317 EUR).

Figure E.1: Gender-specific wages of parents and non-parents.



Notes: This figure shows monthly gross wages of non-single and men (red) and women (blue) between 33 and 37 years of age holding at least a bachelor's degree equivalent (ISCED level 6 or higher) and working full-time, separately for parents and non-parents. *Source*: SOEP, survey years 2010-2019, $n = 2,309$; authors' calculations.

If we assume that ever entering parenthood is related to the expressed career motive of having *time for family*, this pattern closely resembles what we see in wage expectations

[32] Individual wages are observed at multiple periods and thus for multiple ages. Parenthood is defined as ever having a child. Hence, individuals fall in either category (parent/non-parent) irrespective of their age and timing of childbirth.

of high school students. For men, family makes no difference in either real earnings or expectations, whereas for women there is a drop of around 838 EUR (or 20%) in actual earnings and a drop of around 15% in wage expectations (see Figure 2). Given that the difference in expectations reflect net rather than gross wages and already controls for some degree of sorting (e.g. into college major), these differences are potentially very similar.[33] Overall, this comparison suggests that the motherhood penalty, well known to be the largest single explanation for gender inequality in earnings, is already present in the expectations of high school students, i.e. before individuals typically enter parenthood.

What could explain such an early manifestation of the motherhood penalty? First, Germany constitutes a special case as it ranges among the European countries with the highest gender gaps (OECD, 2018) and the highest child penalties (Kleven et al., 2019). Most recently, the COVID-19 pandemic has shown that in (especially West) Germany classical gender roles still intend women to undertake the major share of care work (Jessen et al., 2021, Jessen, 2022). Under such gender norms, it is likely that women with child-bearing plans expect to bear most of the care work and thus give up positions in higher management or leadership positions. These gender norms, by definition, affect men and women differently and thus help in explaining why expected family penalties are perceived exclusively by women. As discussed earlier, the effect of preferences for *time for family* on wage expectations is indeed negative and significant for females, while male wage expectations are virtually not affected by such considerations (see Figure 2). As a result, the corresponding contribution to the gender gap in wage expectations is large and significant.

Second, it is well documented that women value flexibility in working hours more than men (Flabbi and Moro, 2012, Goldin, 2014, Bronson, 2015, Mas and Pallais, 2017, Wasserman, 2019). Since providing flexibility in working hours is costly to the firm, it is generally offered in exchange for lower pay, especially for high-end professionals (Goldin and Katz, 2011). Figure 3 shows that gender differences in coefficients and hence the absolute and relative share explained by expected family penalties are especially strong for expected maximum wages and for higher educational attainment. This is in line with the interpretation that, as a result to *anticipated* child birth or care responsibilities for other family members, women might give up career plans for positions in higher management or leadership positions in order to have a higher flexibility in working hours.

In the light of gender differences in beliefs associated with *time for family* commitments, it is still helpful to assess how these preferences are distributed among men and women. Table 1 shows that the share of women (43.8%) that hold a strong preference for this factor is 8.2 percentage points higher than for men (35.6%). However, the share contribution to the gender gap via differences in such *endowments* is small and insignifi-

[33] Further factors may explain differences either upwards or downwards such as the difference in probability of having entered parenthood by 35 in the expectations vs real data.

cant (Table 2).[34] Given the fact that coefficient effects are much larger than endowment effects and providing flexibility in working hours is costly to the firm, on the one hand lower wage expectations could represent expected discrimination from the demand side and reveal potential preferences for more flexible working hours (Blau and Kahn, 2017). On the other hand, these differences in wage expectations could be driven by the fact that women plan to select more often into *family-friendly* occupations.

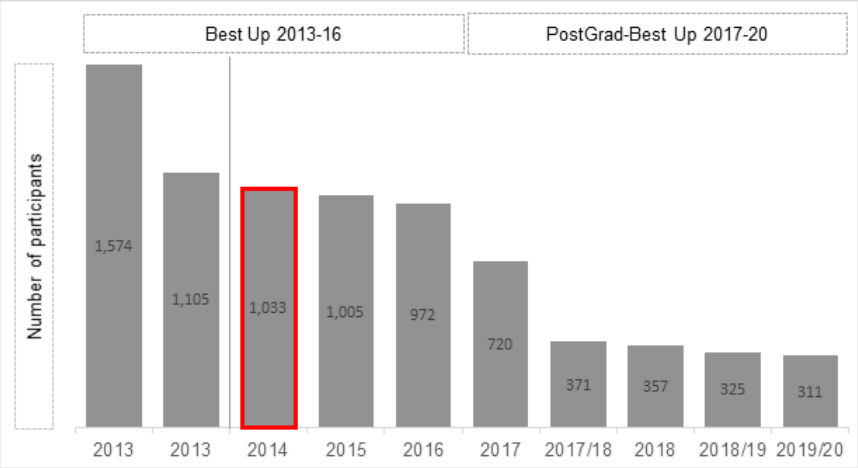
The notion that women are at least not fully aware of the costs of care responsibilities are roughly in line with the literature. Kuziemko et al. (2018) show that, in the US and the UK, women strongly underestimate the employment costs of motherhood. However, while the authors focus on an extensive margin – labor force attachment vs. women's perceived probability to be stay-at-home mothers – we add to this emerging strand of literature by focusing on an intensive margin: (expected) family penalties on wages full-time earnings. On this margin, women seem to hold somewhat more realistic expectations when compared to the reality.[35]

[34] Focusing solely on endowment effects is one reason why Kiessling et al. (2019) and Ehrmantraut et al. (2020) attribute a larger role in explaining the gender gap in expected wages to occupational sorting and negotiation styles. We argue that when one takes coefficient effects into account, the relative contribution of expected family penalties is large, which is more in line with previous literature. This holds especially true for wages associated with higher educational attainment and maximum expectations.

[35] Further results in Appendix Table C.2 show that surveyed wage expectations are more often higher than realized earnings with comparable degrees at the time. This is more often and more strongly the case for women, which could be interpreted as either (over-)optimism or misinformation.

F Other Figures

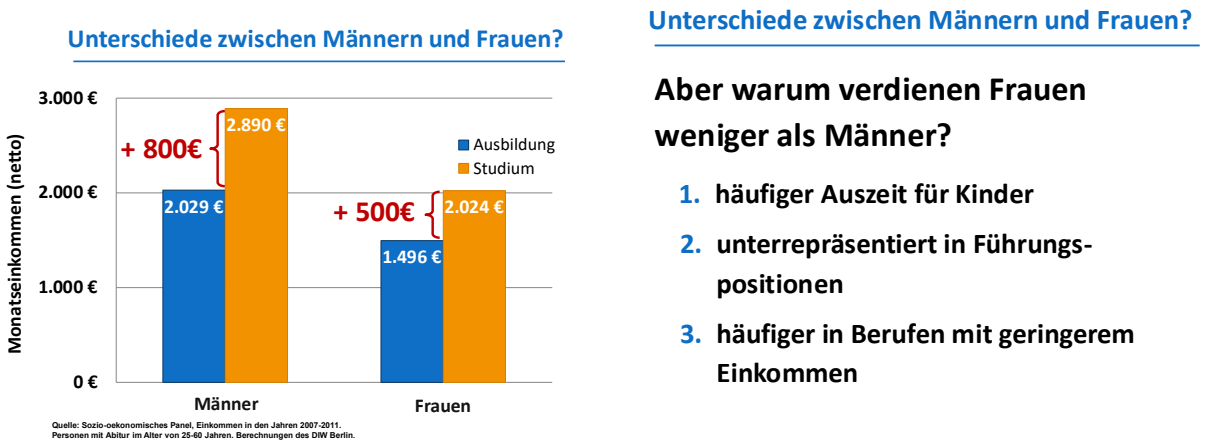
Figure F.1: Participants in Best Up and PostGrad-Best Up (2013-2020).



Notes: The red box indicates wave 3 that includes information on distributional parameters of wage expectations. The vertical line indicates high school leaving exams. Source: Own representation based on Zweck et al. (2019).

Figure F.2: Information slides on gender gap.

(a) (b)



Notes: This figure shows treatment slides on the gender gap as presented in the in-class information workshop on the returns to tertiary education. Slide (a) shows differences in monthly net earnings between jobs associated with a college degree (orange) and a vocational degree (blue) for men (*Männer*) and women (*Frauen*) separately. Slide (b) answers the questions "But why do women earn less than men?" (*Aber warum verdienen Frauen weniger als Männer?*) by giving three reasons: They are "more often on parental leave" (*1. häufiger Auszeit für Kinder*), "underrepresented in leadership positions" (*2. unterrepräsentiert in Führungspositionen*) and "more often in jobs with lower wage" (*3. häufiger in Berufen mit geringerem Einkommen*).