Check for updates

Gender differences in wage expectations: the role of biased beliefs

Stephanie Briel[1] · Aderonke Osikominu[1,2,3,4] · Gregor Pfeifer[3,4,5] ·
Mirjam Reutter[6] · Sascha Satlukal[1]

Abstract

We analyze gender differences in expected starting salaries along the wage expectations distribution of prospective university students in Germany, using elicited beliefs about both own salaries and salaries for average other students in the same field. Unconditional and conditional quantile regressions show 5–15% lower wage expectations for females. At all percentiles considered, the gender gap is more pronounced in the distribution of expected own salary than in the distribution of wages expected for average other students. Decomposition results show that biased beliefs about the own earnings potential relative to others and about average salaries play a major role in explaining the gender gap in wage expectations for oneself.

✉ Sascha Satlukal
sascha.satlukal@uni-hohenheim.de

[1] Department of Economics, University of Hohenheim, Schloss Hohenheim, Stuttgart, Germany

[2] CEPR, London, UK

[3] CESifo, Munich, Germany

[4] IZA, Bonn, Germany

[5] School of Economics, University of Sydney, Social Sciences Building, Camperdown, NSW 2006, Australia

[6] Department of Economics, European University Institute, Villa La Fonte, Fiesole, Italy

1 Introduction

Pay differences between men and women are one of the most intensively investigated phenomena in economics. Although the gender pay gap has been declining over the last decades in developed countries, women still earn considerably less than men (Blau and Kahn 2017). Gender differences do not only exist in realized wages, but they already emerge in expectations of young people about their future salaries (e.g., Filippin and Ichino 2005; Kiessling et al. 2019; Reuben et al. 2017). Hence, one reason why women do not earn as much as men might be lower wage expectations, as individuals tend to accept jobs that match their beliefs (Reuben et al. 2017).

In this paper, we focus on prospective university students' expectations about their own starting salary as well as that of an average graduate in the same field and with the same degree. We provide a comprehensive analysis of the structure of the gender gap along the entire distribution of expected salaries. In a first step, we estimate conditional and unconditional quantile regressions making use of a rich survey data set collected at Saarland University, Germany, that is roughly representative of the German student population. The conditional quantile regressions allow us to examine the gender gap at different percentiles of the distribution of expected salaries among students with the same field of study, intended degree, attitudes about income and other characteristics. The unconditional quantile regressions, in contrast, allow us to address the question how gender affects *ceteris paribus* the salary expectations of students at different percentiles of the unconditional distribution. In a second step, we decompose the gender gap at unconditional quantiles based on the method proposed by Firpo et al. (2009, 2018) in order to evaluate the contributions of different observed characteristics. Finally, we analyze to what extent gender differences in biased beliefs with regard to their own performance and the general labor market prospects contribute to the gender gap in the salary expected for oneself. To this end, we consider two different measures of biased beliefs. Firstly, we include indicators for over- and underplacement that relate students' own salary expectations to their expectations for average others while taking into account their relative performance (grade point average above/below median) in secondary school. Secondly, we analyze how the gender gap is affected by misperceptions of the average starting salary. For this, we compare students' expected average salary with the actual average starting salary in their field of study.

Our results indicate a raw gender gap at all percentiles of the distribution of expected salaries. At the median, we find raw gender gaps of 17 and 12 percent for the expected own and expected average salary, respectively. Accounting for observed characteristics reduces the gender gaps substantially, however, particularly the gap in the expected own salary remains sizable and highly significant. We find larger negative effects of being female at the bottom than at the top of the distribution, indicating a sticky floor structure of the gender gap in wage expectations. Furthermore, our decomposition results show that differential sorting of men and women into fields of study is strongly related to the gender gap in expected own and expected average salaries. However, after additionally accounting for biased beliefs, the contribution of field of study to

the explained gap in the expected own salary shrinks and becomes insignificant at most percentiles. Gender differences in our measures of over- and underplacement together contribute between 15 and 51 percent to the total gender gap. Even larger fractions of the total gender gap can be attributed to misperceptions of the average salary. Consequently, the unexplained part of the gender gap becomes insignificant after taking gender differences in biased beliefs into account.

Our paper contributes to three research areas: gender differences in pay, wage expectations and biased beliefs. First, our study speaks to the literature analyzing the gender gap in actual pay (see, e.g., the meta-analysis by Weichselbaumer and Winter-Ebmer 2005 for an overview). Many studies in this line of research document a U-shaped pattern of the gender pay gap along the conditional and unconditional distributions of earnings, meaning that the gender pay gap is particularly pronounced in the lower (sticky floor effect) and upper tail (glass ceiling effect) of wage distributions (Antonczyk et al. 2010; Christofides et al. 2013). Analyzing the gender wage gap in the entire working population, Collischon (2019) documents a glass ceiling effect in Germany. In contrast, Francesconi and Parey (2018), who exclusively focus on university graduates in Germany, a group closely related to our sample, find larger gender gaps in the lower tail of the wage distribution. Focusing on prospective students, we are able to document that the gender pay gap is already present at an earlier point in the careers of young people: when high school graduates form their expectations about their future salary.

Second, our study contributes to the literature on gender differences in earnings expectations that have been much less under study so far. However, the use of subjective expectations data in economics in general as well as in the context of education and labor economics has increased over the last years (see, e.g., Arcidiacono et al. 2020; Baker et al. 2018; Jacob and Wilder 2011; Reuben et al. 2017; Zafar 2013). Specifically, Kiessling et al. (2019)—who analyze wage expectations of advanced students in Germany—find significant gender gaps in the expected starting salary and expected lifetime earnings. Furthermore, their results suggest that a substantial part of the gap can be attributed to gender differences in sorting into fields of study as well as personal style in pay negotiations. Unlike Kiessling et al. (2019), we focus on the gender gap in expected starting salaries of prospective students, young individuals who are yet to enter university. In contrast to Kiessling et al. (2019), we study the gender gap also along the unconditional wage expectations distribution and compare the gaps in expected own salary and expected average salary. Moreover, our evidence highlights the role of gender differences in biased beliefs for explaining the gender gap in the expected own salary, rendering the contribution of differential sorting into field of study insignificant.

Third, our paper is related to studies analyzing gender differences in beliefs about individual and general economic outcomes. Students' beliefs about average starting salaries in their chosen field of study are related to the concept of optimism. In the economic literature, optimism is often defined as positive expectations about future events that are beyond the control of an individual (Jacobsen et al. 2014; Bjuggren and Elert 2019). Previous studies have shown that there exist gender differences in optimism in various economic domains. Bjuggren and Elert (2019), for instance, document that men are more optimistic than women about the future state of the economy.

In a similar vein, we compare the expected average starting salary between genders and show that male students are more optimistic about future wage levels in their field than their female counterparts.

A further related study is the work of Reuben et al. (2017), who analyze the impact of overconfidence as well as competitiveness and risk taking on earnings expectations of undergraduate students at New York University. Combining experimental and survey data, they show that a substantial part of the gender gap in wage expectations can be attributed to differences in overconfidence and competitiveness. Similarly, we study the role of biased beliefs that are presumably related to overconfidence. Specifically, we evaluate how students perceive their own earnings potential relative to other students and compare this to the relation between their own grade point average in secondary school and the average GPA of students in the same study field. Such a comparison is most closely related to the concept of over- and underplacement. Overplacement is classified as one form of overconfidence in the psychology literature and is defined as an exaggerated belief that one is doing better than others (Moore and Healy 2008). Thus, we denote our measures as over- and underplacement, respectively. Previous research documents that males are more likely than females to overplace themselves in some settings (e.g., Niederle and Vesterlund 2007; Ring et al. 2016).

Moreover, we also analyze gender differences with respect to misperceptions of average salaries. A number of studies have shown that university students are to some extent misinformed about average market salaries (Betts 1996; Jensen 2010; Wiswall and Zafar 2015). Our results suggest that biased beliefs play a major role in explaining gender differences in wage expectations. In fact, we do not find a statistically significant unexplained part of the gender gap after accounting for biased beliefs.

The remainder of the paper is organized as follows. Section 2 describes the data and shows descriptive statistics. Section 3 presents and discusses our estimation results of the conditional and unconditional quantile regressions, as well as of the decomposition of the gender gaps. Section 4 describes our measurements of over- and underplacement as well as of misperceptions of average salaries and discusses their impact on the gender gap in expected own salaries. Section 5 concludes.

2 Data and descriptive evidence

2.1 Data set

We use data from a survey of prospective students at Saarland University, Germany, who applied for enrollment in the academic years 2011 and 2012 (henceforth Student Survey). Prospective students submitting a complete application received an e-mail with the URL of the survey and were asked to fill in the questionnaire truthfully. In 2011, applicants for a program in Business Studies, Humanities, Law Studies, as well as Mathematics and Computer Science received the questionnaire; in 2012, also applicants for a program in Education and Medicine were surveyed. The sample consists of 2,061 students in total, which is a relatively large sample size compared with other studies examining students' wage expectations (e.g., Reuben et al. 2017; Webbink and Hartog 2004). The data were first used by Klößner and Pfeifer (2019),

who also validated the quality of the data.[1] As Klößner and Pfeifer (2019) document, Saarland University and its students are well representative of an average university and the corresponding average student body in Germany. Figures like student/teacher ratio, gender ratio, student age distribution, distribution of graduates across fields, distribution of gender across fields, number of exams passed, grades, duration of studies, etc., are all close to the German average. Saarland University also appears in the middle of international rankings across items such as teaching/learning, research, knowledge transfer, international orientation and regional engagement.

The prospective students first had to state the field of study and the degree (bachelor, master or state examination) they applied for. They were also asked whether they intended to do a further degree afterward (master, second state examination or doctoral degree) and with which degree they intended to earn their first salary. In the next part of the survey, the prospective students were asked to answer several questions regarding monthly gross salaries. First, they had to provide an estimate of the average starting salary of other persons graduating in the same field of study and with the same degree. Second, they had to state their own expected starting salary, again referring to their chosen field and intended degree. Figure 4 in "Appendix" shows an extract of the questionnaire containing these two core questions.

Furthermore, the students had to provide information about personal characteristics and family background. Specifically, we collected information on their gender; age; grade point average in secondary school; the type of secondary school they had graduated from; the federal state in which they had graduated from secondary school; whether their mother and father had a college degree and if so, their major discipline; whether they intended to live at their parents' home while studying; and whether they expected to receive the public financial aid 'BAfoeG'[2] and, if so, how much.

In the last part of the survey, students had to state the branch of business they intended to work in and their work experience in this branch. They also had to answer two questions regarding the value they placed on their future income: the importance of expected income for their choice of college major and the importance of receiving an above average salary.

2.2 Descriptive statistics

Table 1 depicts means and standard deviations of wage expectations and several other relevant variables for male and female students, respectively.[3] The last column of Table 1 also shows *t*-values for the test of equality of means between males and females.

Male students expect on average higher starting salaries both for themselves and for an average other person in the same field and with the same degree compared with

---

[1] Klößner and Pfeifer (2019) evaluate the students' knowledge of the German income tax system and focus on the difference between expected gross and net salaries. They document that prospective students tend to underestimate the progressiveness of the tax system and propose a correction for expected net salaries. Throughout our study, we use the corrected expectations of gross salaries.

[2] The Bundesausbildungsförderungsgesetz (BAfoeG, Federal Training Assistance Act) regulates student grants and loans in Germany that are granted to students with a relatively weak financial background.

[3] Figure 5 in "Appendix" shows the densities of wage expectations.

Table 1 Descriptive statistics

	Males	Females	t-Statistic
(a) Expected gross monthly salaries (Euro)			
Own starting salary	3579.12 (1908.16)	3014.89 (1558.45)	7.14
Average starting salary	3571.00 (1839.61)	3178.37 (1620.37)	5.02
(b) Personal characteristics			
Age (years)	21.22 (3.85)	20.78 (3.50)	2.61
GPA secondary school (scale from 1 to 4)	2.31 (0.60)	2.16 (0.59)	4.91
Relevant work experience (months)	4.36 (17.30)	4.23 (16.22)	0.17
(c) Field of study (percent)			
Business studies	29.77 (45.75)	19.61 (39.72)	5.25
Education	6.32 (24.35)	11.21 (31.56)	−3.95
Humanities	9.43 (29.23)	15.03 (35.75)	−3.90
Law studies	12.18 (32.73)	14.09 (34.81)	−1.27
Math./computer science	16.21 (36.87)	3.57 (18.55)	9.28
Medicine	18.85 (39.13)	29.03 (45.41)	−5.43
Natural sciences	7.24 (25.93)	7.47 (26.30)	−0.20
(d) Importance of expected income for choice of major (percent)			
Very strong	2.18 (14.62)	1.78 (13.24)	0.64
Strong	16.09 (0.37)	14.43 (35.16)	1.03
Neutral	37.93 (48.55)	37.01 (48.30)	0.42
Low	17.13 (37.70)	18.34 (38.71)	−0.71
Very low	26.67 (44.25)	28.44 (45.13)	−0.89
(e) Importance of an above average salary (percent)			
Very important	8.97 (28.59)	5.01 (21.82)	3.41
Important	42.30 (49.43)	35.14 (47.76)	3.28
Neutral	35.29 (47.81)	42.44 (49.45)	−3.00
Unimportant	8.51 (27.91)	11.88 (32.37)	−2.53
Very unimportant	4.94 (21.69)	5.52 (22.84)	−0.58
Observations	870	1178	

The first two columns show means and standard deviations (in parentheses). The last column contains t-values of a test of equality of means between males and females. In Germany, grades are scaled from 1 to 6, with 1 denoting the best grade and 4 the lowest passing grade. *Source* Student survey, own calculations

female students. More specifically, males expect on average to earn 3579€ a month at labor market entry, while the corresponding number for females is 3015€. With regard to average earnings of others in the same field of study and with the same degree, males state on average 3571€, which is significantly higher than the average female statement of 3178€. There is also a gender difference in the size of the expected own salary relative to expected average salary. Whereas the expected own salary is on average higher than the expected average salary for male students, the opposite is the case for female students.
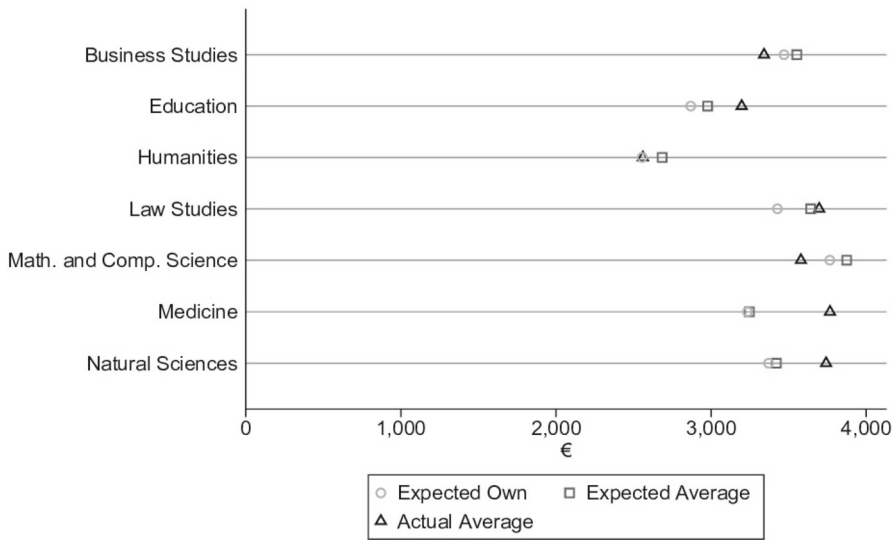
Fig. 1 Expected and actual starting salaries. *Notes:* Circles show mean values of expected own starting salaries. Squares show mean values of expected average starting salaries of students in the same field of study. Triangles show actual average starting salaries. *Source:* PersonalMarkt Services GmbH and Student Survey, own calculations

Moreover, men and women differ substantially with regard to the choice of field of study, as is also found by previous studies (e.g., Zafar 2013; Osikominu et al. 2020). Men in the sample are significantly more likely to apply for a program in Business Studies or in Mathematics and Computer Science. Contrarily, significantly larger shares of women apply for a program in Education, Humanities or Medicine.

There are also differences between male and female students in terms of the value they place on their income after graduation. Male students value an above average salary more than female students. Similarly, male students are more likely to state that future income prospects were important or very important for their choice of college major. This suggests that men are more strongly driven by pecuniary incentives than women who seem to place a higher value on non-pecuniary aspects of college majors and employment opportunities. Similar evidence has also been documented, among others, in Wiswall and Zafar (2018), Osikominu et al. (2020).

Figure 1 displays mean values of expected own starting salaries and expected average starting salaries based on the Student Survey as well as actual average starting salaries differentiated by field of study.[4] Wage expectations differ considerably across majors. While Mathematics and Computer Science students expect the highest starting salaries for themselves and for average others, Education and Humanities students expect the least. The comparison of expected starting salaries with actual starting salaries reveals that prospective students make fairly small estimation errors on aver-

---

[4] Data on actual starting salaries are from PersonalMarkt Services GmbH that offers the largest database of actual salaries for Germany.

age.[5] Figure 6 in "Appendix" shows that across all fields more men than women expect their own salary in excess of the actual average salary.

3 Analyses of the gender gap in wage expectations

3.1 Conditional and unconditional quantile regressions

Since gender gaps in actual salaries vary along the wage distribution (see, e.g., Collischon 2019; Francesconi and Parey 2018), we want to analyze the gender gap in students' expectations at different quantiles of the wage expectations distribution. Hence, we regress students' expected own salary as well as their expected salary for average others on a female dummy and observed characteristics using conditional quantile regressions (CQR). Specifically, we model the logarithm of the respective expected salary (own or average) of individual i, y_i, as:

$$y_i = \mathbf{x}_i \boldsymbol{\beta}_\tau + u_{\tau i} \quad \text{with} \quad Q_\tau(y_i \mid \mathbf{x}_i) = \mathbf{x}_i \boldsymbol{\beta}_\tau \,, \tag{1}$$

where $\mathbf{x}_i$ denotes the vector of explanatory variables, including a constant, and $\boldsymbol{\beta}_\tau$ the corresponding coefficient vector. As explanatory variables we use a female dummy as well as variables indicating the field of study, intended degree, degree applied for, family background, desired branch of business, work experience in desired branch, importance of income, grade point average in secondary school, type secondary school, region where high school was attended, age and survey year. The function $Q_\tau(y_i \mid \mathbf{x}_i)$ denotes the τ-th conditional quantile of the expected salary conditional on the explanatory variables, and $u_{\tau i}$ is an error term with $Q_\tau(u_{\tau i} \mid \mathbf{x}_i) = 0$ (see Koenker and Hallock 2001; Koenker 2005, for details). CQR allow us to evaluate the gender gap along the distribution of expected salaries among students with the same characteristics.

As the conditional quantile function is nonlinear the gender gap at the τ-th *conditional* quantile may differ from the gender gap at the τ-th *unconditional* quantile. Therefore, we also estimate unconditional quantile regressions (UQR), as proposed by Firpo et al. (2009), that allow us to evaluate the gender gap along the unconditional wage expectations distribution. The basic idea of UQR is to model the recentered influence function (RIF) of the τ-th unconditional quantile, Q_τ, of the logarithm of expected salary, y_i, which can be expressed as

$$\text{RIF}(y_i; Q_\tau) = Q_\tau + \frac{\tau - \mathbb{1}\{y_i \le Q_\tau\}}{f_y(Q_\tau)} \,, \tag{2}$$

where $f_y(Q_\tau)$ denotes the marginal density of y_i, and $\mathbb{1}\{\cdot\}$ is the indicator function (Firpo et al. 2009). Assuming that the conditional expectation function of the RIF of Q_τ is linear, we obtain a linear regression model

$$\text{RIF}(y_i; Q_\tau) = \mathbf{x}_i \boldsymbol{\gamma}_\tau + v_{\tau i} \,, \tag{3}$$

---

[5] Klößner and Pfeifer (2019), who use the same data set, analyze estimation errors of students in more detail and find an average estimation error of six percent.

Table 2 Gender gap in unconditional and conditional quantile regressions

	(1) 10%	(2) 25%	(3) 50%	(4) 75%	(5) 90%
(a) Log expected own salary					
RGG	−0.1383**	−0.1902***	−0.1673***	−0.1531***	−0.1751***
	(0.0696)	(0.0338)	(0.0227)	(0.0264)	(0.0441)
UQR	−0.1548**	−0.1204***	−0.1065***	−0.1111***	−0.0696*
	(0.0653)	(0.0372)	(0.0294)	(0.0283)	(0.0401)
CQR	−0.1468**	−0.1074***	−0.1205***	−0.1041***	−0.0879***
	(0.0542)	(0.0343)	(0.0266)	(0.0287)	(0.0341)
(b) Log expected average salary					
RGG	−0.1745**	−0.1420***	−0.1209***	−0.0996***	−0.0859**
	(0.0709)	(0.0335)	(0.0231)	(0.0290)	(0.0352)
UQR	−0.0936	−0.0876**	−0.0520*	−0.0465	−0.0582
	(0.0769)	(0.0344)	(0.0277)	(0.0284)	(0.0382)
CQR	−0.1060*	−0.0573*	−0.0830***	−0.0547*	−0.0330
	(0.0589)	(0.0313)	(0.0286)	(0.0280)	(0.0367)
Observations	2048	2048	2048	2048	2048

The dependent variable is the logarithm of the gross starting salary expected for oneself, panel (a), and for an average graduate with the same intended degree and field of study, panel (b). Rows labeled 'RGG' display the raw gender gap, computed as the difference between the respective quantile in the male and female distribution. Rows labeled 'UQR' display the coefficient on the female dummy in unconditional quantile regressions with control variables, whereas row labeled 'CQR' refer to conditional quantile regressions with control variables. In rows labeled 'UQR' and 'CQR,' we control for field of study, intended degree, degree applied for, family background, desired branch of business, work experience in desired branch, importance of income, grade point average in secondary school, type secondary school, region in Germany where high school was attended, age and survey year. Bootstrap standard errors (400 replications) are shown in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively. *Source* Student Survey, own calculations

where $\boldsymbol{\gamma}_\tau$ denotes the coefficient vector and $v_{\tau i}$ the error term with $\mathbb{E}[v_{\tau i} \mid \mathbf{x}] = 0$. Standard ordinary least square (OLS) regressions of Eq. (3) yield the UQR coefficient estimates that can be interpreted as partial effects on unconditional quantiles of y_i.

Our estimated gender gaps in wage expectations using conditional and unconditional quantile regressions are reported in Table 2. While Panel (a) displays the gender gaps in the expected own salary, Panel (b) shows the gender gaps in the expected salary for an average graduate in the same field of study and with the same degree. First and foremost, we find statistically significant raw gender gaps (RGG) in students' expected salaries for both, themselves and an average graduate, at all inspected quantiles of the distribution, as can be seen in the first row of Panels (a) and (b), respectively. Except for the first decile, the raw gender gap in students' expected own salary, reported in Panel (a), is always larger than the respective gap in their expected salary for an average graduate, reported in Panel (b). While the raw gender gap at the median in case of the expected own salary is 17 percent, the corresponding gap in case of the expected average salary is with 12 percent more than a quarter smaller. Interestingly,

the raw gender gaps we find in students' expected own salaries are similar in size to the gender gap in *actual* salaries of recent university graduates in Germany (Francesconi and Parey 2018).

In terms of the structure of the raw gap along the distribution, we also see differences between expected own and expected average salaries. Whereas the raw gender gap in expected average salaries is declining along the distribution, the raw gap in expected own salaries varies rather unsystematically. Thus, our results for the expected own salary do not completely match the findings of Kiessling et al. (2019), who report lower raw gender gaps at the top of the wage expectations distribution. Similarly, the raw gender gap in actual salaries of university graduates is declining with increasing wage levels. The raw gender wage gap in the whole working population in Germany, in contrast, is increasing along the distribution (Collischon 2019).

The second and third rows of both panels show gender gaps in wage expectations that are adjusted for gender differences in important observed characteristics (e.g., field of study, intended degree, GPA in secondary school, income importance). Consequently, the gender gaps estimated from the quantile regressions reflect gender differences in wage expectations that persist after taking these differences in observed characteristics into account. The two rows differ, however, in terms of the respective distribution being analyzed. The gender gaps displayed in the second row of both panels are estimated by unconditional quantile regressions (UQR) and, thus, can be interpreted as the adjusted gender differences in the unconditional distribution of wage expectations. The gender gaps displayed in the third row of both panels, in contrast, are estimated by conditional quantile regressions (CQR). Hence, these estimates reflect the gender gaps in the conditional wage expectations distributions.

Compared with the raw gender gaps, the adjusted gender gaps at the conditional and unconditional quantiles are, except at the first decile of the expected own salary, always smaller, suggesting that a non-negligible part of the raw gender gaps can be attributed to differences in observed characteristics between men and women. Nevertheless, the gender gaps in the expected own salary remain sizable and statistically significant at all studied quantiles after accounting for differences in observed characteristics. Analogously to the raw gender gaps, the adjusted gender gaps in the expected own salary are larger than the corresponding adjusted gaps in the expected average salary. In fact, the adjusted gender gaps at the bottom and at the top of the conditional and unconditional distribution of the expected average salary are at most weakly statistically significant. However, there are significant adjusted gender gaps in the middle of the expected average salary distribution. Thus, while we find that young women and men differ more in how they perceive their personal earnings prospects, we also see that male students are more optimistic about what a typical graduate from the same field of study can expect to earn. This finding is consistent with previous studies showing that women tend to be more pessimistic about future economic events (e.g., Jacobsen et al. 2014; Bjuggren and Elert 2019).

Overall, we do not find large differences between gender gaps in the conditional and unconditional distributions. Among students with the same characteristics, the gender gap is most pronounced among those who see themselves at the bottom of the conditional distribution of expected own salary, and the gap is the least pronounced at the top. Similarly, our unconditional quantile regressions also suggest a larger negative *cet.*

par. effect of being female at the bottom of the unconditional distribution of expected own salary. Thus, the pattern of the adjusted gaps at conditional and unconditional quantiles mimics a sticky floor effect.

Eventually, our adjusted gender gaps are similar in size and structure to the adjusted gaps found by Kiessling et al. (2019), who estimate the gaps by conditional quantile regressions. However, the adjusted gender gaps in students' own wage expectation do not completely mirror the adjusted gender gaps at conditional quantiles of actual salaries of recent university graduates found by Francesconi and Parey (2018). In contrast to the adjusted gender gaps in students' wage expectations, the adjusted gap in actual salaries of university graduates is a bit smaller and constant across conditional quantiles. This divergence in adjusted gaps between expected and realized earnings might be due to differences in the respective set of covariates used across the different studies.

3.2 Decomposition of gender gaps at unconditional quantiles

To decompose the gender gap at unconditional quantiles of expected salaries, we apply the decomposition method proposed by Firpo et al. (2009, 2018) that is based on the RIF. One advantage of this approach in comparison with other decomposition methods for quantiles (e.g., Machado and Mata 2005; Melly 2005) is that it allows a detailed decomposition of the composition effect. Hence, we can also quantify the contribution of gender differences in, for example, the choice of field of study to the gender gap along the unconditional distribution of expected salaries.

We model the group-specific RIFs of men ($g = 0$) and women ($g = 1$) adopting again a linear specification for the conditional expectation of the RIFs:[6]

$$\text{RIF}(y_{gi}; Q_{\tau g}) = \mathbf{x}_i \boldsymbol{\gamma}_{\tau,g} + v_{\tau gi} \qquad \text{for} \quad g = 0, 1 \, . \tag{4}$$

The vector of explanatory variables, $\mathbf{x}_i$, does not include a female dummy now. After estimating Eq. (4) by OLS for both groups, the overall gender difference at the τ-th quantile, $\hat{\Delta}_O^{\tau}$, can be decomposed as,

$$\hat{\Delta}_O^{\tau} = \underbrace{\bar{\mathbf{x}}_1(\hat{\boldsymbol{\gamma}}_{\tau,1} - \hat{\boldsymbol{\gamma}}_{\tau,0})}_{\hat{\Delta}_U^{\tau}} + \underbrace{(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_0)\hat{\boldsymbol{\gamma}}_{\tau,0}}_{\hat{\Delta}_E^{\tau}} \, , \tag{5}$$

where $\hat{\boldsymbol{\gamma}}_g$ denotes the respective OLS coefficient vector and $\bar{\mathbf{x}}_g$ the respective vector of sample means. Hence, the raw gender difference at the τ-th quantile is decomposed into the unexplained part, $\hat{\Delta}_U^{\tau}$, and the explained part, $\hat{\Delta}_E^{\tau}$.

If we want to interpret the explained and unexplained part of the decomposition truly as composition and wage structure effect, we need to assume that, at every quantile τ,

---

[6] In a sensitivity analysis, we verify that the linear specifications for the RIFs of the log of expected starting salaries are justified empirically using the reweighted regression approach suggested in Firpo et al. (2011). See "Appendix" C for details.

the error term $v_{\tau g}$ is mean independent of gender, G, i.e.,

$$\mathbb{E}[v_{\tau g} \mid \mathbf{x}, G] = \mathbb{E}[v_{\tau g} \mid \mathbf{x}] = 0\,, \quad \forall\, \tau \in [0, 1]\,, \ g = 0, 1\,. \tag{6}$$

This requires conditioning on all variables that are correlated with both the gender-specific outcome variables and gender. In the following analyses, we include a broad range of explanatory variables that we have selected based on economic reasoning and prior evidence on the determinants of expected starting salaries among prospective university students. Even though we do not have access to detailed controls for all underlying aspects of economic preferences and beliefs associated with both pay expectations and gender, to satisfy Assumption (6) it would suffice to condition on variables that can proxy the underlying facets well enough to achieve conditional independence of group-specific outcomes and group membership. As an example, while we cannot directly control for competitiveness, we know that students who choose a program in Education and indicate that pay is not important to them are similarly competitive and typically less competitive than students majoring in Business who state that pay is important to them.[7] Moreover, in Sect. 4, we consider a richer set of control variables that includes in addition measures of students' beliefs about their own earnings potential relative to others and the general earnings prospects in their field.

Table 3 contains the decomposition results at various unconditional quantiles of the expected own starting salary. While Panel (a) shows the aggregate decomposition results, detailed contributions of the most relevant covariates to the explained and unexplained part are given in Panels (b) and (c), respectively. According to Panel (a) and in line with the results in the previous section, a raw gender gap of 14 to 19 percentage points is present along the entire distribution of the expected own salary.[8] The part of the total gap that can be explained by gender differences in the covariates considered varies between six and ten percentage points in absolute terms and is always statistically significant. In relative terms, the explained part ranges between 36 percent at the median and 75 percent at the first decile. The unexplained part ranges between three and twelve percentage points and is statistically significant except at the tenth percentile.

Panel (b) of Table 3 shows the detailed contributions of selected variables to the explained part of the gender gap in the expected own starting salary. Among the covariates considered, gender differences in field of study explain the largest part of the gender gap in the expected own salary: Its contribution fluctuates between four and six percentage points across the quantiles considered and is statistically significant at all but the tenth percentile. In relative terms, differential sorting into field of study explains between 24 (25th quantile and ninth decile) and 41 (first decile) percent of the total gender gap. Gender differences in views on income importance explain up

---

[7] See also Dale and Krueger (2002) who apply a similar strategy to estimate the causal return to attending a more selective college in the USA.

[8] Note that the total gender gaps in Tables 3 and 4 differ somewhat from the corresponding raw gender gaps in Table 2. The reason for this is that the gaps in Tables 3 and 4 are computed as the difference between the gender-specific means of the respective RIF that are linear approximations of the unconditional quantiles of male and female expected salaries.

Table 3 RIF decomposition of the gender gap in expected own salary

	(1) 10%	(2) 25%	(3) 50%	(4) 75%	(5) 90%
(a) Aggregate decomposition					
Total gender gap	0.1354*	0.1904***	0.1949***	0.1597***	0.1720***
	(0.0727)	(0.0337)	(0.0270)	(0.0284)	(0.0433)
Explained part	0.1010**	0.0793***	0.0701***	0.0639***	0.0734***
	(0.0415)	(0.0239)	(0.0203)	(0.0192)	(0.0238)
Unexplained part	0.0344	0.1110***	0.1248***	0.0958***	0.0987**
	(0.0810)	(0.0413)	(0.0316)	(0.0328)	(0.0439)
(b) Contributions of selected covariates to the explained part					
Field of study	0.0556	0.0452**	0.0508***	0.0648***	0.0409**
	(0.0362)	(0.0217)	(0.0163)	(0.0157)	(0.0196)
Income importance	0.0235	0.0066	0.0040	0.0110*	0.0216**
	(0.0154)	(0.0082)	(0.0069)	(0.0064)	(0.0095)
(c) Contributions of selected covariates to the unexplained part					
Field of study	0.0851	−0.0026	−0.0264	0.0130	0.1828*
	(0.1777)	(0.0952)	(0.0740)	(0.0721)	(0.0985)
Income importance	−0.0204	0.0079	−0.0394	0.0491	0.1163
	(0.1150)	(0.0617)	(0.0511)	(0.0486)	(0.0739)
Observations	2048	2048	2048	2048	2048

The dependent variable is the log expected gross starting salary for oneself. Men are chosen as the reference group. We control for field of study (reference group: Business Studies), importance of income (reference group: neutral), intended degree, degree applied for, family background, desired branch of business, work experience in desired branch, grade point average in secondary school, type of secondary school, region in Germany where high school was attended, age and survey year. Bootstrap standard errors (400 replications) are shown in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively. *Source* Student Survey, own calculations

to two percentage points (seven percent) of the total gender gap. However, the effect is statistically significant only at the 75th and the 90th percentile. The contributions of other covariates (not reported in Table 3) are quantitatively less important and not statistically significant.

Table 4 shows the corresponding decomposition results at unconditional quantiles of the expected average starting salary. According to Panel (a), the total gender gap in the expected average salary is with nine to 14 percentage points again smaller than the one in the expected own salary at every quantile considered. Between five and seven percentage points can be explained by gender differences in the included covariates. In relative terms, this corresponds to 36 to 74 percent of the total gap, depending on the quantile considered. The unexplained part lies between two and nine percentage points. While the explained gap is significant at four out of the five quantiles considered, the unexplained gap is significant only at the 25th and 50th percentile. The results of the detailed decomposition of the explained part, shown in Panel (b), reveal that gender differences in field of study contribute the most to explaining the gender gap in expected

Table 4 RIF decomposition of the gender gap in expected average salary

	(1) 10%	(2) 25%	(3) 50%	(4) 75%	(5) 90%
(a) Aggregate decomposition					
Total gender gap	0.0905	0.1341***	0.1439***	0.1069***	0.0962***
	(0.0647)	(0.0317)	(0.0244)	(0.0287)	(0.0339)
Explained part	0.0673	0.0484**	0.0730***	0.0666***	0.0454**
	(0.0429)	(0.0212)	(0.0196)	(0.0193)	(0.0205)
Unexplained part	0.0233	0.0857**	0.0709**	0.0403	0.0508
	(0.0777)	(0.0377)	(0.0292)	(0.0317)	(0.0375)
(b) Contributions of selected covariates to the explained part					
Field of study	0.0387	0.0340*	0.0573***	0.0547***	0.0516***
	(0.0376)	(0.0181)	(0.0174)	(0.0153)	(0.0191)
Income importance	0.0209	0.0092	0.0042	0.0086	0.0031
	(0.0162)	(0.0076)	(0.0064)	(0.0067)	(0.0074)
(c) Contributions of selected covariates to the unexplained part					
Field of study	−0.0587	−0.0077	−0.0601	−0.0204	0.0151
	(0.1879)	(0.0891)	(0.0737)	(0.0760)	(0.0934)
Income importance	0.1004	0.0038	−0.0300	0.0090	−0.0000
	(0.1274)	(0.0578)	(0.0467)	(0.0530)	(0.0644)
Observations	2048	2048	2048	2048	2048

The dependent variable is the log expected gross starting salary for average others. Men are chosen as the reference group. We control for field of study (reference group: Business Studies), importance of income (reference group: neutral), intended degree, degree applied for, family background, desired branch of business, work experience in desired branch, grade point average in secondary school, type of secondary school, region in Germany where high school was attended, age and survey year. Bootstrap standard errors (400 replications) are shown in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively. *Source* Student Survey, own calculations

average salary. In relative terms, they explain between 25 and 54 percent of the total gender gap in the expected average starting salary, which is quantitatively similar to their contribution to the total gap in the expected own salary. Gender differences in views on income importance and other covariates (not reported in Table 4) do not contribute significantly to explaining the total gender gap.

All in all, the decomposition results match the patterns documented in the quantile regressions in Table 2: A substantial part of the raw gender gap in expected salaries can be attributed to differences in observed characteristics between men and women. Moreover, gender differences in field of study play the by far most important role for explaining the gender gap in expected salaries, which parallels findings in the literature on the gender gap in expected and actual pay (e.g., Collischon 2019; Kiessling et al. 2019; Fernandes et al. 2020).

4 The role of biased beliefs

While field of study or occupation have commonly been documented to be important drivers of the gender pay gap, they represent rather black-box controls for underlying aspects of preferences and beliefs that led to a particular choice of field of study or occupation. In the following, we want to open this black box by exploring to what extent gender differences in prospective students' salary expectations reflect gender differences in biased beliefs about their own relative earnings potential and the general earnings prospects in their field.

4.1 Measuring over- and underplacement

Our data allow us to compare the expectations of students about their own starting salary, conditional on field of study and intended degree, with their expectation about the average starting salary of other students in the same field of study and with the same intended degree. Specifically, in the survey, we first elicit a respondent's expectation about the average starting salary of graduates from the same field before we elicit the starting salary the respondent expects for herself (see Fig. 4 in "Appendix"). In so doing, we explicitly allow respondents to anchor the assessment of their own starting salary to their perception of the monetary success in the labor market of an average graduate with the same field of study and intended degree.

The exaggerated belief that one is doing better or worse than others is known as over- or underplacement in the psychology literature, and this terminology was adopted in economic research as well (Larrick et al. 2007; Moore and Healy 2008; Astebro et al. 2014; Duttle 2016). Individual-level differences between the two expected salaries, however, might not necessarily reflect mere over- or underplacement. Instead, they might also entail differences in abilities. In fact, somebody who expects to earn more than the average other graduate in the same program may do so because she obtained above average grades in secondary school. If all students were able to realistically asses their own abilities and based their wage expectations on them, individual-level differences between the two expected salaries would just reflect ability differences. To address this concern and to capture exaggerated misplacement, we take a prospective student's grade point average in secondary school into account when constructing our measurement of over- and underplacement. The GPA in secondary school constitutes a salient summary measure of a student's abilities because its value and the corresponding approximate rank are well known by the students themselves, and it is also an important criterion for admission to the different university programs in Germany. To be precise, we define a prospective student as overplacing themselves who expects to earn more than the average graduate in the same program but whose GPA in secondary school is below the median in their chosen program. Analogously, we define a student as underplacing themselves who expects to earn less than average, but has a GPA in secondary school above the median in their chosen program.

Table 5 displays the means of our over- and underplacement indicators in the full sample, as well as by gender. Female students are with 18 percent as opposed to 12 percent significantly more likely to underplace themselves with respect to their starting

Table 5 Over- and underplacement by gender

	Overall	Males	Females	t-Statistic
(a) Overplacement	0.092	0.136	0.060	5.218***
	(0.289)	(0.343)	(0.237)	
(b) Underplacement	0.157	0.120	0.184	−3.446***
	(0.364)	(0.325)	(0.387)	

Means and standard deviations (in parentheses) of dummies for over- and underplacement. The last column shows t-values of a test of equality of means between males and females. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively. *Source* Student Survey, own calculations
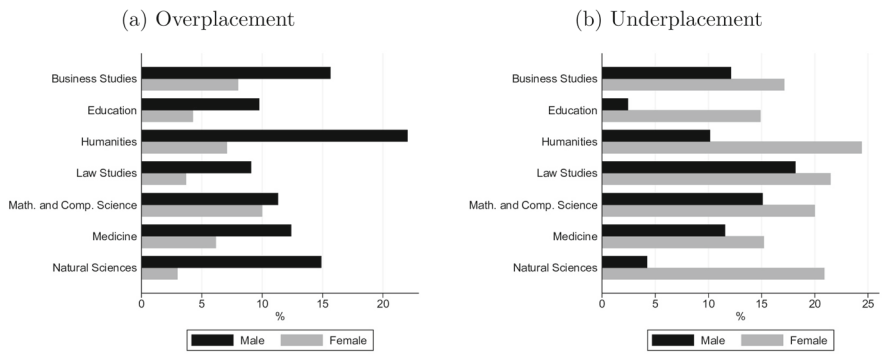
Fig. 2 Over- and underplacement by field of study and gender. *Notes:* **a** Share of students who expect to earn more than average and have a final grade in secondary school that is worse than the median grade in the chosen field of study. **b** Share of students who expect to earn less than average and have a final grade in secondary school that is better than the median grade in the chosen field of study. *Source:* Student Survey, own calculations

salary. Moreover, the share of male students overplacing themselves is with almost 14 percent more than twice as large than the corresponding share of female students with 6 percent. Similar patterns emerge within all fields of study, as can be seen in Fig. 2. Within all fields, males are more likely to overplace themselves, while females are more likely to underplace themselves.

A potential concern with our measurement of over- and underplacement is that, besides ability, there could be other reasons why a student expects to earn more or less than other graduates from the same program earn on average. Prospective students may value financial aspects associated with their study choice differently or desire to work in different industries with differing pay systems. However, in our decomposition of the gender gap, we condition on a broad range of observed characteristics (e.g., income importance, desired branch of business). Hence, this potential issue is not problematic for our identification of the effect of over- and underplacement on the gender gap in expected salaries.

Even though we account for many observed characteristics, there might still be concerns that gender differences in unobserved determinants of wage expectations affect our measures of over- and underplacement. Given that females tend to work fewer hours during their lifetime, one might expect that they intend to work fewer

Table 6 Over- and underestimation of average salary by gender

	Overall	Males	Females	*t*-Statistic
(a) Overest. Average salary	0.345	0.408	0.299	5.116***
	(0.476)	(0.492)	(0.458)	
(b) Underest. Average salary	0.483	0.414	0.534	−5.425***
	(0.500)	(0.493)	(0.499)	

Means and standard deviations (in parentheses) of dummies for over- and underestimation of average salary. The last column shows *t*-values of a test of equality of means between males and females. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively. *Source:* Student Survey, own calculations

hours after graduation and, consequently, expect lower starting salaries. The previous literature, however, suggests otherwise. Neither is there a substantial gender difference in working hours of recent university graduates (Francesconi and Parey 2018), nor are female students less likely to expect to be in full-time employment directly after graduation (Fernandes et al. 2020). Moreover, previous research shows that non-cognitive skills and risk preferences do not contribute to the gender gap in expected starting salaries (Kiessling et al. 2019; Reuben et al. 2017). Therefore, we do not think that such factors have a major impact on the gender differences in our indicators for over- and underplacement with respect to starting salaries.

4.2 Measuring misperception of labor market prospects

Besides biased beliefs about the own salary in relation to the average salary, prospective students might also have misperceptions about the average starting salary in their chosen field of study. In Sect. 2.2, we have shown that students' wage expectations on average are fairly close to the actual salaries in most study fields. However, at the individual level, there might still be substantial differences between the expected average salary and the actual average salary. Hence, we compare students' expectations of the average starting salary with the actual starting salary in their field of study and analyze whether the incidence of biased beliefs about the general labor market prospects in the chosen field of study differs between male and female students. We construct measures of misperception of actual salaries that indicate whether a student over- or underestimates the average actual starting salary by more than ten percent, respectively.

Table 6 displays the means of the indicators for over- and underestimation of the average salary in the full sample, as well as by gender. Firstly, we see that a large number of prospective students either over- or underestimates the average salary by more than ten percent. Male students are with 41 percent as opposed to 30 percent of women significantly more likely to overestimate the average starting salary in their field of study. The share of female students underestimating the average starting salary in their field is with 53 percent twelve percentage points larger than the corresponding share of male students with 41 percent. Moreover, Fig. 3 displays the share of students who over- and underestimate the average salary in each field of study. In all fields apart
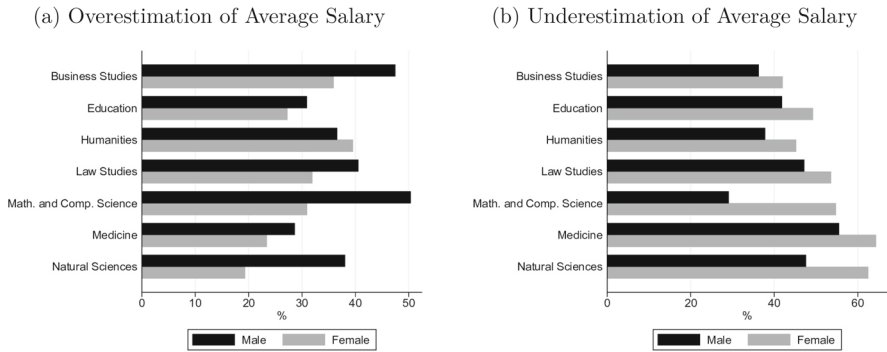
(a) Overestimation of Average Salary   (b) Underestimation of Average Salary

Fig. 3 Over- and underestimation by field of study and gender. *Notes:* **a** share of students who overestimate the average salary by more than ten percent. **b** share of students who underestimate the average salary by more than ten percent. *Source:* Student Survey, own calculations

from Humanities male students are more likely to overestimate the average salary than female students. Females, in contrast, are more likely to underestimate the average salary within all fields.

4.3 Effects on the gender gap in expected own salary

We have shown in Sect. 3 that the gender gap in the expected own salary is larger than the gap in the expected salary for average others. Further, the decomposition reveals that the unexplained part is particularly pronounced for the expected own salary. Thus, we will now study the role of biased beliefs in explaining the larger gap in the expected own salary. To be precise, we decompose the gender gap in the expected own salary along the unconditional distribution, additionally including the indicators for over- and underplacement as well as for over- and underestimation of the average salary as explanatory variables. Table 7 reports corresponding results. Since our measures of over- and underplacement are based on GPA in secondary school, we can only perform the analysis for prospective students who reported their GPA. Therefore, we lose 479 observations.[9]

In line with the complete sample, the gender gap in the expected own starting salary is present at all studied quantiles and varies rather unsystematically: 12 percent at the first decile, 18 percent at the median and 15 percent at the ninth decile. As can be seen in Panel (b) of Table 7, the contribution of overplacement is positive and significantly different from zero at every observed point along the wage distribution, with higher values in the tails of the distribution. Whereas gender differences in overplacement are responsible for 31 percent of the gap at the first decile and for 16 percent of the gap at the ninth decile, they are only responsible for 10 percent of the gap at the median. In contrast to overplacement, the impact of underplacement is significant only at the center of the distribution. At the median and the 75th quantile, the contribution of underplacement is a bit more pronounced than the contribution of overplacement.

---

[9] Baseline results of the conditional and unconditional quantile regressions similar to Table 2 but excluding students with missing GPA are shown in Table 8 in "Appendix."

Table 7 RIF decomposition of the gender gap in expected own salary

	(1) 10%	(2) 25%	(3) 50%	(4) 75%	(5) 90%
(a) Aggregate decomposition					
Total gender gap	0.1205	0.1735***	0.1791***	0.1527***	0.1530***
	(0.0735)	(0.0392)	(0.0313)	(0.0350)	(0.0530)
Explained part	0.2103***	0.1832***	0.1597***	0.1182***	0.1017***
	(0.0673)	(0.0392)	(0.0290)	(0.0221)	(0.0311)
Unexplained part	−0.0898	−0.0097	0.0194	0.0344	0.0512
	(0.0918)	(0.0473)	(0.0293)	(0.0300)	(0.0447)
(b) Contributions of selected covariates to the explained part					
Overplacement	0.0372***	0.0217***	0.0180***	0.0113**	0.0243**
	(0.0134)	(0.0073)	(0.0054)	(0.0047)	(0.0098)
Underplacement	0.0236	0.0171**	0.0211***	0.0116**	0.0075
	(0.0146)	(0.0073)	(0.0071)	(0.0049)	(0.0058)
Overest. Average salary	0.0022	0.0110*	0.0418***	0.0754***	0.0688***
	(0.094)	(0.0066)	(0.0121)	(0.0145)	(0.0157)
Underest. Average salary	0.1395***	0.1082***	0.0751***	0.0084*	0.0047
	(0.0389)	(0.0262)	(0.0150)	(0.0047)	(0.0054)
Field of study	−0.0062	0.0123	0.0215	0.0442***	0.0259
	(0.0446)	(0.0229)	(0.0145)	(0.0133)	(0.0207)
Income importance	0.0065	−0.0015	−0.0067	0.0047	0.0015
	(0.0199)	(0.0088)	(0.0054)	(0.0055)	(0.0088)
(c) Contributions of selected covariates to the unexplained part					
Overplacement	0.0126	−0.0039	−0.0017	−0.0086	−0.0008
	(0.0145)	(0.0072)	(0.0051)	(0.0058)	(0.0099)
Underplacement	−0.0065	0.0027	−0.0171	0.0074	0.0047
	(0.0434)	(0.0203)	(0.0121)	(0.0124)	(0.0182)
Overest. Average salary	−0.0186	0.0124	0.0474**	−0.0176	−0.0741**
	(0.0315)	(0.0196)	(0.0206)	(0.0312)	(0.0367)
Underest. Average salary	−0.1994*	−0.0717	0.0751	0.0838**	0.0171
	(0.1147)	(0.0664)	(0.0471)	(0.0325)	(0.0295)
Field of study	0.1818	−0.0154	−0.0485	−0.0324	0.1617
	(0.2001)	(0.1023)	(0.0633)	(0.0634)	(0.0995)
Income importance	0.0527	0.0198	−0.0252	0.0647	0.0332
	(0.1342)	(0.0619)	(0.0376)	(0.0401)	(0.0688)
Observations	1564	1564	1564	1564	1564

The dependent variable is the log expected gross starting salary for oneself. Men are the reference group. We control for over-/underplacement (reference group: no misplacement), misperception of average salary (reference group: no misperception), field of study (reference group: Business Studies), importance of income (reference group: neutral), intended degree, degree applied for, family background, desired branch of business, work experience in desired branch, GPA in secondary school, type of secondary school system, region in Germany where high school was attended, age and survey year. Bootstrap standard errors (400 replications) shown in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively. *Source* Student Survey, own calculations

Thus, while gender differences in overplacement are more important at the bottom and top of the distribution, differences in underplacement play a more decisive role at the center. Overall, our findings are in line with Reuben et al. (2017) who document that gender differences in overconfidence are partly responsible for higher expected salaries of male students.

Furthermore, our results show that misperceptions of the average salary significantly affect the gender gap in the expected own salary. The contribution of gender differences in the propensity to overestimate the average salary is sizable and highly significant at the median and the upper tail of the distribution. Up to 49 percent of the gender gap can be attributed to overestimation of the average salary. Analogously, the contribution of underestimation of the average salary is sizable and highly significant at the lower tail of the distribution up to the median. At the first decile, the contribution of underestimation of the average salary even exceeds the total gender gap. This suggests that, while the stronger tendency of males to overestimate the average salary explains a substantial part of the gender gap in the upper half of the distribution, the stronger tendency of females to underestimate the average salary is responsible for a substantial fraction of the gap in the lower half. After including our indicators of biased beliefs, the contribution of field of study to the explained part is only statistically significant at the 75th quantile, while the contribution of income importance to the explained part turns insignificant along the entire distribution. This indicates that these variables are correlated with biased beliefs and have partly absorbed their influence on the total gender gap in our previous decomposition analysis in Sect. 3.2. In Figs. 2 and 3 , it can be seen that the incidence of over- and underplacement as well as over- and underestimation of the average salary varies substantially between fields. In line with this argument, other studies that do not account for biased beliefs find significant contributions of differential sorting of male and female students into field of study to the gender gap in wage expectations (Kiessling et al. 2019; Fernandes et al. 2020). Moreover, the unexplained part of the gender gap we find is no longer statistically different from zero at all quantiles considered. This result is in contrast to previous studies examining students' expected salaries, reporting larger and significant unexplained parts of the gender gap (Reuben et al. 2017; Kiessling et al. 2019; Fernandes et al. 2020). Thus, gender differences in biased beliefs seem to be important drivers of the gender gap in wage expectations and should be taken into account.

5 Conclusion

Based on large and informative survey data on prospective students at Saarland University, Germany, we analyze the gender gap along the conditional and unconditional distribution of students' expected salaries both for themselves and for an average graduate in the same field and with the same degree. In addition, we construct indicators of biased beliefs and study their impact on the gender gap.

Our results indicate that there are substantial differences in the distribution of expected salaries between women and men. At the median, the raw gender gap amounts to 17 percent in case of the expected own salary and 12 percent in case of the expect average salary. The size of the raw gender gap we find for the expected own

salary is comparable to the raw gender gap in the actual salary of university graduates (Francesconi and Parey 2018). We also find a raw gender gap in the expected average salary, suggesting that female students are less optimistic about general earnings prospects in their field of study than male students.

Furthermore, there remain sizable and significant gender gaps along the conditional and unconditional distribution of the expected own and average salaries after accounting for observed characteristics. Comparing the gender gaps at conditional quantiles with the corresponding gaps at unconditional quantiles, we do not find major differences. In both cases, our results suggest a larger negative effect of being female at the bottom than at the top of the distribution. Hence, the structure of the adjusted gender gap in wage expectations mimics a sticky floor effect. Moreover, our decomposition results show that an economically and statistically significant part of the gender gap in the expected own salary can be attributed to biased beliefs. Gender differences in over- and underplacement and misperceptions about average salaries are both important drivers of the total gap. On the one hand, gender differences in biased beliefs account for a substantial part of the explained gender gap. On the other hand, the unexplained part of the gender gap turns statistically insignificant after conditioning also on biased beliefs.

Our findings raise the question how inter-individual differences in beliefs about the own earnings potential relative to others as well as in beliefs about average salaries emerge, and whether and how they can be influenced by parents and educators. If such beliefs are malleable, policy interventions aimed at informing students about the labor market prospects associated with different fields of study and at making girls and young women in particular more confident about their future labor market prospects might help to reduce gender differences in pay.

Declarations

Appendix

A Survey questions eliciting salary expectations

Estimated Starting Salaries

In the following, please estimate different starting salaries as well as salaries after five years of work experience. Thereby, please always provide a monthly gross estimate using integer numbers (no commas, points, or spaces).

Note: All deductions, such as income tax, church tax, solidarity surcharge, contributions to the statutory pension insurance, health insurance, nursing care insurance, and unemployment insurance, are made from the gross salary. The remaining salary is called net salary, and is monthly paid out to the employee.

Please estimate the monthly average gross salary for other students who are majoring in the field of study for which you have applied for, assuming the degree with which you stated to earn your first salary.

At labor market entry
After five years of work experience

Please estimate your own monthly gross salary after your graduation.

At labor market entry
After five years of work experience

Next

Fig. 4 Survey questions to elicit expected starting salaries. *Notes:* Extract of the questionnaire the prospective students had to answer during the online survey, which focuses on estimated starting salaries

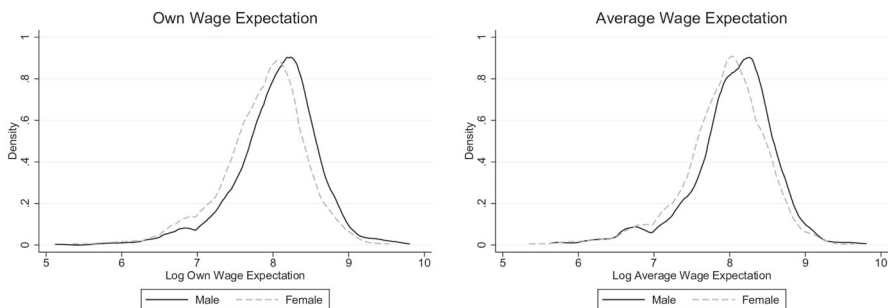## B Additional empirical evidence

Fig. 5 Densities of expected starting salaries. *Notes:* The left subfigure shows kernel density estimates of the logarithm of expected own starting salaries. The right subfigure shows kernel density estimates of the logarithm of expected average starting salaries. In both subfigures, we use an Epanechnikov kernel. *Source:* Student Survey, own calculations
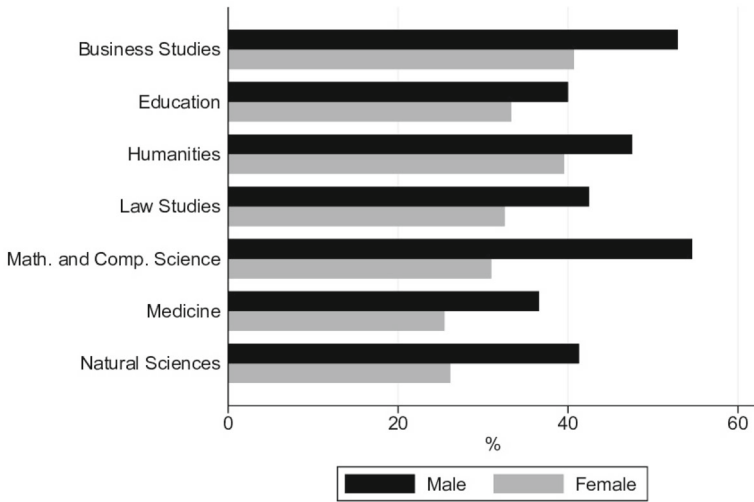
Fig. 6 Expected own starting salary above actual average starting salary. *Notes:* Share of students for which the expected own starting salary exceeds actual average starting salaries in their chosen field of study. *Source:* Student Survey and PersonalMarkt Services GmbH, own calculations

Table 8 Gender gap in unconditional and conditional quantile regressions (excluding observations with missing high school GPA)

	(1) 10%	(2) 25%	(3) 50%	(4) 75%	(5) 90%
(a) Log expected own salary					
RGG	−0.1081	−0.1793***	−0.1633***	−0.1766***	−0.1508***
	(0.0788)	(0.0407)	(0.0265)	(0.0332)	(0.0527)
UQR	−0.1462**	−0.1139***	−0.0989***	−0.1137***	−0.0635
	(0.0732)	(0.0437)	(0.0327)	(0.0326)	(0.0408)
CQR	−0.1098*	−0.1009**	−0.1264***	−0.1133***	−0.0616
	(0.0630)	(0.0425)	(0.0316)	(0.0347)	(0.0415)
(b) Log expected average salary					
RGG	−0.1036	−0.1366***	−0.1347***	−0.1151***	−0.0579
	(0.0870)	(0.0385)	(0.0265)	(0.0323)	(0.0386)
UQR	−0.0560	−0.0909**	−0.0572*	−0.0476	−0.0154
	(0.0852)	(0.0401)	(0.0306)	(0.0346)	(0.0425)
CQR	−0.0832	−0.0374	−0.0847***	−0.0649*	0.0020
	(0.0621)	(0.0356)	(0.0326)	(0.0367)	(0.0368)
Observations	1564	1564	1564	1564	1564

The dependent variable is the logarithm of the gross starting salary expected for oneself, panel (a), and for an average graduate with the same intended degree and field of study, panel (b). Rows labeled 'RGG' display the raw gender gaps, computed as the difference between the respective quantile in the male and female distribution. Rows labeled 'UQR' display the coefficient on the female dummy in unconditional quantile regressions with control variables, whereas row labeled 'CQR' refer to conditional quantile regressions with control variables. In rows labeled 'UQR' and 'CQR,' we control for field of study, intended degree, degree applied for, family background, desired branch of business, work experience in desired branch, importance of income, grade point average in secondary school, type of secondary school, region in Germany where high school was attended, age and survey year. Bootstrap standard errors (400 replications) are shown in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively. *Source:* Student Survey, own calculations

C Empirical support for the linearity assumption in the decomposition analyses

One potential concern w.r.t the decomposition of the unconditional quantiles is that consistent estimation of the explained and the unexplained part requires that the linearity assumption is satisfied in Eq. (4) (Barsky et al. 2002; Firpo et al. 2011, 2018). If the relationship is nonlinear, the τ-th quantile of the counterfactual distribution of log expected salaries that women would have, if they formed their expectations in the same way as men, is not equal to $\mathbb{E}[\mathbf{x} \mid G = 1] \cdot \boldsymbol{\gamma}_{\tau,0}$. Therefore, as an additional robustness check, we apply the reweighted regression approach described by Firpo et al. (2011). The idea of this approach is to reweigh the observations of male students so as to align the distribution of the characteristics of male students to that of female students. The reweighted regression approach allows us to estimate the specification error arising if the regression model is misspecified. A specification error close to zero indicates that the linear model is accurate. In addition, we use the method to calculate the reweighting error. If the reweighting function is consistently estimated, the reweighting error should be close to zero.

Applying the reweighted regression approach in the decomposition of unconditional quantiles, we find no statistically significant specification error at any percentile of the distribution of the expected starting salaries. Moreover, the reweighting error is always very close to zero and statistically insignificant, which indicates that the reweighting factors are consistently estimated.[10] This suggests that the linear specifications seem justified empirically.

References

Antonczyk D, Fitzenberger B, Sommerfeld K (2010) Rising wage inequality, the decline of collective bargaining, and the gender wage gap. Labour Econ 17(5):835–847

Arcidiacono P, Hotz VJ, Maurel A, Romano T (2020) Ex Ante returns and occupational choice. J Political Econ, forthcoming

Astebro T, Herz H, Nanda R, Weber RA (2014) Seeking the Roots of Entrepreneurship: insights from behavioral economics. J Econ Perspect 28(3):49–70

Baker R, Bettinger E, Jacob B, Marinescu I (2018) The effect of labor market information on community college students' major choice. Econ Educ Rev 65:18–30

Barsky R, Bound J, Charles KK, Lupton JP (2002) Accounting for the black–white wealth gap: a nonparametric approach. J Am Stat Assoc 97(459):663–673

Betts JR (1996) What do students know about wages? Evidence from a survey of undergraduates. J Human Res 31(1):27–56

Bjuggren CM, Elert N (2019) Gender differences in optimism. Appl Econ 51(47):5160–5173

Blau FD, Kahn LM (2017) The gender wage gap: extent, trends, and explanations. J Econ Lit 55(3):789–865

Christofides LN, Polycarpou A, Vrachimis K (2013) Gender wage gaps, 'sticky floors' and 'glass ceilings' in Europe. Labour Econ 21:86–102

Collischon M (2019) Is There a glass ceiling over Germany? Ger Econ Rev 20(4):e329–e359

Dale S, Krueger A (2002) Estimating the payoff to attending a more selective college: an application of selection on observables and unobservables. Q J Econ 117(3):1491–1527

Duttle K (2016) Cognitive skills and confidence: interrelations with overestimation, overplacement and overprecision. Bull Econ Res 68(S1):42–55

---

[10] The detailed results are available on request.

Fernandes A, Huber M, Vaccaro G (2020) Gender differences in wage expectations. Technical report arXiv:2003.11496

Filippin A, Ichino A (2005) Gender wage gap in expectations and realizations. Labour Econ 12(1):125–145

Firpo SP, Fortin NM, Lemieux T (2009) Unconditional quantile regressions. Econometrica 77(3):953–973

Firpo SP, Fortin NM, Lemieux T (2011) Decomposition methods in economics. In: Ashenfelter O, Card D (eds) Handbook of labor economics, vol 4. Elsevier, Amsterdam, pp 1–102

Firpo SP, Fortin NM, Lemieux T (2018) Decomposing wage distributions using recentered influence function regressions. Econometrics 6(2):28

Francesconi M, Parey M (2018) Early gender gaps among university graduates. Eur Econ Rev 109:63–82

Jacob BA, Wilder T (2011) Educational expectations and attainment. In: Duncan G, Murnane R (eds) Whither opportunity? Rising inequality, schools, and children's life chances. Russell Sage Press, New York, pp 133–162

Jacobsen B, Lee JB, Marquering W, Zhang CY (2014) Gender differences in optimism and asset allocation. J Econ Behav Org 107:630–651

Jensen R (2010) The (perceived) returns to education and the demand for schooling. Q J Econ 125(2):515–548

Kiessling L, Pinger P, Seegers P, Bergerhoff J (2019) Gender differences in wage expectations: sorting, children, and negotiation styles. CESifo Working Paper No. 7827, CESifo Munich

Klößner S, Pfeifer G (2019) The importance of tax adjustments when evaluating wage expectations. Scandinavian J Econ 121(2):578–605

Koenker R (2005) Quantile regression. Cambridge University Press, Cambridge

Koenker R, Hallock KF (2001) Quantile regression. J Econ Perspect 15(4):143–156

Larrick RP, Burson KA, Soll JB (2007) Social comparison and confidence: when thinking you're better than average predicts overconfidence (and when it does not). Org Behav Human Dec Process 102(1):76–94

Machado JAF, Mata J (2005) Counterfactual decomposition of changes in wage distributions using quantile regression. J Appl Econ 20(4):445–465

Melly B (2005) Decomposition of differences in distribution using quantile regression. Labour Econ 12(4):577–590

Moore DA, Healy PJ (2008) The trouble with overconfidence. Psychol Rev 115(2):502–517

Niederle M, Vesterlund L (2007) Do women shy away from competition? Do men compete too much? Q J Econ 122(3):1067–1101

Osikominu A, Grossmann V, Osterfeld M (2020) Sociocultural background and choice of STEM majors at University. Oxf Econ Pap 72:347–369

PersonalMarkt Services GmbH (no date). http://www.personalmarkt.de/de/. Accessed on 24 Dec 2012

Reuben E, Wiswall M, Zafar B (2017) Preferences and biases in educational choices and labour market expectations: shrinking the black box of gender. Econ J 127(604):2153–2186

Ring P, Neyse L, David-Barett T, Schmidt U (2016) Gender differences in performance predictions: evidence from the cognitive reflection test. Front Psychol 7:1680

Webbink D, Hartog J (2004) Can students predict starting salaries? Yes!. Econ Educ Rev 23(2):103–113

Weichselbaumer D, Winter-Ebmer R (2005) A meta-analysis of the international gender wage gap. J Econ Surv 19(3):479–511

Wiswall M, Zafar B (2015) How do college students respond to public information about earnings? J Human Cap 9(2):117–169

Wiswall M, Zafar B (2018) Preference for the workplace, investment in human capital, and gender. Q J Econ 133(1):457–507

Zafar B (2013) College major choice and the gender gap. J Human Res 48(3):545–595