

1742

Discussion
Papers

The Gender Gap in Wage Expectations: Do Young Women Trade off Higher Wages for Lower Wage Risk?

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

IMPRESSUM

© DIW Berlin, 2018

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
<http://www.diw.de>

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
<http://www.diw.de/discussionpapers>

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
<http://ideas.repec.org/s/diw/diwwpp.html>
<http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html>

The gender gap in wage expectations: Do young women trade off higher wages for lower wage risk?[☆]

Vaishali Zambre

DIW Berlin

June 18, 2018

Abstract

Several studies show that young women start with lower wage expectations than men, even before entering the labor market and that this partly translates into the actual gender wage gap through effects on educational choice and the formation of reservation wages. Building on the theoretical reasoning of compensating differentials proposing that the labor market compensates higher wage risk with higher wages, this study investigates whether the gender gap in wage expectations can be explained by individuals anticipating this form of risk compensation. Relying on a unique survey on German high school graduates in which we elicited information on the entire distribution of expected wages, this study documents that already at this early stage, female students expect to earn around 16% less than their male counterparts. At the same time, they expect lower wage risk as measured by the individual-specific dispersion in wage expectations. I decompose the gender gap into components attributable to socio-demographic factors, academic performance and skills, intended college major choice, career motives, personality traits, economic preferences, measures for students' confidence and expected wage risk. The results indicate that anticipated compensation for wage risk plays a major role in explaining the gender gap in wage expectations suggesting that females have lower wage expectations because they are willing to trade off higher wages for lower wage risk. The results of this study shed light on why young women make different choices regarding education and careers, thereby enhancing our understanding of the observed gender wage gap.

Keywords: wage expectations, gender gap, wage risk, risk compensation

JEL: I28, J18, D04

[☆]I thank C. Katharina Spieß, Sophia Schmitz, Ludovica Gambaro, Felix Weinhardt, and participants of the 2018 Annual Meeting of the American Economic Association for valuable comments and suggestions. I gratefully acknowledge funding from the Einstein Foundation Berlin (A-2010-025 (FU)). The usual disclaimer applies.

Email address: vzambre@diw.de (Vaishali Zambre)

1 Introduction

The gender wage gap remains a highly disputed topic in labor economics as well as in politics. Numerous articles investigate the extent, development, and potential explanations of this gap (for a recent overview see Blau and Kahn, 2016). In Germany the raw gender wage gap amounts to 22%. Compared to other European countries, the extent of the gender gap is remarkably high (e.g. Finke et al., 2017).

Several studies show that this gap is not only prevalent based on actual labor market earnings, but that females start with lower wage expectations even before entering the labor market (Blau and Ferber, 1991; Filippin and Ichino, 2005; Zafar, 2013; Reuben et al., 2015). These early gender differences in wage expectations can be particularly detrimental as they form the ground on which individuals base their decisions on. The gender gap in expected wages may translate into the actual gender wage gap through at least two channels. First, based on human capital theory, lower expected wages reduce the incentives to invest in education. Existing evidence – as well as a supplementary analysis in the Appendix – shows that expected wages are a significant predictor for the choice which level of education to pursue (Hartog et al., 2012; Schweri and Hartog, 2015; Belfield et al., 2016; Attanasio and Kaufmann, 2014, 2017) as well as for college major choice (Zafar, 2013; Arcidiacono et al., 2014; Wiswall and Zafar, 2016a; Ruder and Noy, 2017).¹ Second, lower wage expectations of females prior to labor market entry can become self-fulfilling. These expectations are likely to affect starting wages through the formation of reservation wages, which in turn have a persistent effect on future wage trajectories.² Females holding lower wage expectations might be more likely to accept lower wage offers and less likely to negotiate for higher wages because it matches their expectations. Hence, lower wage expectations may directly result in lower actual wages. In this regard, analyzing the reservation wages of unemployed individuals, Caliendo et al. (2017) show that gender differences in reservation wages can account for a large share of the subsequent gender gap in realized wages.

Thus, analyzing the determinants of gender differences in wage expectations may yield valuable insights on why women make different choices regarding education and careers, thereby enhancing our understanding of the gender gap in actual wages.³ Building on the

¹While not the focus of the current study, the analysis in Appendix B shows that students who expect higher returns to college are more likely to enroll in college. Moreover, students with higher wage expectations are more likely to choose a high paying college major and students who expect higher wage risk are more likely to choose a high wage risk college major.

²Evidence of the dampening effect of lower starting wages is provided by Oreopoulos et al. (2012). The authors show that entering the labor market during a recession has a long-lasting and partly permanent negative effect on wages.

³An additional advantage of analyzing gender differences in wage expectations instead of examining

theoretical reasoning of compensating differentials, the aim of this study is to examine whether the gender gap in wage expectations can partly be explained by anticipated compensation for wage risk, where wage risk is captured by the dispersion in expected wages. The basic idea is straightforward: It has been shown that individuals, who are more risk averse, are more likely to sort into occupations that exhibit lower wage risk, as measured by the wage variation (Bonin et al., 2007). Given that females are, on average, more risk averse than males (see e.g. Croson and Gneezy, 2009), this suggests that females sort into occupations with low wage risk, which – due to compensating differentials – pay less. The positive relationship between higher average wages and higher wage risk is evidenced multiple times based on actual labor market data (e.g. Hartog and Vijverberg, 2007). If individuals anticipate this form of risk compensation,⁴ this may explain why young women expect to earn less than men even before entering the labor market.

In addition, it is precisely the concept of compensating differentials around which many explanations for the actual gender wage gap are built. Gender differences with respect to preferences make some jobs more, and others less, attractive to female employees (Bertrand, 2011). For example, females may value flexibility in working hours more than males, and since providing flexibility in working hours is costly to the firm, it is only offered in exchange for lower pay. In this regard, while not the primary focus of their analysis, Wiswall and Zafar (2016b) confront students from New York University with multiple hypothetical job choice scenarios that varied in expected wages and other job characteristics. Their results show that, among others, females are willing to pay six times more in terms of expected wages for a higher flexibility in working hours or more secure jobs than males. More generally, it is assumed that females are willing to trade off higher wages for other appreciated job attributes. Wage risk, as measured by the dispersion in wages, may constitute one such job attribute.

If females do indeed deliberately choose to trade off higher wages for lower wage risk, we should be able to uncover this relationship based on individuals *ex ante* wage expectations, i.e. before they enter the labor market and self-select into different industries and occupations. Providing evidence that females not only have lower wage expectations but also expect lower wage risk, would strengthen the argument of conscious self-selection that is otherwise solely based on the idea of revealed preferences that assumes we can correctly identify preferences by only examining observed labor market outcomes.

The analysis draws on a unique survey of German high school graduates, in which

actual (realized) labor market earnings is that expectations abstract from labor market equilibrium effects and are not affected by unanticipated events that impact actual job choices.

⁴The question whether students anticipate compensation for wage risk has been addressed by Mazza and Hartog (2011) and Schweri et al. (2011); however, with contradicting results (see Section 3 for details).

we elicited information on the entire distribution of students' wage expectations. This allows me to construct a measure for wage risk based on the individual-specific dispersion of expected wages. As such, I focus on wage risk, as perceived by students when they are making their post-secondary educational choice. The results of this study show that females expect to earn considerably less than their male counterparts. Already at high school graduation, the gender gap in wage expectations amounts to around 16%. At the same time, young women expect lower wage risk. While students of both genders anticipate compensation for wage risk, i.e. expect a positive correlation between average expected wages and expected wage risk, the extent of expected risk compensation does not differ between genders. In order to analyze the role of anticipated compensation for wage risk in explaining the gender gap in wage expectations, I decompose the gender gap into components attributable to socio-demographic factors, academic performance and skills, intended college major choice, career motives, personality traits, economic preferences, measures for students' confidence, and expected wage risk. The results show that differences in expected wage risk between genders can explain around three quarters of the gender gap in expected wages. The decomposition results highlight that anticipated compensation for wage risk plays a major role in explaining the gender gap in wage expectations. Overall, this study sheds light on the self-selection process into different educations and careers and suggests that young women expect to earn less because they are willing to trade off higher wages for lower wage risk.

This study contributes to the existing literature on students' wage expectations (see Manski, 2004). The majority of studies in this strand of the literature focuses on expected average wages, i.e. ask students about a point estimate of their wage expectations.⁵ By contrast, the entire distribution of expected wages or other measures of expected wage risk are rarely elicited (Dominitz and Manski, 1996). Moreover, by not only analyzing average expected wages but broadening the analysis toward a perspective on gender differences while simultaneously considering the role of expected wage risk, constitutes a novel approach that has not been investigated previously. In addition, existing studies have mostly focused on wage expectations among college students; however, in order to

⁵Early studies in this strand of the literature focused on comparing students' expected wages to actual wages of individuals currently in the labor market (Betts, 1996; Wolter, 2000; Wolter and Zbinden, 2002; Huntington-Klein, 2015; Frick and Maihaus, 2016). These studies examine whether the assumption of reference group based expectations (see Manski (1993) for details), as generally made in educational choice models, is suitable. Relatedly, other studies investigate how accurate students' expectations are by comparing their expectations to actual realized wages for the same student, i.e. they follow students into the labor market (Webbink and Hartog, 2004; Jerrim, 2015). Additionally, some studies investigate the determinants of wage expectations and examine whether differences in expected wages reflect heterogeneity that is comparable to heterogeneity based on actual labor market data (see e.g. Brunello et al., 2004; Bonnard et al., 2014; Diaz-Serrano et al., 2016).

adequately examine the role of wage expectations for future choices it is essential to go one step further back and analyze wage expectations of students before they make their educational choice and thereby self-select into different fields and levels of education. Finally, and most importantly, several studies document a gender gap in wage expectations but the empirical evidence on the determinants of this gap is scarce. This study adds to the literature by exploring the role of anticipated compensation for wage risk in explaining the gender gap in wage expectations.

The study by Reuben et al. (2015), who investigate the role of preferences in explaining gender differences in wage expectations, is most closely related to my analysis. Based on a sample of New York University undergraduate students, their analysis documents 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 percent. The authors further show that gender differences in experimentally measured preferences, i.e. overconfidence, competitiveness and risk aversion, explain around 20 percent of the gender gap in expected wages. Another interesting aspect on the gender gap in wage expectations can be drawn from a study conducted by Filippin and Ichino (2005), who collected data on wage expectations from Bocconi University in Milan students as well as actual wages for Bocconi graduates. The authors show that the gender gap implied by students' expected wages is close to the gender gap based on actual wages of Bocconi graduates. This evidence underlines the close relation between expected and actual wages. Moreover, the analysis of Filippin and Ichino (2005) and Filippin (2003) suggest that part of the gender gap in wage expectations is likely due to females expecting to be discriminated in the labor market. When Filippin and Ichino (2005) ask students about the different explanations for the actual gender wage gap, females are more likely than males to attribute the gap to expected discrimination, while more males report differences in characteristics between gender as an explanation. Filippin (2003) derive a model, showing that this form of anticipated discrimination, even if it is entirely unsubstantiated, can result in actual wage differences. This result emphasizes the importance of investigating the determinants of gender differences in wage expectations.

The remainder of the paper is structured as follows. Section 2 introduces the data and provides detailed information on the measurement of wage expectations. In addition, this section presents descriptive evidence for gender differences in wage expectations and other observed characteristics. In Section 3, I provide evidence for the hypothesis that students anticipate compensation for wage risk. Section 4 examines the role of expected wage risk in explaining the gender gap in wage expectations and presents the results of the decomposition analysis. Moreover, this section investigates how sensitive the results

are to different specifications. It closes with a discussion of an alternative hypothesis that may explain why women expect lower wages and lower wage risk showing that this explanation is unlikely to apply. Finally, Section 5 concludes.

2 Data and descriptives

The empirical analysis is based on data from the *Berliner Studienberechtigten Panel (Best Up)* survey (Ehlert et al., 2017). This survey followed students in Berlin from the end of their penultimate year in high school through two years after graduating from high school. Students were surveyed five times over that period. The survey aimed at obtaining a sample of students who were pre-dominantly from non-academic family backgrounds. Thus, the selected schools are located in districts with a high share of low educated individuals (ISCED 0-2) and cover twenty percent of all upper secondary schools in Berlin. As such the sample is neither representative of the German nor Berlin student population. Except for the first survey, which was a paper and pencil survey conducted in schools, the subsequent surveys were administered on-line. Of the 1578 students surveyed in the first wave, 1105 participated in the second, 1034 in the third, 1005 in the fourth, and 972 in the fifth wave. The data includes detailed information on student characteristics, their educational aspirations, and their actual educational choices.⁶

The module on labor market expectations was included in wave three of the Best Up panel survey, i.e. just after students graduated from high school and were about to make their decision about their post-secondary education. Out of the 1034 students who participated in the third wave of the survey, a large share of students did not answer the question on their expected wage. Response rates are around 60% for each education scenario and slightly lower (56%) when considering only students with complete responses to all education scenarios. Existing studies on students' wage expectations primarily use data that was specifically collected to analyze wage expectations. As such, item non-response in these studies are lower (although sample selection is still a major concern). In contrast, this study builds on data where questions on wage expectations were included in a larger survey and not the primary focus of the questionnaire. Hence, higher item non-response rates are not surprising.⁷

⁶In the context of the data collection two interventions (one information and one financial intervention) were conducted. One concern might be that students may have adjusted their expected wages in response to the information intervention. However, I do not find that the information intervention had a significant effect on students' expected wages. Nevertheless, throughout this study I include an indicator variable accounting for the information intervention. For further details on the financial intervention see Peter et al. (2016a); and for the information intervention see Peter et al. (2016b) and Peter and Zambre (2017).

⁷The literature on reservation wages, for example, reports item non-response rates of similar magnitude if questions are included in surveys that do not exclusively focus on this issue (e.g. Caliendo et al., 2009).

However, comparing individual characteristics across students who answered the module on wage expectations for at least one education scenario with those who did not, suggests that item non-response in our questionnaire does not occur randomly (see Table A.1 in the Appendix). Students with a better high school GPA, students with higher scores on the verbal cognition test, and students who intend to enroll in college are statistically significantly more likely to provide information on their expected wages. For the current analysis, however, it is more important whether response behavior differs among genders. As seen in Table A.1, females who provided information on their wage expectations are more likely to have higher scores on the cognition tests, are more likely to intend to enroll in college and choose a STEM major, and are slightly more open than females 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, are less neurotic, less agreeable, more risk averse, and more likely to have confidence in their own abilities than males who did not answer the module. Overall, it seems that students who provided information on their wage expectations are slightly positively selected in terms of their academic performance and skills, as well as with regards to their educational aspirations. However, comparing females and males, selection patterns seem to be rather similar. Nevertheless, in Section 2 I investigate how a potential non-random response rate affects the gender gap in wage expectations using a selection corrected model as suggested by Heckman (1979).

2.1 *Measuring expectations about future wages*

In order to elicit information on students' expected wage distribution, the literature offers two main approaches. The first asks students about their expected *median* wage and some probabilities to earn more/less than X% of the median (Schweri and Hartog, 2015; Mazza and Hartog, 2011; Wolter, 2000). In contrast, the second approach asks about the minimum and maximum expected wage as well as additional information on the probability distribution (Guiso et al., 2002; Attanasio and Kaufmann, 2014). As it is disputable whether the concept of the *median* and the difference of it to the mean is fully understood by our target group, we implemented the second approach that is described in more detail below.⁸

For three different education scenarios, students are asked about the range of their individual wage distribution and the probability mass to the right of the midpoint of this range. More specifically, we asked students to state their net minimum (y_m) and net

⁸The results of a small pre-test of students of similar age as the target group suggested that most students do not perceive a difference between the mean and the median.

maximum (y_M) wage they expect to earn at the age of 35 conditional on working full time and having earned a) a vocational degree, b) a Bachelor's degree, and c) a Master's degree. We then asked students about the probability to earn more than the midpoint of their range, $p = Pr(Y > \frac{y_m + y_M}{2})$, where the midpoint was calculated by the computer.

The literature on wage risk typically measures risk by the dispersion around average wages within a specific education or occupation group. Following this approach, expected wage risk is measured as the variance of expected wages conditional on obtaining a particular educational degree.

Yet, 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_m) to the midpoint (y_{mid}) and from the midpoint (y_{mid}) to the maximum (y_M)). In this study, I 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.⁹

Based on these three pieces of information on the individual wage distribution (y_m, y_M, p) and the distributional assumption, average expected wages and expected wage risk for each student i and education scenario $d = 1, 2, 3$ (i.e. vocational, Bachelor's or Master's degree) can be calculated as:¹⁰

$$\begin{aligned} E(y) &= \int_{y_m}^{y_M} y f(y) dy \\ &= \int_{y_m}^{y_{mid}} (1-p) y f(y) dy + \int_{y_{mid}}^{y_M} p y f(y) dy \end{aligned} \quad (1)$$

$$\begin{aligned} Var(y) &= \int_{y_m}^{y_M} y^2 f(y) dy - E(y)^2 \\ &= \int_{y_m}^{y_{mid}} (1-p) y^2 f(y) dy - E(y)^2 + \int_{y_{mid}}^{y_M} p y^2 f(y) dy - E(y)^2 \end{aligned} \quad (2)$$

where y_m and y_M are the expected minimum (m) and maximum (M) wage of student i at age 35 conditional on working full time with educational degree d . y_{mid} indicates the midpoint between these two points ($\frac{y_m + y_M}{2}$) and p represents the respective probability to earn more than $\frac{y_m + y_M}{2}$. Given that we elicited wage expectations conditional on full time employment, biases arising from different labor supply expectations are ruled out by construction.

⁹In the robustness section, I investigate the sensitivity of this assumption by assuming that wage expectations are uniformly distributed over the two intervals.

¹⁰For further details see Guiso et al. (2002), Appendix D.

In the analysis, I pool wage expectations across the three education degrees and exclude students whose responses suggest that they did not entirely understand the question. These are students who assigned the entire probability mass either to the lower or the upper part of the support of their individual wage distribution, i.e. to the left or the right of the midpoint (N=62). Further, to ensure that the analysis is not driven by a few outliers, I exclude students whose average expected wage fall in either the highest or the lowest one percentile of the cross-sectional distribution of expected wages for each education degree (N=38). Further, I only keep students who provided complete information on all covariates. The final sample for the empirical analysis consists of 1430 observations, corresponding to 516 students, of whom 207 are males and 309 females.

2.2 Gender differences in wage expectations and observed characteristics

Gender differences in wage expectations. I start by looking at the distribution of expected average wages across individuals. Figure 1 depicts the cross-sectional distributions of (individual) expected average wages separately for males and females as well as for each education scenario. Several 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. At the same time, the higher the educational degree, the more dispersed the distribution. This points toward increasing heterogeneity in expected average wages with increasing levels of education. Second, the distribution for both groups, males and females, is skewed to the right, indicating that students are less likely to expect very high wages. Third, and finally, in all education scenarios the distribution of males is shifted to the right and exhibits a thicker right tail, thus implying that males expect higher wages than females on average and are more likely to expect exceptionally high wages. The Kolmogorov-Smirnov test confirms that the distributions of cross-sectional average expected wages differ significantly by gender in each education scenario.

[Figure 1 about here]

In line with previous findings in the literature, Panel A of Table 1 documents a large gender gap in expected average wages. Even before entering the labor market females expect to earn considerably less than their male counterparts. While males expect to earn on average around 3,212 EUR per month, females expect to earn 2,695 EUR. This difference implies a gender gap in wage expectations of around 16%. It also shows that students perceive a considerable amount of wage risk as measured by the standard deviation of

expected wages, where females expect lower wage risk than males.¹¹

Gender differences in observed characteristics. The literature offers a range of hypotheses that (may) explain the gender wage gap.¹² In general, if we assume that students are able to correctly identify factors that are more or less rewarded in the labor market, then all explanations that are typically brought forward with regard to the actual gender wage gap may also help to explain the gender gap in wage expectations. Including an extensive set of covariates in the analysis serves two main purposes. First, if expected wage risk plays a role in explaining gender differences in wage expectations, this correlation should be robust to accounting for other individual characteristics. These may be correlated with expected wages as well as expected wage risk and simultaneously vary by gender. Second, including a set of additional explanatory variables enables me to set the explanatory power of anticipated compensation for wage risk in relation to the importance of other characteristics that differ by gender.

[Table 1 about here]

Table 1 presents averages in observed characteristics separately for males and females as well as the difference between them. I include six sets of covariates that represent different explanations for the gender gap in wage expectations: (1) *Baseline characteristics* account for socio-demographic differences. In regards to these, males and females do not differ significantly, with the exception of attended high school type. Females are around 7 percentage points (pp) more likely to have attended the most prestigious high school track (*Gymnasium*). (2) *Performance and cognitive skills* reflect standard human capital variables and comprise students' final high school GPA as well as test scores on a verbal and figural cognition test.¹³ The final high school GPA signals students their ability (which may or may not be correct) and we would expect higher performing students to anticipate that their higher ability (or signal thereof) is rewarded in the labor market. Looking at academic performance and cognitive skills, somewhat surprisingly, females score significantly lower on the verbal cognition test; however, the difference is, at around one point, rather small and corresponds to less than half of a standard deviation. (3) *Intended college major* accounts for the well-documented wage differences between

¹¹Note that females wage expectations are lower for each education scenario and that expected wage risk increases with the level of educational degrees for both genders (see Table A.2 in the Appendix), which is in line with the empirical findings on actual labor market data (Koerselman and Uusitalo, 2014).

¹²For a comprehensive overview of existing explanations see Blau and Kahn (2016).

¹³Note that in Germany final high school GPA ranges from one to four and a higher GPA corresponds to lower performance. The cognition tests were conducted in the first wave of the panel survey. Higher scores on the cognitive tests indicate higher skills.

college majors as well as differences in college major choice by gender. As we asked students about their wage expectations for themselves, these expectations should at least partly be influenced by the type of occupation they aspire to or the college major they intend to enroll in (Montmarquette et al., 2002; Arcidiacono et al., 2014). To minimize the possibility that the observed gender gap in wage expectations is entirely due to gender differences in (future) major choice, I control for students' intended college major. In that sense, I am focusing on gender differences within (intended) college major.¹⁴ 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. Regarding intended college major, the well-known pattern is found. Females are more likely to intend to enroll in language and cultural as well as medical studies and significantly less likely to intend to enroll in science, technology, engineering, and mathematics (STEM) majors. (4) *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 are likely to differ by gender (Humlum et al., 2012). In particular, one might expect that females anticipate future career breaks due to child bearing and rearing (Chevalier, 2007). If they incorporate family formation plans into their wage expectations, this may explain their lower wage expectations compared to males. Although I do not have direct information on these plans, the survey includes a battery of questions (11 items) capturing how important different career aspects are to students' future job choice. Among others, students report how important it is for their future job choice to have a job that leaves enough time for family commitments. The different items can be grouped into four main categories (see Weinhardt and Schupp, 2011):¹⁵ (a) Extrinsic motives; (b) Intrinsic motives; (c) Social motives; and (d) Work-life balance related motives. The grouping of the 11 items into these categories can be inferred from Table 1. These career motives differ significantly between genders. Interestingly, males and females do not differ significantly with respect to their overall extrinsic motivation, although females are around 9 pp less likely than males to state that a high income is very important for their job choice. Moreover, females put more weight on intrinsic, social, and work-life balance oriented motives than males. For females' job choice it is more

¹⁴The 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, I 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.

¹⁵The grouping of items follows the recommendation of Weinhardt and Schupp (2011). However, one item was excluded because an exploratory factor analysis revealed that it did not load on the correct factor.

important to have an interesting job, have a job with frequent social interaction, and a job in which you help others. Interestingly, having a job that is important for society is equally important to males and females. Females are also more likely to report that it is very important to have a job that offers good health as well as safety conditions and leaves sufficient time for family commitments. (5) *Personality traits and preferences* account for differences in non-cognitive skills and preferences that may affect the career path that students strive for. A growing literature emphasizes the importance of psychological attributes, preferences, and personality traits in explaining educational choices (Koch et al., 2014) as well as labor market outcomes (Heckman and Kautz, 2012) and document gender differences with respect to these non-cognitive skills (for an overview see Bertrand, 2011).¹⁶ To capture differences in non-cognitive skills and preferences, I include personality traits, locus of control, risk aversion, and time preferences. Personality traits are measured by an adjusted version of the Five Factor Model that covers openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (Big Five) (McCrae and Costa, 1996).¹⁷ Comparing personality traits, females are more open, more conscientious, more neurotic, and more agreeable than males. The extent of extraversion is the only dimension of personality that does not differ significantly between gender. The locus of control indicates how strongly an individual believes that what happens is a consequence of her own actions (internal) or outside her control, i.e. due to luck and fate (external).¹⁸ Females show a higher external locus of control, indicating that they perceive their life to be more affected by circumstances outside their control than males. Risk aversion is measured by asking students to rate their willingness to take risks on a 11-point scale, a validated measure capturing the likelihood of risky behavior in different domains (Dohmen et al., 2011). As is commonly found in the literature, Table 1

¹⁶For example, Grove et al. (2011) find that, based on U.S. data, the inclusion of measures on non-cognitive skills and work preferences significantly increase the explained part of the gender pay gap for a sample of individuals with a Master’s degree.

¹⁷Definitions of the five dimensions as cited in Almlund et al. (2011) are: *openness* defines the tendency to be open to new aesthetic, cultural, or intellectual experiences; *conscientiousness* mirrors the tendency to be organized, responsible, and hard-working; *extraversion* describes an orientation of one’s interests and energies toward the outer world of people and things rather than the inner world of subjective experience and is characterized by positive affect and sociability; *agreeableness* shows the tendency to act in a cooperative, unselfish manner; and *neuroticism* is a chronic level of emotional instability and proneness to psychological distress. Each dimension is represented by three statements that are answered on a 7-point Likert type scale ranging from 1 “does not apply at all” to 7 “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, I generate summation scores for each personality dimension.

¹⁸These two measures are based on eight different items, capturing the extent to which individuals agree (on a 7-point Likert type scale) with statements such as “*The possibilities I have in life are dependent on social circumstances.*”

shows that females are more risk averse than males. Finally, time preferences account for inter-temporal choice behavior and are approximated by a measure for patience and present bias.¹⁹ However, males and females do not differ significantly with regards to their time preferences. (6) *Self-confidence* accounts for gender differences in the assurance to succeed in the labor market that could result in higher expected wages. 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 7-point scale. Additionally, students were asked how likely they think it is that they could successfully graduate from university. Answers are given on a 4-point scale ranging from very low to very high. I generate a binary variable that equals one if students rate their chances as very high and zero otherwise. This measure may yield some indication about students' confidence in their own abilities. As can be seen in Table 1, females have considerably less confidence in their own abilities and a less positive perception of themselves when compared to males.

In Section 4, I investigate to what extent these differences are related to gender differences in expected average wages.

3 Do students anticipate compensation for wage risk?

The argument of this study rests on the assumption that students anticipate compensation for wage risk. Thus, before focusing on whether anticipated compensation for wage risk can explain the gender gap in wage expectations, I provide corroborating evidence on this relationship. Risk compensation is typically uncovered by first estimating a standard Mincer wage equation and then using the variance of the residuals within an education-occupation group as a measure for wage risk. In a second step, this risk measure is then included in the Mincer wage equation yielding the risk augmented Mincer wage equation (see Hartog, 2011).

In line with this approach, there are two studies examining whether students anticipate compensation for wage risk, i.e. expect a positive correlation between higher average wages and higher wage risk (Mazza and Hartog, 2011; Schweri et al., 2011); however, with mixed results. Schweri et al. (2011) gathered information on the expected wage distribution of first year economics students in Berne, finding that higher expected average wages are

¹⁹Patience is measured by asking students to rate their general patience on an 11-point scale from very impatient to very patient. This survey measure is shown to correlate well with individuals' choices in an incentive based choice experiment eliciting time preferences (Vischer et al., 2013). The measure for present bias is based on the average degree of agreement (7-point scale) with the following two statements: *"I rather have fun today and do not think about tomorrow"* and *"I pass on something today, in order to be able to afford more tomorrow"* (reversed item).

indeed significantly positively related to expected wage risk. By contrast, using the same survey questions as Schweri et al. (2011) on Dutch high school students in their last year, Mazza and Hartog (2011) cannot find a similar relationship between expected wages and wage risk. Thus, additional evidence on this relationship is required. This section determines whether the results of Schweri et al. (2011) can be replicated using a different sample and an alternative measure for wage risk.²⁰

Following the analysis of Schweri et al. (2011) and Mazza and Hartog (2011), I analyze whether students expect compensation for wage risk by estimating the following model:

$$\ln(y_{id}) = \alpha + \beta_1 risk_{id} + X_i' \beta_2 + \mu_d + \varepsilon_{id}, \quad (3)$$

where y_{id} represents the expected wage of student i conditional on earning educational degree d ($d = 1, 2, 3$, i.e. vocational, Bachelor's or Master's degree). $risk_{id}$ reflects expected wage risk as measured by the dispersion in expected wages. More specifically, expected wage risk is measured by the log of the individual-specific variance of expected wages as described in Section 2, such that β_1 can be interpreted as an elasticity and is the coefficient of main interest. If students expect compensation for wage risk, the estimate of β_1 should be positive. X_i is a vector of control variables including baseline characteristics, academic performance and cognitive skills as described in Section 2 and μ_d represents indicator variables for each education scenario. Finally, ε_{id} is an individual error term and captures the remaining variation. As the error terms among the same student are likely to be correlated, I cluster standard errors at the individual level.

The results are presented in Table 2 and provide supporting evidence for the findings of Schweri et al. (2011). While the first three columns are based on expectations pooled over the three different educational degrees, column (4), (5), and (6) show the estimates for each educational degree separately.

[Table 2 about here]

As shown in column (1) the coefficient for expected wage risk is positive and highly significant. An increase in expected wage risk, as measured by the variance of the expected wages, by one percent increases expected average wages by 0.160 percent.²¹ Adding

²⁰Both studies use the median version (see Section 2) to elicit students' expected wage distribution and use the probability mass assigned to the outer tails of the distribution as a measure for the variance. They further analyze the effect of skewness of the expected wage distribution as an additional measure for wage uncertainty. By contrast, this study elicited the distribution of expected wages based on another approach that is detailed in Section 2.

²¹For an in-depth discussion on why a positive relationship between the mean and the variance in wages is not a mechanical relict, see Hartog (2011).

baseline characteristics and controlling for academic performance and skills only affects this relationship marginally (see column 2). In column (3), I further test whether the relationship between expected average wages and expected wage risk differs by gender. This information is essential for evaluating the role of expected wage risk in explaining gender differences in expected wages. If, for example, females expect a significantly higher compensation for wage risk than males (due to their higher risk aversion), including expected wage risk to explain gender differences in expected wages would not only capture differences in expected earning risk but also the different extent of anticipated risk compensation. Hence, in column (3), I estimate Equation 3 and additionally include an interaction term between expected wage risk and the female indicator. The coefficient on the interaction term is close to zero and not statistically significant, suggesting that males and females do not differ in their extent of anticipated compensation for wage risk.

In order to investigate whether the relationship between expected average wages and expected wage risk is different for the three education scenarios, I repeat the estimation for each education scenario separately (columns 4-6). The coefficient on wage risk is positive and highly significant in all education scenarios. Interestingly, the coefficient on expected wage risk increases with the level of the education scenario; it is lowest for expected wages with a vocational degree (0.113), followed by expected wages with a Bachelor's degree (0.179), and highest for expected wages with a Master's degree (0.191). Anticipated compensation for wage risk seem to increase with the education level, suggesting that students perceive higher levels of education to be more risky and thus expect an additional compensation for taking this risky decision. This argument fits well with the more theoretical rationale that students also expect compensation for postponing earnings while staying in education longer (Hartog, 2011). In addition, based on data covering 16 different countries, the study by Pereira and Martins (2002) similarly shows a positive relationship between average wages and wage risk with increasing levels of education.

The estimates reported in Table 2 are similar to the reported elasticity of 0.125 percent in the study by Schweri et al. (2011). Given the differences in eliciting information on the expected wage distribution and measurement of expected wage risk, this finding is remarkable. The elasticity estimates as shown in Table 2 are also comparable to estimates on risk compensation using actual labor market data. In an overview Hartog (2011) reports that most estimates for risk compensation fall between 0.1 - 0.2;²² as such, the elasticity estimates in Table 2 are contained in this interval.

Overall, this analysis shows that students anticipate compensation for wage risk at

²²These estimates are based on occupation specific risk measures. Hartog (2011) notes, however, that education based risk measures may lead to lower elasticity estimates.

similar elasticities as found based on actual labor market earnings. Having established this relationship, I now turn to analyzing the gender gap in wage expectations.

4 The role of expected wage risk in explaining the gender gap in expected wages

In order to investigate how the gender gap in wage expectations can be explained, I perform a standard Blinder-Oaxaca decomposition (Oaxaca, 1973; Blinder, 1973). I start by estimating separate regression for males and females:

$$\ln(y_{id}) = \alpha_f + X_i' \beta_f + \varepsilon_{id} \quad (4)$$

$$\ln(y_{id}) = \alpha_m + X_i' \beta_m + \varepsilon_{id} \quad (5)$$

where y_{id} is the (log) expected average wage of student i conditional on earning educational degree d ($d = 1, 2, 3$, i.e. vocational, Bachelor's or Master's degree) and X_i is a vector of explanatory variables containing – apart from the measure of (log) expected wage risk as defined in Section 2 – a broad set of different student characteristics, cognitive and non-cognitive skills, economic preferences, intended college major as well as variables capturing different career motives (see Section 2 for more details). ε_{id} is an individual error term, which is again clustered at the individual level.

The average difference in expected wages between males and females can then be expressed as:

$$(\bar{Y}_f - \bar{Y}_m) = (\bar{X}_f - \bar{X}_m)' \beta_m + \bar{X}_m' (\beta_f - \beta_m) - (\bar{X}_f - \bar{X}_m)' (\beta_f - \beta_m) \quad (6)$$

where the first part on the right-hand side captures the part of the difference that can be explained by mean differences in observed characteristics and is typically labeled *the explained component*; the second term measures the extent to which these characteristics affect males' and females' expected mean wages differently; and the third term represents the interaction between differences in mean characteristics and differences in coefficients. The latter two make up *the unexplained component* of the difference. In this analysis I focus exclusively on the explained component. The result of the decomposition will depend on whether males or females are used as a reference group. Equation 6 is expressed from the viewpoint of males, i.e. it uses the male coefficient to weight the group differences in characteristics and likewise uses male characteristics to weight group differences in coefficients. However, Equation 6 could also be expressed from the viewpoint of females. Thus, in my preferred specification I follow the recommendation of Neumark (1988) and

use the coefficient from a pooled regression, such that the average difference in expected wages between males and females can be expressed as:

$$(\bar{Y}_f - \bar{Y}_m) = (\bar{X}_f - \bar{X}_m)' \beta^* + \bar{X}_f' (\beta_f - \beta^*) + \bar{X}_m' (\beta^* - \beta_f) \quad (7)$$

where β^* represents the estimates from a pooled regression over males and females.

4.1 Decomposition results

The results of the decomposition are presented in Table 3. In order to examine the role of different explanations in accounting for the gender gap in wage expectations, columns (1)-(8) in Table 3 display how a particular set of covariates affect the gender gap in expected wages. The raw gender gap amounts to 0.177 log points indicating that females expect to earn around 16% less than males.

Adding baseline characteristics increases the gender gap by 2.4% (column 1). Similarly, accounting for differences in performance at school and cognitive skills marginally increase the gender gap by 0.5% (column 2).

Assuming that students are aware of the wage differences across college majors, differences in expected wages should reflect gender differences in intended field of study.²³ As shown in column (3) students' intended college major explain 6.0% of the gender gap in expected wages. This shows that most of the gender differences in expected average wages occur within (intended) college major. This finding is line with the results reported in Reuben et al. (2015).

The role of differences in career motives is shown in column (4). Including how important particular aspects are for students' job choice accounts for 4.2% of the gender gap. Being more extrinsically motivated is significantly related to higher expected wages, suggesting that students whose job choice is motivated by earning a high income, having good promotion possibilities or a highly recognized job expect higher average wages. Similarly, being more socially or work-life balance motivated is related to lower expected wages.

Accounting for differences in personality traits leads to a reduction in the gender gap of 7.1% (column 5). Being more neurotic or agreeable is negatively correlated with expected average wages, showing that students who have a stronger tendency to act cooperatively and unselfishly expect to earn less and may thus indeed settle for lower wages. A higher internal locus of control appears to be positively related to expected wages; however, this

²³Note that for wage expectations with a vocational degree, intended college major may not adequately reflect the gender specific sorting into different occupations. However, it may still capture gender differences in subject-specific choices that similarly applies to occupational choices.

correlation misses the ten percent threshold for statistical significance.

In column (6) the role of differences in preferences is shown. Differences in risk aversion, patience and time preference lead to a decrease of the gender gap of 1.4%. In particular, being more risk averse is negatively correlated with expected wages, showing that more risk averse students expect lower wages.

The results in column (7) underline the importance of students' confidence in their abilities as well as themselves. Taking the measures for confidence into account decreases the gender gap by 8.7%. Gender differences in confidence are even more important than differences in intended college major. Being very sure that one could successfully complete a college degree is significantly related to higher expected wages. Similarly, being more self-confident is positively related to expected wages.

Column (8) of Table 3 reveals that expected wage risk plays a crucial role in explaining the gender gap in expected average wages. Including the measure of expected wage risk, as measured by the (log) variance of the individual expected wage distribution, accounts for about three-quarters of the gender gap. Accounting for differences in expected wage risk decreases the gender gap in expected wages from 0.177 to 0.0413 log points (or from 16% toward 4%). Recall that the extent of anticipated compensation for wage risk does not vary significantly by gender (see Table 2) and, hence, the reduction in the gender gap is solely driven by differences in expected wage risk. Analyzing the role of wage risk in explaining the gender gap in wage expectations for each education scenario separately reveals an interesting observation. The role of anticipated risk compensation appears to increase with higher educational degrees. While differences in expected wage risk can account for around 87% of the gender gap in expected wages with a Master's degree, this share equals 56% for expected wages with a Bachelor's degree, and 47% for expected wages with a vocational degree (see Table A.3). This observation suggests that anticipated compensation for wage risk is more relevant in the upper part of the wage distribution than in the lower part.

Including all explanatory variables except expected wage risk (column 9), decreases the gender gap in wage expectations by around 23%. Comparing the decomposition results when including all explanatory variables except expected wage risk (column 9) and all variables, i.e. including expected wage risk (column 10) shows that anticipated compensation for wage risk adds around two-thirds to the explained part of the gender gap in expected wages. In order to examine the specific contribution of expected wage risk to explain the gender gap while simultaneously including all other variables, Table A.4 in the Appendix presents the results of a detailed decomposition, where the detailed

contributions of each set of predictors to the explained part of the gender gap are shown.²⁴ These results confirm that differences in expected wage risk are the single most important factor in explaining the gender gap in wage expectations.²⁵

Finally, by including the full set of explanatory variables in column (10) 89.5% of the gender gap in expected average wages can be explained, yielding an insignificant estimate of gender gap that is close to zero.

Overall, the decomposition analysis shows that including expected wage risk substantially increases the explained part of the gender gap in expected average wages. Expected wage risk can explain three times as much of the gender gap as all other included variables combined, highlighting that expected wage risk is unlikely to be captured by students' characteristics that may be more easily observed. Given the role of differences in expected wage risk in explaining the gender gap in expected wages, Table 3 provides supportive evidence for the hypotheses that females expect to trade off higher wages for lower wage risk.

4.2 Sensitivity analysis

The key finding of the empirical analysis is that anticipated compensation for wage risk explains a considerable part of the gender gap in expected average wages. This section investigates how sensitive this finding is to different specifications. A summary of the sensitivity analysis is presented in Table 4. I use the estimate of the reduction in the gender gap in expected average wages when accounting for expected wage risk (see Table 3, column 8) as a reference point to which alternative specifications are compared. This estimate is displayed again in the first row of Table 4. Column (1) presents the raw gender gap in log expected wages, while column (2) transforms this gap in a percentage difference. Column (3) displays the explained part and column (4) the unexplained part of the gender gap. Column (5) shows how much of the gender gap (in %) can be explained by accounting for differences in expected wage risk.

[Table 4 about here]

In order to calculate moments of the individual expected wage distribution, I followed the approach of Guiso et al. (2002) and Attanasio and Kaufmann (2014). I assumed that expected wages are distributed triangularly over the two intervals from the minimum to

²⁴Note that in the case of the detailed decomposition, the other variables are accounted for by the fact that the regression coefficients used to weight the group differences are based on partial correlations, i.e. obtained from a regression including the full set of covariates.

²⁵In the detailed decomposition, differences in expected wage risk can account for more than three-quarters of the gender gap in wage expectations (see Table A.4).

the midpoint and from the midpoint to the maximum. To verify that the main finding is not driven by the distributional assumption, I repeat the estimation and assume a uniform distribution. The results are presented in row (2). The first column shows that based on this alternative distributional assumption the gender gap is with 0.1763 log points only marginally smaller. Adding the measure of expected wage risk reduces the gap by 76.2%, which is almost the same as 76.7% in row (1). Hence, the finding that anticipated compensation for wage risk can explain a very high share of the gender gap in expected average wages does not hinge on the distributional assumption.

Next, I use the range between expected maximum and minimum wage as an alternative measure for expected wage risk. The estimates in row (3) show that using this alternative measure of wage risk does not change the role of expected wage risk in explaining the gender gap in averages expected wages. Expected wage risk still accounts for 77.2% of the gender gap.

In the remainder of Table 4, I analyze how different sample restrictions and item non-response affect the role of expected wage risk in explaining the gender gap in expected average wages. In row (4) I include only students who answered the complete module on expected wages for all education scenarios. This results in a slightly smaller gender gap (0.1693 log points) and a slight decrease in the share of the gender gap that is explained by differences in expected risk (75.7%).

In the main specification, I only exclude wage expectations that fell in the upper or lower one percentile of the cross-sectional distribution to account for outliers. In row (5) I alternatively exclude all observations that fall in the upper and lower five percentiles, while in row (6) I do not account for any outliers. As shown, focusing only on wage expectations between the 5th and the 95th percentile of the cross-sectional distribution reduces the raw gender gap (0.1454 log points) as well as the share that is accounted for by differences in expected wage risk (64.8%). In contrast, not accounting for any outliers increases the gender gap in expected wages (0.2227 log points), while the share of the gap that can be explained by differences in expected wage risk does not change much (76.2%). In the analysis, I excluded all students with missing information on any covariates in order to analyze the role of alternative explanations for the gender gap in expected wages. In row (7), I include those observations, which yields a slightly lower gender gap of 16.76 log points and a lower share of that gap that is explained by expected wage risk (69.4%). In row (8) I use all available information on expected wages which results in a higher gender gap (21.24 log points) and a slightly smaller share that can be accounted for by differences in expected wage risk. Finally, as outlined in Section 2, a large share of students did not answer the questions on their wage expectations. Table A.1 suggests that this item non-

response may not be at random. Thus, in the last column of Table 4 I estimate a selection corrected model, as suggested by Heckman (1979) and adjust the decomposition results. As exclusion restrictions I use the time each student spent on the questionnaire as well as a binary variable indicating whether a student is a stable panel member, which is defined as having participated in all five waves of the panel survey. These variables may contain information on how dedicated students are to the survey and, hence, how thoroughly they fill in the questionnaire.²⁶ Adjusting the decomposition for selective response behavior, yields a higher gender gap in average expected wages (0.2244 log points) and a reduction in the share of the gender gap that is attributable to differences in expected wage risk (64.9%).

Overall, the estimates in Table 4 suggest that the role of expected wage risk in explaining gender differences in expected average wages does not depend on the exact specification. Although the percentage contribution of expected wage risk in explaining the gender gap in expected wages varies, the general conclusion is unaffected: Expected wage risk plays a crucial role in explaining the gender gap in expected wages.

4.3 Are females just better informed?

The analysis presented so far provides evidence for the hypothesis that females deliberately trade off higher wages for lower wage risk. In this section I discuss an alternative hypothesis that could explain why females expect to earn less and, at the same time, expect lower wage risk. I provide evidence that this explanation is unlikely to apply.

It is typically assumed that students base their wage expectations on what they observe from currently active labor market participants. Further, several studies show that students tend to overestimate their own future wage in relation to what current employees earn (Botelho and Pinto, 2004; Jerrim, 2015). Thus, it might be that females expect to earn less than males because they are better informed, i.e. their expected wages are closer to what can be observed in the labor market. Being better informed may also entail lower expected wage risk. In this case, the lower wage expectations and lower expected wage risk of young women may just reflect more accurate information.

In order to explore this alternative hypothesis, I examine the importance of ‘misinformation’ considering all education scenarios. In Table 5, I show students’ expected average wages in the estimation sample (Panel A) and additionally report actual average wages as calculated from the German Micro Census (Panel B).²⁷

²⁶Both variables predict response behavior on a statistically significant basis.

²⁷For the calculation of average wages I only use the years 2010-2012. As we asked students about their expected wage at the age of 35, the figures presented are based on individuals between 33 and 37 with an *Abitur*. Note that in the German Micro Census individuals only report their net income within

Difficulties in the comparison between expected wages and actual wages arise because the number of active labor market participants in Germany with a Bachelor’s or a Master’s degree is rather small for historical reasons.²⁸ It might be even less clear to students how the labor market rewards these relatively ‘new’ degrees in comparison to the previously awarded degrees. This may further contribute to the uncertainty students face when forming expectations about their future wages. However, considering the years of education necessary to earn these different degrees, wages with a Bachelor’s degree should be closer to wages with a vocational degree, while wages with a Master’s degree should be closest to wages with the previous German tertiary degrees. As shown in Panel B of Table 5, where I additionally report average wages for individuals with a tertiary university degree (i.e. pooled over all tertiary degrees), this is indeed the case.

Reverting to the question whether the observed gender differences in expected wages could be partly due to females being ‘better’ informed, Table 5 reveals that this is unlikely. Panel C of Table 5 shows that the majority of students overestimate their own wage (or are likely to overestimate their wage).²⁹ Focusing on gender differences, females are even more likely to overestimate their wages than males in all education scenarios. To account for the possibility of over- and underestimation, I additionally calculate the percentage absolute deviation between expected average wages and actual population wages on an individual basis (see Panel D).³⁰ Students’ expected wages with a vocational degree deviate on average by 28% for females and 23% for males. With respect to higher educational degrees, these deviations increase. Comparing the deviations of expected wages with a Bachelor’s [Master’s] degree from population wages with the same degree, shows that females’ deviations on average amount to 37% [70%], while males’ deviations equal on average 33% [54%]. Thus, from Table 5 I conclude that females are certainly not better informed than males; if anything, females appear to be more poorly informed.

specific bins, which might not necessarily consist only of labor market income. To account for this, I only include individuals who work full time and report that their main source of income derives from labor market income. Moreover, I use the midpoint of each bin to proxy actual wages.

²⁸The introduction of the two-tier system of Bachelor’s and Master’s degrees was only initiated in the academic year 2000/2001. This replaced the German degrees formerly awarded, like the *Diplom* or the *Magister*. Both degrees were awarded after completing 4-5 years of schooling; this is, roughly a Master’s degree.

²⁹The only exception are expected wages with a vocational degree. One explanation for the smaller difference in expected wages and actual wages with a vocational degree relates to the sample of students under consideration. Recall that the survey oversampled students from lower educated backgrounds. Given this background it may be easier for students to observe individuals with a vocational degree than individuals holding a higher education degree. Hence, their information on population wages with a vocational degree may be more accurate than their wage expectations with a higher education degree.

³⁰The percentage absolute deviation is calculated as: $\Delta_i = (|y_{id}^e - \bar{y}_d^p|) / \bar{y}_d^p$, where y_{id}^e represent student i ’s expected wage with education degree d and \bar{y}_d^p indicates actual average population wages with education degree d .

Consequently, the gender gap in wage expectations and the role of expected wage risk in accounting for this gap cannot be explained by females holding more realistic wage expectations than males.

5 Conclusion

Several studies show that females start out with lower wage expectations, even before entering the labor market, which partly translates into the actual gender wage gap. The main channels run through the effect of expected wages on educational choice and the formation of reservation wages. While previous studies have focused on wage expectations of college students, this study examines wage expectations of high school graduates, thereby analyzing gender differences in wage expectations before students self-select into different college majors and/or levels of education. A supplementary analysis (see Appendix B) has shown that wage expectations contain information about future educational choices that are relevant for actual wage differences between genders in the labor market. Thus, analyzing the determinants of the gender gap in wage expectations can contribute to a better understanding of the actual gender wage gap.

This study investigates gender differences in wage expectations and explores the role of anticipated compensation for wage risk in explaining the gender gap in wage expectations. The focus on wage risk stems from the recurrent finding that women are more risk averse than men. In addition, based on labor market data it is repeatedly shown that – due to compensating wage differentials – higher average wages are positively related to higher wage risk. If students anticipate this form of risk compensation, this may explain why women have lower earnings expectations. Based on a unique survey in which we elicited information on the entire distribution of students' wage expectations, it is possible to construct a measure of expected wage risk by computing the individual-specific variance of expected wages.

The results of the empirical analysis can be summarized in three key findings: First, already at high school graduation female students expect to earn around 16% less than their male counterparts but simultaneously expect lower wage risk. This observation cannot be explained by females being better informed about actual labor market wages. Second, students anticipate compensation for wage risk, i.e. higher expected average wages are positively correlated with higher expected wage risk. This finding supports the evidence provided by Schweri et al. (2011). Additionally, the extent of anticipated risk compensation does not differ by gender, i.e. despite higher risk aversion, female students do not anticipate higher compensation for wage risk. And third, a considerably large share of the gender gap in wage expectations can be explained by differences in

expected wage risk. This share amounts to three-quarters of the gender gap in expected average wages. Given the extensive set of additional covariates included in the analysis to capture alternative explanations for the gender gap, the importance of expected wage risk in explaining the gender gap in expected average wages is emphasized.

Overall, the results of this study suggest that females expect to earn less because they are willing to trade off higher wages for lower wage risk. This finding provides supporting evidence for a conscious selection of females into lower paying careers and/or occupations and industries in exchange for lower wage risk. While the results of this article are descriptive in nature, they provide additional evidence on why women self-select into lower paying careers and postulate a fruitful perspective for future research. A particularly interesting question is whether the self-selection process into different majors and/or occupations could be altered if wage perspectives would be perceived as more stable. Providing causal evidence in this direction could have important policy implications.

References

- Almlund, M., Duckworth, A. L., Heckman, J., and Kautz, T. (2011). Chapter 1 - Personality, Psychology and Economics. In Eric A. Hanushek, S. M. and Woessmann, L., editors, *Handbook of The Economics of Education*, volume 4 of *Handbook of the Economics of Education*, pages 1–181. Elsevier.
- Arcidiacono, P., Hotz, V. J., Maurel, A., and Romano, T. (2014). Recovering ex ante returns and preferences for occupations using subjective expectations data. *NBER Working Paper No. 20626*. National Bureau of Economic Research.
- Attanasio, O. P. and Kaufmann, K. M. (2014). Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender. *Journal of Development Economics*, 109:203–216.
- Attanasio, O. P. and Kaufmann, K. M. (2017). Education choices and returns on the labor and marriage markets: Evidence from data on subjective expectations. *Journal of Economic Behavior & Organization* (forthcoming).
- Belfield, C., Boneva, T., Rauh, C., and Shaw, J. (2016). Money or fun? Why students want to pursue further education. *IZA Discussion Paper No. 10136*. Institute for the Study of Labor.
- Bertrand, M. (2011). New perspectives on gender. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4, part B, chapter 17, pages 1543–1590. Elsevier.
- Betts, J. R. (1996). What do students know about wages? evidence from a survey of undergraduates. *The Journal of Human Resources*, 31(1):27–56.
- Black, D. A., Haviland, A. M., Sanders, S. G., and Taylor, L. J. (2008). Gender wage disparities among the highly educated. *The Journal of Human Resources*, 43(3):630–659.
- Blau, F. D. and Ferber, M. A. (1991). Career plans and expectations of young women and men: The earnings gap and labor force participation. *The Journal of Human Resources*, 26(4):581–607.
- Blau, F. D. and Kahn, L. M. (2016). The gender wage gap: Extent, trends, and explanations. *IZA Discussion Paper No. 9656*. Institute for the Study of Labor.

- Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *The Journal of Human Resources*, 8(4):436–455.
- Bonin, H., Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2007). Cross-sectional earnings risk and occupational sorting: The role of risk attitudes. *Labour Economics*, 14(6):926–937.
- Bonnard, C., Giret, J.-F., and Mener, M. L.-L. (2014). Educational intentions, cognitive skills and earnings expectations of french undergraduates. *Applied Economics Letters*, 21(18):1293–1296.
- Botelho, A. and Pinto, L. C. (2004). Students’ expectations of the economic returns to college education: Results of a controlled experiment. *Economics of Education Review*, 23(6):645–653.
- Brunello, G., Lucifora, C., and Winter-Ebmer, R. (2004). The wage expectations of European business and economics students. *The Journal of Human Resources*, 39(4):1116–1142.
- Caliendo, M., Gambaro, L., and Haan, P. (2009). The impact of income taxation on the ratio between reservation and market wages and the incentives for labour supply. *Applied Economics Letters*, 16(9):877–883.
- Caliendo, M., Lee, W.-S., and Mahlstedt, R. (2017). The gender wage gap and the role of reservation wages: New evidence for unemployed workers. *Journal of Economic Behavior & Organization*, 136:161– 73.
- Chevalier, A. (2007). Education, occupation and career expectations: Determinants of the gender pay gap for UK graduates. *Oxford Bulletin of Economics and Statistics*, 69(6):819–842.
- Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–474.
- Dehne, M. and Schupp, J. (2007). Persönlichkeitsmerkmale im Sozio-oekonomischen Panel (SOEP): Konzept, Umsetzung und empirische Eigenschaften. *DIW Research Notes No. 26*. German Institute for Economic Research.
- Destatis (2012). Bildung und Kultur. Studierende an Hochschulen, WS 2011/2012. *Fachserie 11, Reihe 4.1*. Statistisches Bundesamt, Wiesbaden.

- Destatis (2017). Nichtmonetäre hochschulstatistische Kennzahlen. *Fachserie 11 Reihe 4.3.1*. German Federal Statistical Office, Wiesbaden.
- Diaz-Serrano, L., Hartog, J., Nilsson, W., van Ophem, H., and Yang, P. (2016). Student earnings expectations: Heterogeneity or noise? *CESifo Working Paper No. 6110*. version as of 19-09-2016.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinant, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.
- Dominitz, J. and Manski, C. F. (1996). Eliciting student expectations of the returns to schooling. *The Journal of Human Resources*, 31(1):1–26.
- Ehlert, M., Peter, F., Finger, C., Rusconi, A., Solga, H., Spieß, K. C., and Zambre, V. (2017). Berliner-Studienberechtigten-Panel (Best Up) – Data Description and Documentation. *DIW Data Documentation (forthcoming)*. German Institute for Economic Research.
- Filippin, A. (2003). Discrimination and workers’ expectations. *IZA Discussion Paper No. 823*. Institute for the Study of Labor.
- Filippin, A. and Ichino, A. (2005). Gender wage gap in expectations and realizations. *Labour Economics*, 12(1):125–145.
- Finke, C., Dumpert, F., and Beck, M. (2017). Verdienstunterschiede zwischen Männern und Frauen. Eine Ursachenanalyse auf Grundlage der Verdienststrukturerhebung 2014. *Wirtschaft und Statistik*, 2/2017. Statistisches Bundesamt, Wiesbaden.
- Frick, B. and Maihaus, M. (2016). The structure and determinants of expected and actual starting salaries of higher education students in Germany: Identical or different? *Education Economics*, 24(4):374–392.
- Grove, W. A., Hussey, A., and Jetter, M. (2011). The gender pay gap beyond human capital. *The Journal of Human Resources*, 46(4):827–874.
- Guiso, L., Jappelli, T., and Pistaferri, L. (2002). An empirical analysis of earnings and employment risk. *Journal of Business & Economic Statistics*, 20(2):241–253.
- Hartog, J. (2011). A risk augmented Mincer earnings equation? Taking stock. In Solomon W. Polachek, K. T., editor, *Research in Labor Economics*, volume 33, chapter 4, pages 129–173. Emerald Group Publishing Limited.

- Hartog, J., Ding, X., and Liao, J. (2012). Is earnings uncertainty relevant for educational choice? An empirical analysis for China. *Education Economics*, 22(5):471–483.
- Hartog, J. and Vijverberg, W. P. (2007). On compensation for risk aversion and skewness affection in wages. *Labour Economics*, 14(6):938 – 956.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.
- Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4):451–464.
- Humlum, M. K., Kleinjans, K. J., and Nielsen, H. S. (2012). An economic analysis of identity and career choice. *Economic Inquiry*, 50(1):39–61.
- Huntington-Klein, N. (2015). Subjective and projected returns to education. *Journal of Economic Behavior & Organization*, 117:10–25.
- Jerrim, J. (2015). Do college students make better predictions of their future income than young adults in the labor force? *Education Economics*, 23(2):162–179.
- Koch, A., Nafziger, J., and Nielsen, H. S. (2014). Behavioral economics of education. *IZA Discussion Paper No. 8470*. Institute for the Study of Labor.
- Koerselman, K. and Uusitalo, R. (2014). The risk and return of human capital investments. *Labour Economics*, 30:154–163.
- Machin, S. and Puhani, P. A. (2003). Subject of degree and the gender wage differential: evidence from the UK and germany. *Economics Letters*, 79(3):393–400.
- Manski, C. F. (1993). Adolescent econometricians: How do youth infer the returns to schooling? In Clotfelter, C. T. and Rothschild, M., editors, *Studies of Supply and Demand in Higher Education*, chapter 2, pages 43–60. University of Chicago Press.
- Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72(5):1329–1376.
- Mazza, J. and Hartog, J. (2011). Do they understand the benefits from education? Evidence on Dutch high school students’ earnings expectations. *IZA Discussion Paper No. 5714*. Institute for the Study of Labor.
- McCrae, R. and Costa, P. J. (1996). In (Ed.), p. . (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In Wiggins, J., editor, *The five factor model of personality: Theoretical perspectives*, chapter 3, pages 51–87. New York: Guilford.

- Montmarquette, C., Cannings, K., and Mahseredjian, S. (2002). How do young people choose college majors? *Economics of Education Review*, 21(6):543–556.
- Neugebauer, M., Neumeyer, S., and Alesi, B. (2016). More diversion than inclusion? Social stratification in the Bologna system. *Research in Social Stratification and Mobility*, 45:51–62.
- Neumark, D. (1988). Employers’ discriminatory behavior and the estimation of wage discrimination. *The Journal of Human Resources*, 23(3):279–295.
- Oaxaca, R. (1973). Male-Female wage differentials in urban labor markets. *International Economic Review*, 14(3):693–709.
- Oreopoulos, P., von Wachter, T., and Heisz, A. (2012). The short- and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics*, 4(1):1–29.
- Pereira, P. T. and Martins, P. S. (2002). Is there a return-risk link in education? *Economics Letters*, 75(1):31–37.
- Peter, F. H., Rusconi, A., Solga, H., and Spieß, C. K. (2016a). Anschubfinanzierung reicht nicht: Neue Studie untersucht, was zur Entscheidung fürs Studium beiträgt. *WZB-Mitteilungen*, 153:32–35.
- Peter, F. H., Rusconi, A., Solga, H., Spieß, C. K., and Zambre, V. (2016b). Informationen zum Studium verringern soziale Unterschiede bei der Studienabsicht von AbiturientInnen. *DIW Wochenbericht*, 26:555–565. German Institute for Economic Research.
- Peter, F. H. and Zambre, V. (2017). Intended college enrollment and educational inequality: Do students lack information? *Economics of Education Review*, 60:125 – 141.
- Reuben, E., Wiswall, M., and Zafar, B. (2015). Preferences and biases in educational choices and labour market expectations: Shrinking the black box of gender. *The Economic Journal*. doi:10.1111/eoj.12350.
- Ruder, A. I. and Noy, M. V. (2017). Knowledge of earnings risk and major choice: Evidence from an information experiment. *Economics of Education Review*, 57:80–90.
- Schweri, J. and Hartog, J. (2015). Do wage expectations influence the decision to enroll in nursing college? *IZA Discussion Paper No. 9120*. Institute for the Study of Labor.

- Schweri, J., Hartog, J., and Wolter, S. C. (2011). Do students expect compensation for wage risk? *Economics of Education Review*, 30(2):215–227.
- Vischer, T., Dohmen, T., Falk, A., Huffman, D., Schupp, J., Sunde, U., and Wagner, G. G. (2013). Validating an ultra-short survey measure of patience. *Economics Letters*, 120(2):142 – 145.
- Webbink, D. and Hartog, J. (2004). Can students predict starting salaries? Yes! *Economics of Education Review*, 23(2):103–113.
- Weinhardt, M. and Schupp, J. (2011). Multi-Itemskalen im SOEP Jugendfragebogen. *DIW Data Documentation No. 60*. German Institute for Economic Research.
- Wiswall, M. and Zafar, B. (2016a). Human capital investments and expectations about career and family. *NBER Working Paper No. 22543*. National Bureau of economic research.
- Wiswall, M. and Zafar, B. (2016b). Preference for the workplace, human capital, and gender. *NBER Working Paper No. 22173*. National Bureau of economic research.
- Wolter, S. C. (2000). Wage expectations: A comparison of Swiss and US students. *Kyklos*, 53(1):51–69.
- Wolter, S. C. and Zbinden, A. (2002). Labour market expectations of Swiss university students. *International Journal of Manpower*, 23(5):458–470.
- Zafar, B. (2013). College major choice and the gender gap. *The Journal of Human Resources*, 48(3):545–595.

6 Tables and figures

Table 1: Gender differences in wage expectations and individual characteristics

	Female	Male	Difference	p-value
Panel A: Wage expectations (pooled over degrees)				
Expected wage (in EUR)	2694.67	3211.63	-516.96***	(0.000)
Expected wage risk (measured by the standard deviation)	323.99	453.53	-129.54***	(0.000)
Panel B: Individual characteristics				
Baseline characteristics				
Migration background	0.479	0.420	0.059	(0.190)
Non-academic fam.backgr.	0.599	0.585	0.014	(0.749)
Academic high school	0.320	0.251	0.069*	(0.091)
Integrated high school	0.346	0.411	-0.064	(0.139)
Performance and cognitive skills				
Final high school GPA	2.435	2.493	-0.058	(0.283)
Verbal cognitive skills	9.832	10.918	-1.086***	(0.000)
Figural cognitive skills	11.223	10.966	0.257	(0.281)
Intended college major				
Language and Culture studies	0.126	0.043	0.083***	(0.001)
Social Sciences	0.055	0.053	0.002	(0.927)
Business and Economics	0.146	0.106	0.039	(0.193)
STEM	0.168	0.401	-0.233***	(0.000)
Teaching	0.055	0.048	0.007	(0.738)
Law and Management Sciences	0.049	0.068	-0.019	(0.357)
Medicine	0.152	0.077	0.075**	(0.011)
Psychology	0.065	0.048	0.016	(0.436)
Arts and Sports	0.042	0.053	-0.011	(0.559)
Others	0.029	0.029	0.000	(0.993)
No coll.edu. intended	0.084	0.068	0.017	(0.493)
Missing information	0.029	0.005	0.024**	(0.050)
Career motives				
Extrinsically motivated	3.098	3.068	0.030	(0.574)
High income	0.308	0.396	-0.088**	(0.040)
Good promotion possibilities	0.394	0.374	0.020	(0.643)
Recognition	0.278	0.222	0.056	(0.153)
Intrinsically motivated	3.350	3.225	0.125**	(0.012)
Interesting job	0.660	0.585	0.076*	(0.081)
Independent working	0.291	0.295	-0.003	(0.933)
Socially motivated	2.925	2.609	0.316***	(0.000)
Social interaction	0.350	0.188	0.161***	(0.000)
Important for society	0.178	0.193	-0.015	(0.662)
Help Others	0.312	0.169	0.143***	(0.000)
Work-Life-Balance motivated	3.213	3.032	0.180***	(0.000)
Spare time	0.142	0.179	-0.036	(0.267)
Good health/safety conditions	0.615	0.444	0.171***	(0.000)
Time for family commitments	0.427	0.350	0.078*	(0.078)
Personality traits				
Openness	5.146	4.972	0.174*	(0.059)
Extraversion	4.873	4.834	0.039	(0.739)
Conscientiousness	5.024	4.707	0.317***	(0.001)
Neuroticism	4.646	3.623	1.022***	(0.000)
Agreeableness	5.388	5.021	0.367***	(0.000)
External Locus of Control	3.425	3.176	0.249***	(0.005)
Internal Locus of Control	5.371	5.382	-0.011	(0.875)
Preferences				
Risk aversion	4.540	4.184	0.357*	(0.073)
Patience	5.809	6.077	-0.268	(0.184)
Time preference for present	3.264	3.348	-0.084	(0.496)
Confidence				
Confidence in own ability	0.220	0.324	-0.104***	(0.009)
Self-confidence	4.731	5.188	-0.457***	(0.001)
N	309	207		

Notes: This table presents differences in wage expectations and individual characteristics between males and females. Means and mean differences are based on a two-sided t-test. Due to item non-response the number of observations may vary. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Expected compensation for wage risk

Dep.var.: log expected mean wages	Pooled over degrees			Master's degree	Bachelor's degree	Vocational degree
	(1)	(2)	(3)	(4)	(5)	(6)
Expected wage risk	0.160*** (0.007)	0.158*** (0.007)	0.165*** (0.011)	0.191*** (0.013)	0.179*** (0.014)	0.113*** (0.016)
Bachelor's degree	0.250*** (0.012)	0.248*** (0.012)	0.248*** (0.012)			
Master's degree	0.475*** (0.015)	0.476*** (0.016)	0.475*** (0.016)			
Expected wage risk * Female			-0.013 (0.013)	-0.021 (0.017)	-0.032* (0.019)	0.016 (0.020)
Controls	No	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.6596	0.6716	0.6719	0.5950	0.4938	0.3500
N	1,430	1,430	1,430	469	476	485

Notes: Estimates are based on Equation 3 and include the following control variables: gender, migration, and academic family background, type of attended high school, final high school GPA, and two measures of verbal and figural cognitive ability. Expected wage risk is measured by the log of the variance of expected wages. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Oaxaca-Blinder decomposition

	Raw gap	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Unexplained Gap	0.1770***	0.1813***	0.1779***	0.1663***	0.1696***	0.1645***	0.1745***	0.1616***	0.0413*	0.1371***	0.0186
Explained Gap		-0.0043	-0.0009	0.0107	0.0074	0.0125	0.0025	0.0154**	0.1357***	0.0399*	0.1584***
% share explained by ...		-2.43	-0.51	6.05	4.18	7.06	1.41	8.70	76.67	22.54	89.49
Explanatory variables:											
(1) Baseline characteristics		✓								✓	✓
(2) Performance and skills			✓							✓	✓
(3) Intended college major				✓						✓	✓
(4) Career motives					✓					✓	✓
(5) Personality traits						✓				✓	✓
(6) Preferences							✓			✓	✓
(7) Confidence								✓		✓	✓
(8) Expected wage risk									✓		✓

Notes: This table presents estimates of decomposing the gender gap in expected average (log) wages based on the Oaxaca-Blinder decomposition. The number of observations equals N=1,430. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Sensitivity analysis

	Raw gender gap	Raw gender gap in %	Explained part	Unexplained part	% explained by expected wage risk	N
	(1)	(2)	(3)	(4)	(5)	(6)
(1) Main	0.1770***	16.2	0.1357***	0.0413*	76.66	1,430
(2) Uniform distribution	0.1763***	16.2	0.1344***	0.0419*	76.23	1,430
(3) Measuring wage risk by the range between y_m and y_M	0.1770***	16.2	0.1366***	0.0404*	77.18	1,430
(4) Complete response to expectation module	0.1693***	15.6	0.1282***	0.0411*	75.72	1,373
(5) Excluding upper/lower 5 percentiles	0.1454***	13.5	0.0942***	0.0512**	64.79	1,306
(6) Including upper/lower 1 percentile	0.2227***	20.0	0.1696***	0.0531**	76.16	1,462
(7) Including observations with missing covariates	0.1676***	15.4	0.1163***	0.0513**	69.39	1,741
(8) Maximum number of observations	0.2124***	19.1	0.1542***	0.0582**	72.60	1,779
(9) Adjusting for item-nonresponse	0.2244***	20.1	0.1456***	0.0788***	64.88	1,779

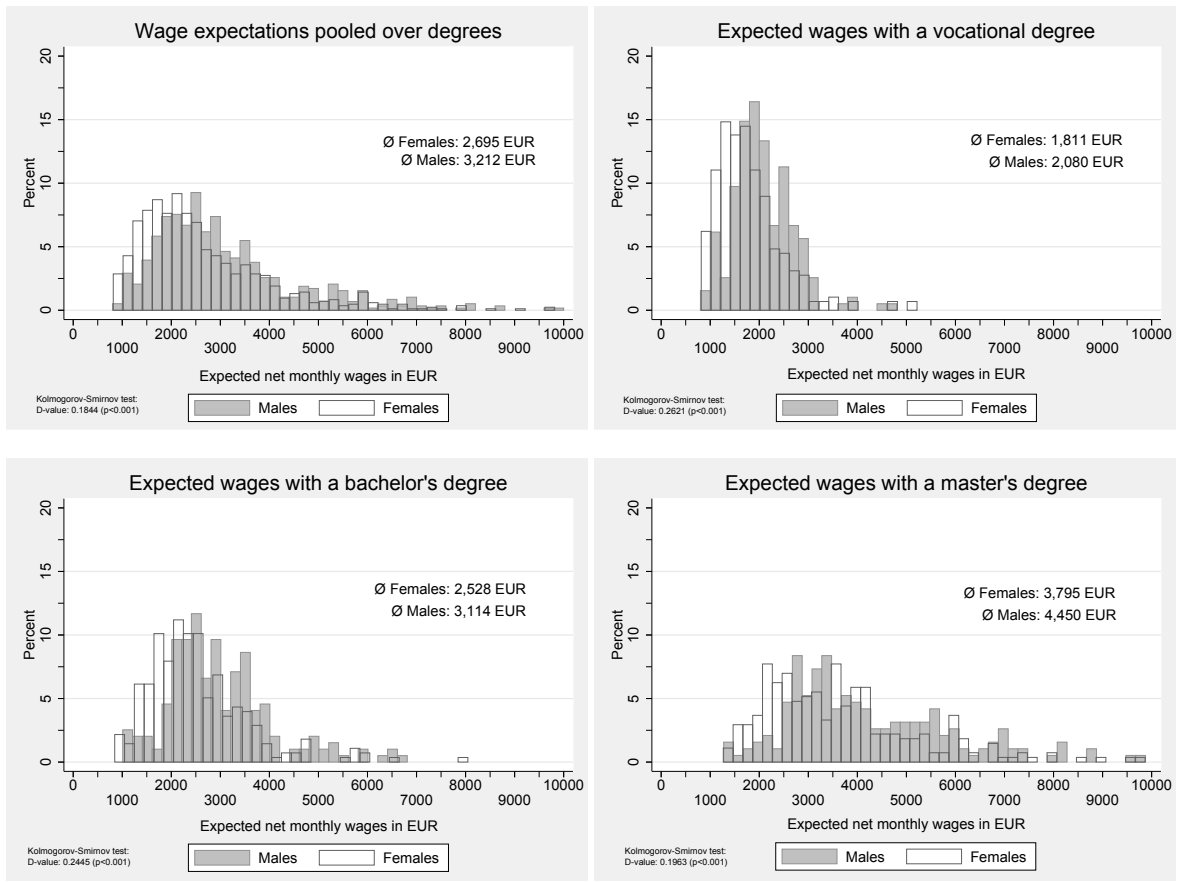
Notes: This table presents estimates of the Oaxaca-Blinder decomposition using different specifications and sample restrictions as indicated by row names. While column (1) reports the gender gap in expected wages in log points, column (2) transforms this gap in %. Column (5) indicates how much of the gender gap can be explained by including expected wage risk. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Are females better informed?

	Females	Males	Difference
Panel A: Students' average wage expectations (in EUR)			
...with a vocational degree	1,811	2,080	-269***
...with a Bachelor's degree	2,528	3,114	-586***
...with a Master's degree	3,795	4,450	-655***
Panel B: Actual average (population) wages (in EUR)			
...with a vocational degree	1,760	2,278	-518
...with a Bachelor's degree	2,070	2,739	-669
...with a Master's degree	2,323	3,092	-769
...with a tertiary degree	2,325	3,002	-677
Panel C: Share of students overestimating wages (in %)			
...with a vocational degree	44.5	32.3	12.2***
...with a Bachelor's degree	65.0	54.3	10.7**
...with a Master's degree	80.0	74.2	5.8
Panel D: Percentage deviations between expected and actual population wages			
...with a vocational degree	28.1	23.2	4.9**
...with a Bachelor's degree	36.9	33.2	3.7
...with a Master's degree	69.7	53.6	16.1**

Notes: This table presents differences between males and females. The difference in means is based on a two-sided t-test. The percentage deviation in Panel D is calculated as: $\Delta_i = (|y_{id}^e - \bar{y}_d^p|) / \bar{y}_d^p$, where y_{id}^e represent student i 's expected wage with education degree d and \bar{y}_d^p indicates actual average population wages with education degree d . Calculations are based on data from Best Up as well as from the German Micro Census (2010-2012). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1: Expected average wages by education scenario and gender



Notes: This figure shows the cross-sectional distribution of expected average wages with different educational degrees. Observations that fall in the upper or lower one percentile of the respective distributions are excluded. For illustration purposes wage expectations exceeding 10,000 EUR per month are not depicted.

Appendix A: Additional tables

Table A.1: Comparing non-respondents and respondents

	Full sample		Females		Males	
	Non-Response	Response (Difference)	Non-Response	Response (Difference)	Non-Response	Response (Difference)
Baseline characteristics						
Female	0.617	-0.050				
Migration background	0.519	-0.029	0.530	-0.023	0.500	-0.033
Non-academic fam.backgr.	0.643	-0.042	0.653	-0.045	0.628	-0.036
Academic high school	0.271	0.040	0.287	0.053	0.245	0.029
Integrated high school	0.378	-0.025	0.380	-0.043	0.374	-0.000
Vocational oriented high school	0.352	-0.016	0.333	-0.010	0.381	-0.029
Performance and cognitive skills						
Final high school GPA	2.569	-0.085**	2.542	-0.078	2.615	-0.103*
Verbal cognitive skills	9.423	0.701***	8.654	0.984***	10.671	0.089
Figural cognitive skills	10.802	0.222	10.734	0.454*	10.911	-0.104
Intended college major						
Language and Culture studies	0.089	-0.001	0.114	0.014	0.048	-0.012
Social Sciences	0.060	-0.008	0.089	-0.032	0.014	0.033*
Business and Economics	0.122	0.007	0.114	0.019	0.136	-0.011
STEM	0.216	0.043	0.101	0.062**	0.401	-0.017
Teaching	0.031	0.018	0.046	0.011	0.007	0.032*
Law and Management Sciences	0.063	0.001	0.080	-0.023	0.034	0.037
Medicine	0.107	0.010	0.139	0.013	0.054	0.017
Psychology	0.065	-0.011	0.084	-0.027	0.034	0.016
Arts and Sports	0.036	0.013	0.030	0.014	0.048	0.009
Others	0.026	0.000	0.034	-0.007	0.014	0.011
No coll.edu. intended	0.151	-0.065***	0.139	-0.050*	0.170	-0.088***
Missing information	0.034	-0.008	0.030	0.006	0.041	-0.027*
Career motives						
Extrinsically motivated	3.050	0.036	3.079	0.020	3.002	0.064
Intrinsically motivated	3.259	0.026	3.297	0.046	3.200	0.010
Socially motivated	2.903	-0.077*	2.994	-0.042	2.757	-0.098
Work-Life-Balance motivated	3.166	-0.024	3.200	0.020	3.111	-0.072
Personality traits						
Openness	4.967	0.110	5.000	0.155*	4.911	0.060
Extraversion	4.830	0.024	4.909	-0.027	4.698	0.118
Conscientiousness	4.901	-0.028	5.085	-0.049	4.594	0.063
Neuroticism	4.341	-0.117	4.601	0.025	3.904	-0.218*
Agreeableness	5.339	-0.102	5.405	-0.014	5.231	-0.199*
External Locus of Control	3.271	0.027	3.292	0.103	3.236	-0.070
Internal Locus of Control	5.377	-0.031	5.391	0.002	5.354	-0.071
Preferences						
Riskaversion	4.163	0.255*	4.407	0.140	3.766	0.483**
Patience	6.051	-0.062	5.833	0.016	6.400	-0.228
Time preference for present	3.402	-0.088	3.348	-0.053	3.490	-0.149
Confidence						
Confidence in own ability	0.197	0.055*	0.181	0.026	0.222	0.087*
Self-confidence	4.834	0.068	4.643	0.062	5.164	0.003
Educational aspiration						
Intended college enrollment	0.680	0.098***	0.692	0.091**	0.660	0.112**
<hr/>						
N	384	649	237	368	147	281
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 those who did not. Means and mean differences are based on a two-sided t-test.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Wage expectations for each education scenario

	Female	Male	Difference	p-value	N
Expectations with a vocational degree					
Expected wage (in EUR)	1810.61	2079.6	-269.04***	(0.000)	485
Expected wage risk	215.70	262.20	-46.50**	(0.012)	485
Expectations with a Bachelor's degree					
Expected wage (in EUR)	2528.22	3113.85	-585.63***	(0.000)	476
Expected wage risk	278.48	413.81	-135.32***	(0.000)	476
Expectations with a Master's degree					
Expected wage (in EUR)	3794.62	4449.73	-655.11***	(0.002)	469
Expected wage risk	484.03	686.60	-202.57***	(0.004)	469

Notes: This table presents gender differences in wage expectations. Wage risk is measured by the standard deviation of the individual specific expected wage distribution. Means and mean differences are based on a two-sided t-test. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Decomposition results for each education scenario

Expected wages	Pooled over degrees (1)	...with a Master's degree (2)	...with a Bachelor's degree (3)	...with a vocational degree (4)
Raw Gap	0.1770***	0.1612***	0.1979***	0.1552***
Explained Gap	0.1357***	0.1400***	0.1099***	0.0734***
Unexplained Gap	0.0413*	0.0212	0.0880***	0.0819***
% share explained by expected wage risk	76.66	86.85	55.53	47.29
N	1430	469	476	485

Notes: This table presents estimates of decomposing the gender gap in expected average (log) wages into one part that can be explained by differences in expected wage risk and an unexplained part based on the Oaxaca-Blinder decomposition. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Detailed decomposition results

	Contribution (1)	% of gap (2)
Difference	0.177	100
Explained		
Baseline characteristics	-0.0038	-2.12
Performance and cognitive skills	-0.0018	-1.04
Intended college major	0.0054	3.05
Career motives	0.0057	3.2
Personality traits	0.0089	5.03
Preferences	0.0001	0.07
Confidence	0.0064	3.61
Expected wage risk	0.1375	77.68
Total explained	0.1584	89.48

Notes: This table presents estimates of the detailed Oaxaca-Blinder decomposition using pooled coefficients as weighting scheme. The number of observations equals $N=1,430$. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix B: Relevance of wage expectations

The relevance of the gender gap in wage expectations depends on whether wage expectations are significantly associated with other relevant choices that can partly account for the actual gender wage gap. Lower wage expectations of female students may affect the actual gender wage gap through lower incentives to invest in education and/or through the formation of reservation wages which are likely to affect the decision to accept wage offers. While I cannot address the latter, I can complement existing evidence (see e.g. Hartog et al., 2012; Schweri and Hartog, 2015; Belfield et al., 2016; Attanasio and Kaufmann, 2014, 2017; Zafar, 2013; Arcidiacono et al., 2014; Wiswall and Zafar, 2016a; Ruder and Noy, 2017) and investigate whether wage expectations are related to (future) educational choices. For a sub-sample of students I can inspect whether wage expectations, or more specifically expected returns to college, are related to the decision to enroll in college. In addition, I can determine whether students who expect higher wages or who expect higher wage risk do indeed enroll in a college major that pays more or exhibit higher wage dispersion based on actual labor market data.³¹ Recall, that the information on students' wage expectations were elicited *before* students enrolled in college. Thus, ex-post rationalizations are of no concern.

Considering the decision to enroll in college, note first that in Germany, female high school graduates are less likely to enroll in college compared to their male counterparts. Among the 2015 cohort, for example, 71% of all male high school graduates enrolled in college (within one year after high school graduation), while among female high school graduates only 63% enrolled (Destatis, 2017). To determine whether wage expectations are relevant for the enrollment choice, I regress a binary variable indicating college enrollment on students' expected returns to college (defined as the difference between (log) expected wages with a Master's degree and (log) expected wages with a vocational degree)³², background characteristics, students' final high school GPA and two measures for cognitive skills. Table B.1 column (1) shows that higher expected returns predict college enrollment, i.e. students with higher expected returns to college are more likely to enroll in college. However, given the small sample size, the coefficient is only marginally statistically significant.

³¹Gender differences in college major choice are known to play an important role in explaining the actual gender wage gap (Machin and Puhani, 2003; Black et al., 2008).

³²In Germany, the majority of Bachelor's students continue to earn a Master's degree (e.g. Neugebauer et al., 2016). Thus, expected wages with a Master's degree should be most relevant for students' higher education choice.

Table B.1: Relevance of wage expectations

Dep.var:	Enroll in college (1)	Choose a high paying college major (2)	Choose a high risk college major (3)
Female	-0.084 (0.061)	-0.129** (0.048)	-0.120** (0.045)
Expected return to college	0.094* (0.047)		
Average expected wage		0.114* (0.065)	
Expected wage risk			0.031** (0.013)
Migration background	0.108** (0.043)	0.221*** (0.051)	0.229*** (0.040)
Non-academic fam.backgr.	0.007 (0.033)	-0.035 (0.043)	0.011 (0.056)
Attended high school type: (ref. vocational oriented high school)			
Academic high school	0.138** (0.060)	0.033 (0.068)	0.056 (0.078)
Integrated high school	0.136*** (0.043)	-0.055 (0.060)	-0.031 (0.069)
Final high school GPA (ranges from 1-4)	-0.268*** (0.031)	-0.020 (0.054)	0.014 (0.061)
Verbal cognitive skills (ranges from 0-20)	0.005 (0.027)	0.017 (0.027)	0.013 (0.027)
Figural cognitive skills (ranges from 0-20)	-0.033 (0.022)	0.049* (0.024)	0.042** (0.019)
Constant	1.168*** (0.109)	-0.288 (0.547)	0.080 (0.263)
Adj. R^2	0.1403	0.0712	0.07421
N	422	308	308

Notes: This table presents estimates from a linear probability model. Expected returns to college are calculated as the difference between (log) expected wages with a Master's degree and a vocational degree. Expected wages refer to wage expectations with a Master's degree and expected wage risk is measured by the (log) of the variance of the individual wage distribution conditional on earning a Master's degree (see Section 2 for details). Note that a higher final high school GPA indicates lower performance. Standard errors in parentheses are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In order to rank students' college major choice according to actual wages and wage risk, I use data from the German Micro Census (2005-2012), a representative data set covering one percent of the German population. Based on this data, I calculate average hourly wages for each college major.³³ For illustration purposes, I group college majors into high, medium, and low paying fields according to their rank in terms of hourly wages. I then match students' chosen study programs to the classification of majors as implemented in the Micro Census and generate an indicator variable that equals one if a student chooses a high paying major and zero otherwise.

To rank college majors in terms of their wage risk, in a first step I estimate a Mincer wage equation among college graduates by regressing monthly wages on age, age², a linear time trend, and a set of college major fixed effects. In a second step, I use the variance of the residuals within each college major as a measure for wage risk. Analogous to the

³³Note that in the German Micro Census individuals only report their net income within specific bins, which must not necessarily consist of labor market income. To account for this, I only include individuals who work full time and report that their main source of income derives from labor market income. Moreover, I use the midpoint of each bin to proxy actual wages and divide monthly wages by the hours an individual typically works during a month. I further restrict the sample to individuals with *Abitur*.

grouping of college majors according to wages, I similarly group college majors into high, medium, and low wage risk fields according to their rank in terms of wage risk. Again, I generate an indicator variable that equals one if a student chooses a high wage risk college major and zero otherwise.

Based on these two groupings of college majors, I regress students' college major choice on their (log) wage expectations with a Master's degree and include the same control variables as before. The results are presented in Table B.1. As can be seen in column (2), female students are around 13 percentage points (pp) less likely to choose a high paying college major and 12 pp less likely to choose to a high wage risk major (column 3). The coefficient on expected wages in column (2) and expected wage risk in column (3) are statistically significant and suggest that students with higher wage expectations are more likely to choose a high paying college major. Similarly, students who expect higher wage risk, are more likely to choose a high wage risk college major.

Overall, the analysis provides supporting evidence for the hypothesis that wage expectations contain information about future educational choices that are relevant for actual wage differences between genders in the labor market. Thus, analyzing the determinants of the gender gap in wage expectations can contribute to a better understanding of the actual gender wage gap.