

Essays on Gender Differences in Pay

Dissertation zur Erlangung des Doktorgrades der
Wirtschaftswissenschaften (Dr. oec.)

Fakultät Wirtschafts- und Sozialwissenschaften

Universität Hohenheim

Institut für Volkswirtschaftslehre

vorgelegt von

Sascha Satlukal

aus Stuttgart

2024

Date of Oral Examination: March 11, 2025

Dean: Prof. Dr. Jörg Schiller

Examination Chair: Prof. Dr. Alfonso Sousa-Poza

Supervisor and First Reviewer: Prof. Dr. Aderonke Osikominu

Second Reviewer: Prof. Dr. Thomas Beißinger

Acknowledgements

First and foremost, I would like to express my deep gratitude to my supervisor, Prof. Dr. Aderonke Osikominu. Thank you for inspiring me for econometrics and empirical economic research and for giving me the opportunity to write a doctoral thesis. Thank you for your invaluable guidance, encouragement, and unwavering support during my doctoral phase. Your expertise and insights have been instrumental in shaping this research.

I am also grateful to the second reviewer, Prof. Dr. Thomas Beißinger, for assessing my dissertation and for inspiring me for labor economics during my study time.

Moreover, I am indebted to my colleague, Prof. Dr. Ulrich Scheurle, for all his support, trust, and encouragement throughout the last years. I am deeply thankful for your help in all matters and for our countless discussions and coffee breaks together.

I would also like to thank my colleague Tim Ruberg for his help and support during our study time as well as our doctoral phase together.

I am equally thankful to Martina Rabe and Dr. Frauke Wolf for being dear colleagues and for their help in all matters in the last years.

Further, I am thankful to my former colleagues and co-authors, Stephanie Briel, Dr. Gregor Pfeifer, and Prof. Dr. Mirjam Stockburger. Your help and support at the start of my dissertation project was invaluable and your contributions and insights have significantly enhanced the quality of our work together.

I would also like to thank my co-author, Prof. Dr. Marina Töpfer, for her collaboration and support. I am thankful for your significant and valuable contributions to our research project and I have greatly appreciated our collaboration.

Finally, I would like to extend my deepest appreciation to my family, in particular to my wife, my children, and my parents. Your love, encouragement, and patience have been my greatest source of strength throughout this journey.

Thank you all for your significant contributions to my academic and personal growth.

Contents

List of Tables	iii
List of Figures	v
1 General Introduction	1
References	12
2 Gender Differences in Wage Expectations: The Role of Biased Beliefs	17
2.1 Introduction	18
2.2 Data and Descriptive Evidence	22
2.2.1 Data Set	22
2.2.2 Descriptive Statistics	23
2.3 Analyses of the Gender Gap in Wage Expectations	26
2.3.1 Conditional and Unconditional Quantile Regressions	26
2.3.2 Decomposition of Gender Gaps at Unconditional Quantiles	31
2.4 The Role of Biased Beliefs	36
2.4.1 Measuring Over- and Underplacement	37
2.4.2 Measuring Misperception of Labor Market Prospects	40
2.4.3 Effects on the Gender Gap in Expected Own Salary	41
2.5 Conclusion	44
References	45
Appendix	49
A Survey Questions Eliciting Salary Expectations	49
B Additional Empirical Evidence	50
C Empirical Support for the Linearity Assumption in the Decomposition Analyses	52
3 Gender Differences in Reservation Wages: New Evidence for Germany	53
3.1 Introduction	54
3.2 Estimation Method	58
3.2.1 Parameter of Interest	58
3.2.2 Linear Regression	61
3.2.3 Blinder-Oaxaca Decomposition	62
3.2.4 Augmented Inverse Probability Weighting	63
3.2.5 Data-Driven Model Specifications	65
3.3 Data, Sample and Descriptive Statistics	68
3.3.1 Reservation Wage Information and Sample Composition	68
3.3.2 Descriptive Statistics	72

3.4	Unexplained Gender Gap in Reservation Wages	76
3.4.1	Reservation Wage Determinants	77
3.4.2	Decomposing the Gender Gap	80
3.4.3	The Role of Various Estimators	83
3.5	Heterogeneity in the Gender Gap	86
3.5.1	Gender Reservation Wage Gaps across the Distribution	86
3.5.2	Gender Reservation Wage Gaps for Selected Subgroups	89
3.6	Conclusion	93
	References	94
	Appendix	101
4	Gender Differences in Labor Market Expectations and Wages	115
4.1	Introduction	116
4.2	Data	120
4.3	Bias in Labor Market Expectations	123
4.3.1	Job Loss and Job Finding Expectations	123
4.3.2	Predicting Job Loss and Job Finding Probabilities	124
4.3.3	Bias in Expected Probabilities	128
4.3.4	Predictive Power of Expectations	129
4.3.5	Gender Differences	130
4.4	Labor Market Expectations and Wages	133
4.4.1	Relating Bias in Labor Market Expectations to Wages	134
4.4.2	Contribution of Labor Market Expectations to Gender Gaps in Wages	138
4.5	Effect Heterogeneity	143
4.5.1	Collective or Individual Wage Agreement	143
4.5.2	Heterogeneity Across Age and Educational Attainment	145
4.5.3	Heterogeneity Across the Wage Distribution	150
4.6	Conclusion	152
	References	153
	Appendix	157
	A Additional Results	157
	B Robustness Checks	171
	C Predictors and Control Variables	175
5	General Conclusion	177

List of Tables

2.1	Descriptive Statistics	25
2.2	Gender Gap in Unconditional and Conditional Quantile Regressions . .	29
2.3	RIF Decomposition of the Gender Gap in Expected Own Salary	34
2.4	RIF Decomposition of the Gender Gap in Expected Average Salary . .	35
2.5	Over- and Underplacement by Gender	38
2.6	Over- and Underestimation of Average Salary by Gender	40
2.7	RIF Decomposition of the Gender Gap in Expected Own Salary	43
2.B.1	Gender Gap in Unconditional and Conditional Quantile Regressions (Excluding Observations with Missing High School GPA)	51
3.1	Variation in Sample Size after Imposition of Sample Restrictions	69
3.2	Summary Statistics by (Non-)Missing Reservation Wage	71
3.3	Summary Statistics	75
3.4	Reservation Wage Regressions, Selected Controls	78
3.5	Reservation Wage Decomposition, Male Reference Category	80
3.6	Estimates of Gender Gap in Reservation Wages across the Distribution, Baseline Model	87
3.A.1	Estimates of Unexplained Gender Gap in Reservation Wages with and without controlling for Family Background	101
3.A.2	Summary Statistics (Registered as Unemployed Sample)	101
3.A.3	Reservation Wage Regressions	101
3.A.4	Reservation Wage Decomposition, Female Reference Category	105
3.A.5	Estimates of Unexplained Gender Gap in Reservation Wages, Baseline Model Specification	106
3.A.6	Estimates of Unexplained Gender Gap in Reservation Wages, ML Model Specifications	106
3.A.7	Estimates of the Gender Gap in Reservation Wages (ML Model) and of the Aggregate RIF-Decomposition (Baseline Model) across the Dis- tribution	107
3.A.8	Estimates of Gender Gap in Reservation Wages within Subgroups . . .	108
4.1	Summary Statistics of Job Loss and Job Finding Probabilities	125
4.2	Predictive Power of Labor Market Expectations	130
4.3	Gender Difference in Job Loss Expectations	131
4.4	Gender Difference in Job Finding Expectations	132
4.5	Bias in Job Loss Expectation and Wages	135
4.6	Bias in Job Finding Expectation and Reservation Wages	137
4.7	Wage Decomposition	140

4.8	Reservation Wage Decomposition	142
4.9	Collective vs. Individual Wage Agreement	145
4.10	Wage Effect by Age and Education	146
4.11	Wage Decomposition by Age and Education	147
4.12	Reservation Wage Effect by Age and Education	148
4.13	Reservation Wage Decomposition by Age and Education	149
4.A.1	Stacking Weights used for Predictions	162
4.A.2	Wage Effect by Gender	165
4.A.3	Reservation Wage Effect by Gender	165
4.A.4	Collective vs. Individual Wage Agreement (With Personality Traits) . .	166
4.A.5	Wage Effect by Age and Education (With Personality Traits)	167
4.A.6	Wage Decomposition by Age and Education (With Personality Traits)	168
4.A.7	Reservation Wage Effect by Age and Education (With Personality Traits)	169
4.A.8	Reservation Wage Decomposition by Age and Education (With Personality Traits)	169
4.B.1	Gender Difference in Job Loss Expectations (With Exact Predicted Prob.)	171
4.B.2	Gender Difference in Job Finding Expectations (With Exact Predicted Prob.)	171
4.B.3	Bias in Job Loss Expectation and Wages (With Exact Predicted Prob.)	172
4.B.4	Bias in Job Finding Expectation and Reservation Wages (With Exact Predicted Prob.)	172
4.B.5	Wage Decomposition (With Exact Predicted Prob.)	173
4.B.6	Reservation Wage Decomposition (With Exact Predicted Prob.)	174
4.C.1	Predictors and Control Variables	175

List of Figures

2.1	Expected and Actual Starting Salaries	26
2.2	Over- and Underplacement by Field of Study and Gender	39
2.3	Over- and Underestimation by Field of Study and Gender	41
2.A.1	Survey Questions To Elicit Expected Starting Salaries	49
2.B.2	Densities of Expected Starting Salaries	50
2.B.3	Expected Own Starting Salary Above Actual Average Starting Salary .	50
3.1	Unexplained Gender Gap in Reservation Wages, Various Estimators . .	84
3.2	Raw and Unexplained Gender Gaps in Reservation Wage within Sub- groups	90
3.A.1	Unexplained Gender Gap in Reservation Wages, Various Estimators (Registered as Unemployed Sample)	109
3.A.2	Raw and Unexplained Gender Gaps in Reservation Wage within Sub- groups (Registered as Unemployed Sample)	110
3.A.3	Unexplained Gender Reservation Wage Gap AIPW Estimates With Various ML Methods	111
3.A.4	Distribution of Propensity Score with Baseline Model	111
3.A.5	Distribution of Propensity Score with ML	112
3.A.6	Distribution of Hourly Reservation Wage by Gender	112
3.A.7	Pattern of the Unexplained Gender Reservation Wage Gap across the Distribution, LRM	113
4.1	Distribution of Expected and Predicted Job Loss Probabilities	123
4.2	Distribution of Expected and Predicted Job Finding Probabilities . . .	124
4.3	Effect of Job Loss Bias Across Wage Distribution	150
4.4	Effect of Job Finding Bias Across Reservation Wage Distribution . . .	151
4.A.1	Distribution of Expected and Predicted Job Loss Probabilities, Females	157
4.A.2	Distribution of Expected and Predicted Job Loss Probabilities, Males .	157
4.A.3	Distribution of Expected and Predicted Job Finding Probabilities, Fe- males	158
4.A.4	Distribution of Expected and Predicted Job Finding Probabilities, Males	158
4.A.5	Job Loss Probabilities Within Subgroups	159
4.A.6	Job Finding Probabilities Within Subgroups	160
4.A.7	Job Loss Bias Within Subgroups	161
4.A.8	Job Finding Bias Within Subgroups	162
4.A.9	Predicted Job Loss Probabilities by Base Learner	163
4.A.10	Predicted Job Finding Probabilities by Base Learner	164
4.A.11	Effect of Job Loss Bias Across Wage Distribution (Full-Time Only) . .	170

4.A.12	Effect of Job Finding Bias Across Reservation Wage Distribution (Looking for Full-time Only)	170
--------	--	-----

Chapter 1

General Introduction

Achieving gender equality in the labor market has been on the political agenda for a number of decades. In this time, women have made remarkable advances in the labor market which has led to a narrowing of gender gaps in education, labor force participation, labor supply, and earnings (Goldin, 2014). Nevertheless, there still remain significant gender disparities in the labor market. In particular, the gender wage gap, i.e. the difference in average wages between women and men, persists (Blau and Kahn, 2017). According to the most recent estimates, women earn 20 percent less than men globally, 13 percent less in the European Union, and 18 percent less in Germany (International Labour Organization, 2018; Eurostat, 2021). The observed earnings differences between women and men are, therefore, still subject of intensive economic research and public debate. One reason for the focus on the gender wage gap is that it can be interpreted as a summary statistic of overall gender differences in work, since wages are shaped by workers' education, training, and work experience (Goldin, 2014). Analyses of gender wage gaps usually differentiate between raw gender gaps and unexplained gender gaps.¹ Whereas raw gender gaps constitute observed wage differentials, unexplained gender gaps are estimates of earnings differences between women and men with equal productivity. Most studies find unexplained gender gaps in wages that are smaller than the raw gaps, but are still substantial (Blau and Kahn, 2017). That is, women earn on average less than men with identical observable characteristics. Unexplained gender gaps are often interpreted as

¹Raw and unexplained gender gaps are also often denoted as unadjusted and adjusted gender gaps.

wage discrimination against women (i.e., unequal pay for equal work), but can also be an indicator for gender heterogeneity in unobserved wage determinants. Hence, there continues to be an intense discussion in the economic literature on what factors are responsible for the persistent wage disparity between men and women. This dissertation consists of three empirical studies that all contribute to this literature on gender differences in pay.

Explanations for the gender wage gap have traditionally concentrated on productivity differences due to the lower human capital accumulation of women. For a long time, women have considerably lacked men in educational attainment, labor force participation and work experience. Differences in human capital have, thus, been a major source of the gap in wages between women and men during that time. However, these gender differences do no longer exist in the same magnitude today. Amid the increasing labor market attachment of women in the last decades, gender gaps in work experience have closed substantially. In terms of educational attainment, women have even slightly overtaken men (Goldin, 2014). As a consequence, differences in human capital accumulation now only account for a small part of the gender wage gap (Blau and Kahn, 2017).

But, there still remain differences in the type of education women and men attain, which contribute to subsequent pay disparities. Gender differences in college major choice, for example, are shown to explain a large portion of the gender wage gap among college graduates (Black et al., 2008). Women are still underrepresented in the completion of college degrees in STEM (Science, Technology, Engineering, Mathematics) fields, in which on average higher wages can be earned than in other fields (Patnaik et al., 2021). Besides disparities in college major choice, women do also sort into lower-paid occupations than men afterwards, even among individuals with the same field of study (Sloane et al., 2021). Moreover, there also exist substantial gender differences in industry sorting. For instance, the empirical analysis of Blau and Kahn (2017) suggests that different sorting into occupations and industries between women and men together account for half of the gender wage gap in the US. Hence, gender differences in sorting are still important drivers of the gender wage gap.

Another important factor in explaining gender inequality in earnings is the negative effect of having children on wages of women, often referred to as child penalty of mothers. Since the arrival of children still has a much stronger impact on the working career of mothers than of fathers, motherhood leads to lower long-term earnings of women relative to men (Kleven et al., 2019). According to some studies, female and male wages are very similar at labor market entry and only start to diverge with the arrival of the first child (Bertrand et al., 2010; Kleven et al., 2019). Giving birth negatively affects earnings of women through career interruptions, shorter working hours, and less accumulated work experience (Bertrand et al., 2010). What is more, mothers may switch to more family-friendly jobs that offer more flexible working schedules but lower wages (Kleven et al., 2019). The role of shorter working hours and a more flexible working schedule of women is also highlighted by Goldin (2014), who shows that firms in many sectors have an incentive to disproportionately reward working long and particular hours. Thus, she argues that a change in the structure and remuneration of jobs is required to reach full gender convergence in the labor market.

More recently, researchers have also started to assess the role of gender differences in personality traits and noncognitive skills. Previous evidence shows that men and women differ substantially in psychological attributes, willingness to compete, negotiation skills, and risk aversion (Bertrand, 2011; Croson and Gneezy, 2009). All of these factors may directly or indirectly (through their effect on educational and occupational choices) affect the gender wage gap. Overall, the empirical results of such analyses suggest that personality traits and noncognitive skills have rather low to moderate effects on the gender wage gap (Blau and Kahn, 2017). A further aspect that is considered as explanation for gender differences in labor market outcomes are gender norms. Views on what women and men should do may still be shaped by the traditional family division of labor, and therefore may restrain the progress of women in the labor market (Bertrand, 2011). Gender norms are found to be predictive of female employment rates and gender wage gaps across countries (Fortin, 2005).

Chapter 2 and 4 of this dissertation focus on an additional factor that may contribute to gender differences in pay, namely the role of beliefs about ability and labor market prospects. It is well documented that in many settings people tend to be too confident, i.e. overconfident, about their ability, performance, or desirable personality traits (Santos-Pinto and de la Rosa, 2020). For instance, studies show that people are overconfident about their cognitive skills (Reuben et al., 2017) and work productivity (Hoffman and Burks, 2020). The psychological literature distinguishes between three forms of overconfidence (Moore and Healy, 2008). Whereas overestimation is defined as an exaggerated belief about one's absolute ability or performance, overplacement denotes an exaggerated belief about one's ability or performance relative to others. Overprecision, in contrast, is an excessive certainty about the accuracy of one's beliefs. The distinction into these three forms of overconfidence has also been adopted in the economic literature (e.g., Adamecz-Völgyi and Shure, 2022; Duttie, 2016). Previous research shows that women often exhibit a lower level of (over)confidence (e.g., Niederle and Vesterlund, 2007). That is, women tend to be less confident about their absolute and relative abilities compared to men. The lower confidence of females has potential implications for the gender wage gap, as it has been shown that it affects educational and occupational choices (Adamecz-Völgyi and Shure, 2022; Buser et al., 2014). Moreover, gender differences do not only appear in beliefs about ability and performance, but also directly emerge in individuals' expectations about their labor market prospects. Contrary to traditional models of the labor market, that usually impose rational expectations assumptions, there is growing empirical evidence that individuals often have biased beliefs about their labor market prospects (Mueller and Spinnewijn, 2023). For example, people tend to be too optimistic about their remaining unemployment duration (Spinnewijn, 2015), but are too pessimistic regarding their job security (Dickerson and Green, 2012). Likewise, students are shown to have distorted beliefs about their future wages (Cortés et al., 2023). Yet, also with regard to labor market prospects women are often less confident than men, which of course may be linked to their general lower level of (over)confidence. Female students and female

job seekers are significantly less optimistic about their future earnings than their male counterparts (e.g., Reuben et al., 2017; Dawson, 2017). Importantly, these differences in beliefs can lead to different labor market choices and can cause gender disparities in labor market outcomes. Cortés et al. (2023) show that as consequence of their lower earnings expectations female college graduates set lower reservation earnings and accept jobs earlier when searching for their first job. This in turn leads to lower starting salaries for females compared to males. Thus, the lower confidence of women regarding their earnings potential is partly responsible for the early-career gender wage gap among the highly educated. Chapter 2 provides an empirical study that assesses how biased beliefs about one’s own earnings potential relative to others affects the gender gap in wage expectations of prospective university students in Germany. In Chapter 4, I evaluate subjective beliefs about job insecurity of workers and about job finding chances of job-seekers in Germany. Furthermore, I show how gender differences in these beliefs are related to gender gaps in wages and reservation wages, respectively.

Alongside lower confidence in their own labor market prospects, women do also tend to be less optimistic about general economic outcomes. Having positive expectations about future events that are beyond one’s control has been referred to as optimism in the economic literature (Bjuggren and Elert, 2019; Jacobsen et al., 2014). In many domains women display a lower level of optimism than men. For instance, research suggests that women are on average less optimistic about the future state of the economy and the stock market (Jacobsen et al., 2014). Chapter 2 additionally studies the role of biased beliefs about average wage levels in explaining the gender gap in expected own wages of students. To do so, students’ elicited beliefs about average starting salaries of others in their study field are compared to actual starting salaries in the respective field.

The finding that women already expect lower earnings prior to (re-)entering the labor market than men has led researchers to assess the gender gap in wage expectations more closely. As expectations about future wages can become self-fulfilling, gender disparities in earnings expectations are of high relevance for the gender gap in actual pay (Reuben

et al., 2017). Having lower wage expectations may induce one to accept lower wage offers and to be more reluctant when negotiating for a higher salary. As discussed above, there is also first empirical evidence that lower expectations of females indeed lead to lower realized earnings (Cortés et al., 2023). Reuben et al. (2017) even argue that analyzing the source of the gender gap in wage expectations is potentially more helpful than looking at realized wages for the understanding of gender differences in educational and career choices. Interestingly, the size of the gender gap in wage expectations of college students found in previous studies resembles the observed gender gap in salaries among recent college graduates (Francesconi and Parey, 2018; Kiessling et al., 2024). This gives rise to the presumption that factors impacting the gender gap in actual pay are anticipated by individuals when forming their wage expectations. Hence, besides the lower confidence and optimism of females, other factors shown to be drivers of the gender wage gap may also be responsible for the gender inequality in wage expectations. Among other things, gender differences in wage expectations can reflect the fact that women tend to choose careers with lower earnings potential than men. As is the case for the gender gap in realized earnings, a considerable part of the differential in wage expectations can be attributed to different choices of college majors and different aspirations of occupation and industry (Kiessling et al., 2024). Moreover, women may already be aware of how their intentions regarding family formation and child bearing impacts their professional career before they enter the labor market. It has been shown that gender differences in expectations about child rearing and career interruptions can explain a significant portion of the gender gap in expected and realized wages (Chevalier, 2007; Fernandes et al., 2021). Likewise, the gender gaps in expected and realized wages are found to be related to women's stronger preferences for family-friendly jobs (Wiswall and Zafar, 2017). Chapter 2 examines the gender gap in wage expectations among prospective university students, while considering biased beliefs, college major choice, and attitudes towards the importance of income as explanatory factors.

Most studies analyzing wage expectations focus on young people with no previous

labor market experience. But, there is also some evidence that female job seekers of all ages expect lower wages than comparable men (Dawson, 2017). Further, previous research finds a gender gap in reservation wages of unemployed job seekers (e.g., Brown et al., 2011; Le Barbanchon et al., 2020). That is, women are willing to work at lower wages than men, which presumably is related to the gender gap in wage expectations. Since reservation wages play a key role in the job search process, differences in reservation wages may also transmit to pay differences. In fact, Caliendo et al. (2017) show that the gender difference in reservation wages of unemployed individuals can explain the majority of the subsequent gap in reemployment wages. Therefore, gender disparities in reservation wages are of significance for the understanding of the gender wage gap. Chapter 3 contributes an empirical study estimating the unexplained gender gap in reservation wages of job seekers in Germany. In Chapter 4, I investigate, among other things, how reservation wages and the corresponding gender gap are affected by perceptions of one's job finding chances.

Moreover, previous literature has tried to shed more light on the causes of the wage inequality between women and men by studying heterogeneity of the gender wage gap. First and foremost, there is heterogeneity with respect to age as the gender wage gap is shown to widen over the life-cycle (Goldin et al., 2017). In addition, earning differences between women and men depend on individual characteristics such as educational attainment, marital status, having children at home, and job-related characteristics (Bach et al., 2023). For example, the gender wage gap is larger among married individuals and parents. In Chapter 3 and 4 the gender gaps in reservation wages and wages are analyzed across various subgroups that are defined according to attributes like age, marital status, education, and labor market experience. Besides heterogeneity across individual characteristics, differences in the gender gap do also arise between different points in the wage distribution. In most countries the gender gap is particularly pronounced in the lower and/or upper tail of the wage distribution (Christofides et al., 2013). A widening of the gap at the bottom is often denoted as sticky floor effect in the literature, while a widening of the gap at the top is referred to as glass ceiling effect. In Germany, the gender

gap increases along the wage distribution, i.e. there is evidence for a glass ceiling effect (Collischon, 2019). However, when only considering college graduates previous research documents a sticky floor effect (Francesconi and Parey, 2018). Chapter 2 examines the gender gap at unconditional and conditional quantiles of the wage expectation distribution. Similarly, Chapter 3 and 4 present results across unconditional reservation wage and wage distributions.

Estimating and analyzing gender wage gaps requires suitable estimation approaches. The majority of studies on the gender wage gap use the decomposition method proposed by the seminal work of Blinder (1973) and Oaxaca (1973) or related decomposition techniques (see Fortin et al. (2011) for a review). The basic idea of this method is to decompose the difference in the mean of an outcome variable (e.g. wages) between two groups (e.g. males and females) into an explained and an unexplained part.² To do so, the outcome variable is first regressed on a set of observed explanatory variables within both groups. In the second step, the explained and unexplained part are then computed based on the two resulting coefficient vectors and the sample means of explanatory variables. The explained part represents the fraction of the outcome differential that can be attributed to group differences in the observed characteristics. The unexplained part, in contrast, represents the fraction that is due to different returns to observed characteristics and to differences in unobserved factors between the two groups. Beyond this partition of the outcome differential, the two components can be further broken down into the contributions of single variables or groups of variables. The application of this method for the analysis of gender wage gaps is attractive for several reasons. First, the aggregate decomposition, i.e. the partition into explained and unexplained part, provides an estimate of the unexplained gender wage gap. Second, the detailed decomposition, i.e. the partition into contributions of single or groups of variables, shows quantitatively how much of the gender wage gap can be explained by differences in certain characteristics. For instance, it can reveal how much of the gender wage gap is attributed to differences in labor market experience. In

²The explained and unexplained parts are also often denoted as ‘composition’ and ‘wage structure’ effects (Fortin et al., 2011).

addition, it indicates which characteristics are remunerated differently between women and men and how this affects the gender wage gap. In Chapter 3 and 4 this approach is used to decompose gender gaps in average reservation wages and wages.

Decomposition techniques can also be applied to various distributional measures such as unconditional quantiles (Fortin et al., 2011). Thus, just like mean differentials, gender differences at unconditional quantiles of wage distributions can be decomposed into explained and unexplained parts. A commonly used estimation approach for this task is the method proposed by Firpo et al. (2009, 2018). The method rests upon unconditional quantile regressions that use recentered influence functions of quantiles as outcome variables. It also allows one to break down the explained and unexplained part into contributions of single variables. This method is applied in all three studies of this dissertation to decompose gender gaps across wage distributions.

The unexplained component of the gender wage gap can also be interpreted as an average treatment effect on the treated (Fortin et al., 2011). Hence, researchers have alternatively considered estimators that come from the program evaluation literature for the estimation of unexplained gender wage gaps (e.g., Black et al., 2008; Ñopo, 2008). These include regression and matching methods as well as inverse probability weighting. Since most of these estimators are less reliant on functional form assumptions than standard decomposition methods, they lower the risk of misspecifications and biased estimates. Previous literature suggests that the estimator choice can significantly affect the magnitude of the estimated gender wage gap (Strittmatter and Wunsch, 2023). Chapter 3 provides a comparison of different estimates of the unexplained gender gap in reservation wages estimated by a linear regression, a standard decomposition, and augmented inverse probability weighting.

To obtain unbiased estimates of the unexplained gender wage gap it is necessary to control for all relevant wage determinants as well as for variables with marked gender differences. Most of the literature relies on economic reasoning and prior empirical evidence for the selection of the relevant control variables. But, manually selecting among

an increasing set of potential wage determinants can be burdensome. Moreover, it is challenging to specify the correct functional form, as the relationship between covariates and wages may be complex. Therefore, researchers have started to apply supervised machine learning methods that are capable of selecting wage determinants in a data-driven way. An estimation approach that uses machine learning methods for model selection is the post-double-selection procedure proposed by Belloni et al. (2014a,b). It selects the most important predictors of the outcome variable and the treatment variable using the least absolute shrinkage operator in a first step, and then estimates the treatment effect with a linear regression including the union of selected predictors as control variables in a second step. This approach has recently been used to estimate unexplained gender wage gaps and its heterogeneity (Bonaccolto-Töpfer and Briel, 2022; Bach et al., 2023). An alternative approach, that does not restrict the functional form in any way, is double/debiased machine learning introduced by Chernozhukov et al. (2017, 2018) which combines augmented inverse probability weighting with the use of flexible machine learning methods. In Chapter 3 the unexplained gender gap in reservation wages is estimated with the post-double selection procedure as well as with double/debiased machine learning.

The next three chapters present the empirical studies of this dissertation. In Chapter 2 I investigate together with my co-authors Stephanie Briel, Aderonke Oskiomini, Gregor Pfeifer, and Mirjam Stockburger how biased beliefs affect the gender gap in wage expectations. To address this research question we utilize a survey among prospective students at Saarland University in Germany. Among other things, the survey elicits respondents' expectations of their own starting salary as well as of the average starting salary of others in their field of study. We start our analysis by computing unexplained gender gaps at conditional and unconditional quantiles of the expected own salary and expected average salary distributions. Our results indicate marked gender gaps in expectations of the own and average salary, which are, however, more pronounced in case of the own salary. We also see that gender gaps are larger at the lower end of the wage expectation distributions. Next, we decompose the gender gaps at unconditional quantiles and demonstrate that the

choice of the study field can explain a considerable portion of the raw gaps. Finally, we construct two measurements of biased beliefs and study their role for the gender gap in the expected own salary. The first measurement confronts students' perceived earnings potential relative to others with their relative performance in secondary school. The second measurement compares students' expectations of average starting salaries with observed starting salaries of university graduates. We find that biased beliefs about the relative earnings potential and average salaries together can explain a substantial portion of the gender gap across the wage expectation distribution. Moreover, the unexplained gender gap in wage expectation becomes statistically insignificant once we control for biased beliefs. Thus, this study mainly contributes to the literature by showing that gender differences in beliefs about relative earnings potential and average salaries are major drivers of the gender gap in wage expectations.

The second study in Chapter 3 deals with the gender difference in reservation wages. Marina Töpfer and I estimate the unexplained gender gap in reservation wages of non-employed job seekers using survey data from the German Socio-Economic Panel. Based on the stated monthly reservation wages and intended working hours of survey respondents we first compute hourly reservation wages and document that females in our sample set lower reservation wages on average. To estimate the corresponding unexplained gender gap we apply different parametric and semiparametric estimators. Furthermore, we use both conventional and data-driven model specifications for the estimation. Therefore, a major contribution of our study is that we compare the results of different estimation approaches. Our different estimates of the unexplained gender gap range between 5 and 8 percent and are all statistically significant. Hence, our finding suggest that female job seekers set lower reservation wages than males with similar observed characteristics. In addition, we compute the gender reservation wage differential across the distribution and across various subgroups and find remarkable heterogeneity of the gender gap. The disparity in reservation wages between women and men is larger at the top of the reservation wage distribution, among the high-skilled, and among individuals who live in a household

with a child.

In the third research project, presented in Chapter 4, I examine gender differences in beliefs about job insecurity and job finding chances and their relation to gender gaps in wages and reservation wages. For this analysis I again use survey data from the German Socio-Economic Panel which regularly asks survey participants about their perception of the probability of losing or finding a job within the next two years. To compare the subjective beliefs of respondents to objective measures, I predict the respective probabilities at the individual level using a large set of predictors and machine learning methods. My results reveal that workers substantially overestimate the probability of losing their job, while job seekers slightly overestimate the probability of finding a job. But, women are significantly more pessimistic about their job security and their job finding chances than men. Next, I relate the elicited beliefs to wages of workers and reservation wages of job seekers. I find a negative effect of the perceived job loss probability on wages and a positive effect of the perceived job finding probability on reservation wages. In the final step, I estimate the contribution of the observed gender differences in beliefs to the gender gaps in wages and reservation wages. The decomposition results indicate positive contributions in both cases. Hence, my study shows that the gender differences in labor market expectations can be an explanatory factor for the gender wage gap. Moreover, I highlight that the effect of job loss expectations on wages is more pronounced for workers without a collectively agreed wage and for college graduates.

Finally, Chapter 5 concludes this dissertation with a summary of the main findings and a discussion of their implications.

References

- Adamecz-Völgyi, A. and Shure, N. (2022). The gender gap in top jobs – The role of overconfidence. *Labour Economics*, 79:102283.
- Bach, P., Chernozhukov, V., and Spindler, M. (2023). Heterogeneity in the US gender wage gap. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 187(1):209–230.

- Belloni, A., Chernozhukov, V., and Hansen, C. (2014a). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2):29–50.
- Belloni, A., Chernozhukov, V., and Hansen, C. (2014b). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2):608–650.
- Bertrand, M. (2011). Chapter 17 - New Perspectives on Gender. volume 4 of *Handbook of Labor Economics*, pages 1543–1590. Elsevier.
- Bertrand, M., Goldin, C., and Katz, L. F. (2010). Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors. *American Economic Journal: Applied Economics*, 2(3):228–55.
- Bjuggren, C. M. and Elert, N. (2019). Gender differences in optimism. *Applied Economics*, 51(47):5160–5173.
- Black, D. A., Haviland, A. M., Sanders, S. G., and Taylor, L. J. (2008). Gender Wage Disparities among the Highly Educated. *Journal of Human Resources*, 43(3):630–659.
- Blau, F. D. and Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3):789–865.
- Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources*, 8(4):436–455.
- Bonaccolto-Töpfer, M. and Briel, S. (2022). The gender pay gap revisited: Does machine learning offer new insights? *Labour Economics*, 78:102223.
- Brown, S., Roberts, J., and Taylor, K. (2011). The gender reservation wage gap: Evidence from British Panel data. *Economics Letters*, 113(1):88–91.
- Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, Competitiveness, and Career Choices *. *The Quarterly Journal of Economics*, 129(3):1409–1447.
- 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*, 136:161–173.
- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., and Newey, W. (2017). Double/Debiased/Neyman Machine Learning of Treatment Effects. *American Economic Review*, 107(5):261–65.

- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68.
- 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.
- Christofides, L. N., Polycarpou, A., and Vrachimis, K. (2013). Gender wage gaps, ‘sticky floors’ and ‘glass ceilings’ in Europe. *Labour Economics*, 21:86–102.
- Collischon, M. (2019). Is There a Glass Ceiling over Germany? *German Economic Review*, 20(4):e329–e359.
- Cortés, P., Pan, J., Pilossoph, L., Reuben, E., and Zafar, B. (2023). Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab*. *The Quarterly Journal of Economics*. qjad017.
- Croson, R. and Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature*, 47(2):448–74.
- Dawson, C. (2017). The upside of pessimism. Biased beliefs and the paradox of the contented female worker. *Journal of Economic Behavior and Organization*, 135:215–228.
- Dickerson, A. and Green, F. (2012). Fears and realisations of employment insecurity. *Labour Economics*, 19(2):198–210.
- Duttie, K. (2016). Cognitive Skills and Confidence: Interrelations with Overestimation, Overplacement and Overprecision. *Bulletin of Economic Research*, 68(S1):42–55.
- Eurostat (2021). Gender Pay Gap Statistics. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Gender_pay_gap_statistics. Accessed: 2024-01-12.
- Fernandes, A., Huber, M., and Vaccaro, G. (2021). Gender differences in wage expectations. *PLOS ONE*, 16(6):1–24.
- Firpo, S. P., Fortin, N. M., and Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, 77(3):953–973.
- Firpo, S. P., Fortin, N. M., and Lemieux, T. (2018). Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics*, 6(2).

- Fortin, N., Lemieux, T., and Firpo, S. (2011). Chapter 1 - Decomposition Methods in Economics. volume 4 of *Handbook of Labor Economics*, pages 1–102. Elsevier.
- Fortin, N. M. (2005). Gender Role Attitudes and the Labour-market Outcomes of Women across OECD Countries. *Oxford Review of Economic Policy*, 21(3):416–438.
- Francesconi, M. and Parey, M. (2018). Early gender gaps among university graduates. *European Economic Review*, 109:63 – 82.
- Goldin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4):1091–1119.
- Goldin, C., Kerr, S. P., Olivetti, C., and Barth, E. (2017). The Expanding Gender Earnings Gap: Evidence from the LEHD-2000 Census. *American Economic Review*, 107(5):110–14.
- Hoffman, M. and Burks, S. V. (2020). Worker overconfidence: Field evidence and implications for employee turnover and firm profits. *Quantitative Economics*, 11(1):315–348.
- International Labour Organization (2018). Global Wage Report 2018/2019: What lies behind gender pay gaps. Technical report.
- Jacobsen, B., Lee, J. B., Marquering, W., and Zhang, C. Y. (2014). Gender differences in optimism and asset allocation. *Journal of Economic Behavior and Organization*, 107:630–651.
- Kiessling, L., Pinger, P., Seegers, P., and Bergerhoff, J. (2024). Gender differences in wage expectations and negotiation. *Labour Economics*, 87:102505.
- Kleven, H., Landais, C., and Søgaaard, J. E. (2019). Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209.
- Le Barbanchon, T., Rathelot, R., and Roulet, A. (2020). Gender Differences in Job Search: Trading off Commute against Wage. *The Quarterly Journal of Economics*, 136(1):381–426.
- Moore, D. A. and Healy, P. J. (2008). The Trouble with Overconfidence. *Psychological Review*, 115(2):502–517.
- Mueller, A. I. and Spinnewijn, J. (2023). Chapter 22 - Expectations data, labor market, and job search. In Bachmann, R., Topa, G., and van der Klaauw, W., editors, *Handbook of Economic Expectations*, pages 677–713. Academic Press.

- Niederle, M. and Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics*, 122(3):1067–1101.
- Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14(3):693–709.
- Ñopo, H. (2008). Matching as a Tool to Decompose Wage Gaps. *The Review of Economics and Statistics*, 90(2):290–299.
- Patnaik, A., Wiswall, M., and Zafar, B. (2021). Chapter 15 - College Majors 1. In McCall, B. P., editor, *The Routledge Handbook of the Economics of Education*. Routledge.
- Reuben, E., Wiswall, M., and Zafar, B. (2017). Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender. *The Economic Journal*, 127(604):2153–2186.
- Santos-Pinto, L. and de la Rosa, L. E. (2020). *Overconfidence in Labor Markets*, pages 1–42. Springer International Publishing, Cham.
- Sloane, C. M., Hurst, E. G., and Black, D. A. (2021). College Majors, Occupations, and the Gender Wage Gap. *Journal of Economic Perspectives*, 35(4):223–48.
- Spinnewijn, J. (2015). Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs. *Journal of the European Economic Association*, 13(1):130–167.
- Strittmatter, A. and Wunsch, C. (2023). Labor Market Sorting and the Gender Pay Gap Revisited. Mimeo, available at <http://dx.doi.org/10.2139/ssrn.4624731>.
- Wiswall, M. and Zafar, B. (2017). Preference for the Workplace, Investment in Human Capital, and Gender*. *The Quarterly Journal of Economics*, 133(1):457–507.

Chapter 2

Gender Differences in Wage Expectations: The Role of Biased Beliefs

Stephanie Briel (University of Hohenheim)

Aderonke Osikominu (University of Hohenheim, CEPR, CESifo, IZA)

Gregor Pfeifer (University of Sydney, CESifo, IZA)

Sascha Satlukal (University of Hohenheim)

Mirjam Stockburger (Justus Liebig University Giessen)

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

JEL Codes: J16; D84; D91; C21

Keywords: Gender Pay Gap, Wage Expectations, Biased Beliefs, Decomposition Analysis, Conditional Quantile Regression, Unconditional Quantile Regression (RIF-Regression)

Acknowledgments: We are grateful for valuable remarks from Martin Biewen, Bernd Fitzenberger, and Anthony Strittmatter as well as seminar and conference participants at the University of Tübingen, the University of Hohenheim, the Journées LAGV, the International Workshop on Applied Economics of Education, the Econometrics Conference on “Economic Applications of Quantile Regressions 2.0”.

Status: Published - Briel, S., Osikominu, A., Pfeifer, G., Reutter, R., Satlukal, S. Gender differences in wage expectations: the role of biased beliefs. *Empirical Economics* 62, 187 - 212 (2021). Available at: <https://doi.org/10.1007/s00181-021-02044-0>.

2.1 Introduction

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

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

(grade point average above/below median) in secondary school. Secondly, we analyze how the gender gap is affected by misperceptions of the average starting salary. For this, we compare students' expected average salary with the actual average starting salary in their field of study.

Our results indicate a raw gender gap at all percentiles of the distribution of expected salaries. At the median, we find raw gender gaps of 17 and 12 percent for the expected own and expected average salary, respectively. Accounting for observed characteristics reduces the gender gaps substantially, however, particularly the gap in the expected own salary remains sizable and highly significant. We find larger negative effects of being female at the bottom than at the top of the distribution, indicating a sticky floor structure of the gender gap in wage expectations. Furthermore, our decomposition results show that differential sorting of men and women into fields of study is strongly related to the gender gap in expected own and expected average salaries. However, after additionally accounting for biased beliefs, the contribution of field of study to the explained gap in the expected own salary shrinks and becomes insignificant at most percentiles. Gender differences in our measures of over- and underplacement together contribute between 15 and 51 percent to the total gender gap. Even larger fractions of the total gender gap can be attributed to misperceptions of the average salary. Consequently, the unexplained part of the gender gap becomes insignificant after taking gender differences in biased beliefs into account.

Our paper contributes to three research areas: gender differences in pay, wage expectations, and biased beliefs. First, our study speaks to the literature analyzing the gender gap in actual pay (see, e.g., the meta-analysis by Weichselbaumer and Winter-Ebmer, 2005 for an overview). Many studies in this line of research document a U-shaped pattern of the gender pay gap along the conditional and unconditional distributions of earnings, meaning that the gender pay gap is particularly pronounced in the lower (sticky floor effect) and upper tail (glass ceiling effect) of wage distributions (Antonczyk et al., 2010; Christofides et al., 2013). Analyzing the gender wage gap in the entire working

population, Collischon (2019) documents a glass ceiling effect in Germany. In contrast, Francesconi and Parey (2018), who exclusively focus on university graduates in Germany, a group closely related to our sample, find larger gender gaps in the lower tail of the wage distribution. Focusing on prospective students, we are able to document that the gender pay gap is already present at an earlier point in the careers of young people: when high school graduates form their expectations about their future salary.

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

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

economic domains. Bjuggren and Elert (2019), for instance, document that men are more optimistic than women about the future state of the economy. In a similar vein, we compare the expected average starting salary between genders and show that male students are more optimistic about future wage levels in their field than their female counterparts.

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

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

The remainder of the paper is organized as follows. Section 2.2 describes the data and shows descriptive statistics. Section 2.3 presents and discusses our estimation results of the conditional and unconditional quantile regressions, as well as of the decomposition of

the gender gaps. Section 2.4 describes our measurements of over- and underplacement as well as of misperceptions of average salaries, and discusses their impact on the gender gap in expected own salaries. Section 2.5 concludes.

2.2 Data and Descriptive Evidence

2.2.1 Data Set

We use data from a survey of prospective students at Saarland University, Germany, who applied for enrollment in the academic years 2011 and 2012 (henceforth Student Survey). Prospective students submitting a complete application received an e-mail with the URL of the survey and were asked to fill in the questionnaire truthfully. In 2011, applicants for a program in Business Studies, Humanities, Law Studies, as well as Mathematics and Computer Science received the questionnaire; in 2012, also applicants for a program in Education and Medicine were surveyed. The sample consists of 2,061 students in total, which is a relatively large sample size compared with other studies examining students' wage expectations (e.g., Reuben et al., 2017; Webbink and Hartog, 2004). The data were first used by Klößner and Pfeifer (2019), who also validated the quality of the data.¹ As Klößner and Pfeifer (2019) document, Saarland University and its students are well representative of an average university and the corresponding average student body in Germany. Figures like student/teacher ratio, gender ratio, student age distribution, distribution of graduates across fields, distribution of gender across fields, number of exams passed, grades, duration of studies, etc., are all close to the German average. Saarland University also appears in the middle of international rankings across items such as teaching/learning, research, knowledge transfer, international orientation, and regional engagement.

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

The prospective students first had to state the field of study and the degree (Bachelor, Master, or State Examination) they applied for. They were also asked whether they intended to do a further degree afterwards (Master, Second State Examination, or Doctoral Degree) and with which degree they intended to earn their first salary. In the next part of the survey, the prospective students were asked to answer several questions regarding monthly gross salaries. First, they had to provide an estimate of the average starting salary of other persons graduating in the same field of study and with the same degree. Second, they had to state their own expected starting salary, again referring to their chosen field and intended degree. Figure 2.A.1 in the Appendix shows an extract of the questionnaire containing these two core questions.

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

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

2.2.2 Descriptive Statistics

Table 2.1 depicts means and standard deviations of wage expectations and several other relevant variables for male and female students, respectively.³ The last column of Table 2.1

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

³Figure 2.B.2 in the Appendix shows the densities of wage expectations.

also shows t -values for the test of equality of means between males and females.

Male students expect on average higher starting salaries both for themselves and for an average other person in the same field and with the same degree compared with female students. More specifically, males expect on average to earn 3,579 € a month at labor market entry, while the corresponding number for females is 3,015 €. With regard to average earnings of others in the same field of study and with the same degree, males state on average 3,571 €, which is significantly higher than the average female statement of 3,178 €. There is also a gender difference in the size of the expected own salary relative to expected average salary. Whereas the expected own salary is on average higher than the expected average salary for male students, the opposite is the case for female students.

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

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

Figure 2.1 displays mean values of expected own starting salaries and expected average starting salaries based on the Student Survey as well as actual average starting salaries differentiated by field of study.⁴ Wage expectations differ considerably across majors. While Mathematics and Computer Science students expect the highest starting salaries

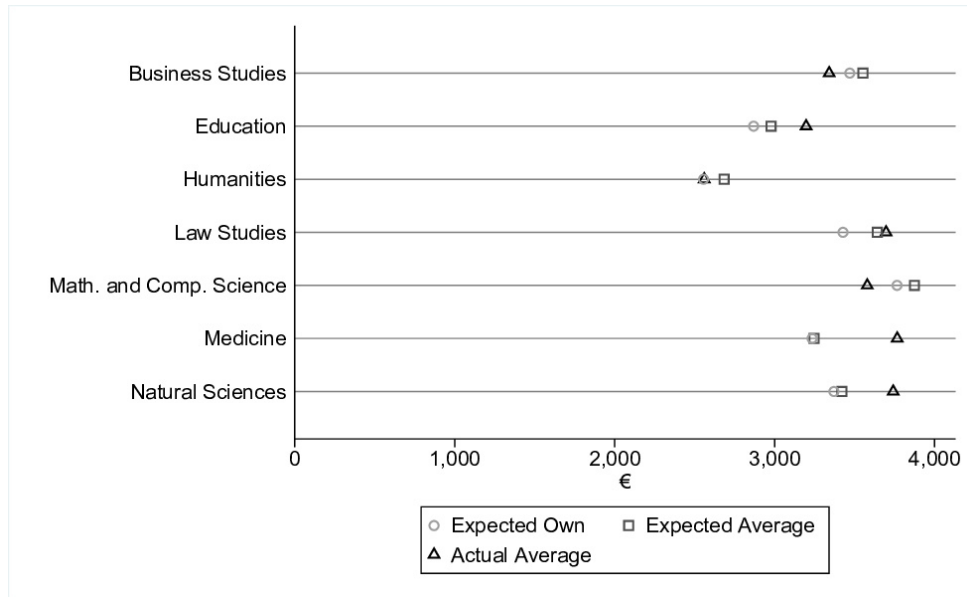
⁴Data on actual starting salaries are from PersonalMarkt Services GmbH that offers the largest database of actual salaries for Germany.

Table 2.1: Descriptive Statistics

	Males	Females	<i>t</i> -Statistic
(a) Expected Gross Monthly Salaries (Euro)			
Own Starting Salary	3,579.12 (1,908.16)	3,014.89 (1,558.45)	7.14
Average Starting Salary	3,571.00 (1,839.61)	3,178.37 (1,620.37)	5.02
(b) Personal Characteristics			
Age (Years)	21.22 (3.85)	20.78 (3.50)	2.61
GPA Secondary School (Scale from 1 to 4)	2.31 (0.60)	2.16 (0.59)	4.91
Relevant Work Experience (Months)	4.36 (17.30)	4.23 (16.22)	0.17
(c) Field of Study (Percent)			
Business Studies	29.77 (45.75)	19.61 (39.72)	5.25
Education	6.32 (24.35)	11.21 (31.56)	-3.95
Humanities	9.43 (29.23)	15.03 (35.75)	-3.90
Law Studies	12.18 (32.73)	14.09 (34.81)	-1.27
Math./Comp. Science	16.21 (36.87)	3.57 (18.55)	9.28
Medicine	18.85 (39.13)	29.03 (45.41)	-5.43
Natural Sciences	7.24 (25.93)	7.47 (26.30)	-0.20
(d) Importance of Expected Income for Choice of Major (Percent)			
Very Strong	2.18 (14.62)	1.78 (13.24)	0.64
Strong	16.09 (0.37)	14.43 (35.16)	1.03
Neutral	37.93 (48.55)	37.01 (48.30)	0.42
Low	17.13 (37.70)	18.34 (38.71)	-0.71
Very Low	26.67 (44.25)	28.44 (45.13)	-0.89
(e) Importance of an Above-Average Salary (Percent)			
Very Important	8.97 (28.59)	5.01 (21.82)	3.41
Important	42.30 (49.43)	35.14 (47.76)	3.28
Neutral	35.29 (47.81)	42.44 (49.45)	-3.00
Unimportant	8.51 (27.91)	11.88 (32.37)	-2.53
Very Unimportant	4.94 (21.69)	5.52 (22.84)	-0.58
Observations	870	1,178	

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

Figure 2.1: Expected and Actual Starting Salaries



Notes: Green circles show mean values of expected own starting salaries. Orange squares show mean values of expected average starting salaries of students in the same field of study. Red triangles show actual average starting salaries. *Source:* PersonalMarkt Services GmbH and Student Survey, own calculations.

for themselves and for average others, Education and Humanities students expect the least. The comparison of expected starting salaries with actual starting salaries reveals that prospective students make fairly small estimation errors on average.⁵ Figure 2.B.3 in the Appendix shows that across all fields more men than women expect their own salary in excess of the actual average salary.

2.3 Analyses of the Gender Gap in Wage Expectations

2.3.1 Conditional and Unconditional Quantile Regressions

Since gender gaps in actual salaries vary along the wage distribution (see, e.g., Collischon, 2019; Francesconi and Parey, 2018), we want to analyze the gender gap in students' expectations at different quantiles of the wage expectations distribution. Hence, we regress

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

students' expected own salary as well as their expected salary for average others on a female dummy and observed characteristics using conditional quantile regressions (CQR). Specifically, we model the logarithm of the respective expected salary (own or average) of individual i , y_i , as:

$$y_i = \mathbf{x}_i \boldsymbol{\beta}_\tau + u_{\tau i} \quad \text{with} \quad Q_\tau(y_i | \mathbf{x}_i) = \mathbf{x}_i \boldsymbol{\beta}_\tau, \quad (2.1)$$

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

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

$$\text{RIF}(y_i; Q_\tau) = Q_\tau + \frac{\tau - \mathbb{1}\{y_i \leq Q_\tau\}}{f_y(Q_\tau)}, \quad (2.2)$$

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

linear, we obtain a linear regression model

$$\text{RIF}(y_i; Q_\tau) = \mathbf{x}_i \boldsymbol{\gamma}_\tau + v_{\tau i}, \quad (2.3)$$

where $\boldsymbol{\gamma}_\tau$ denotes the coefficient vector and $v_{\tau i}$ the error term with $\mathbb{E}[v_{\tau i} | \mathbf{x}] = 0$. Standard ordinary least square (OLS) regressions of eq. (2.3) yield the UQR coefficient estimates that can be interpreted as partial effects on unconditional quantiles of y_i .

Our estimated gender gaps in wage expectations using conditional and unconditional quantile regressions are reported in Table 2.2. While Panel (a) displays the gender gaps in the expected own salary, Panel (b) shows the gender gaps in the expected salary for an average graduate in the same field of study and with the same degree. First and foremost, we find statistically significant raw gender gaps (RGG) in students' expected salaries for both, themselves and an average graduate, at all inspected quantiles of the distribution, as can be seen in the first row of Panels (a) and (b), respectively. Except for the first decile, the raw gender gap in students' expected own salary, reported in Panel (a), is always larger than the respective gap in their expected salary for an average graduate, reported in Panel (b). While the raw gender gap at the median in case of the expected own salary is 17 percent, the corresponding gap in case of the expected average salary is with 12 percent more than a quarter smaller. Interestingly, the raw gender gaps we find in students' expected own salaries are similar in size to the gender gap in *actual* salaries of recent university graduates in Germany (Francesconi and Parey, 2018).

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

Table 2.2: Gender Gap in Unconditional and Conditional Quantile Regressions

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

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

gender wage gap in the whole working population in Germany, in contrast, is increasing along the distribution (Collischon, 2019).

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

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

consistent with previous studies showing that women tend to be more pessimistic about future economic events (e.g., Jacobsen et al., 2014; Bjuggren and Elert, 2019).

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

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

2.3.2 Decomposition of Gender Gaps at Unconditional Quantiles

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

unconditional distribution of expected salaries.

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

$$\text{RIF}(y_{gi}; Q_{\tau g}) = \mathbf{x}_i \boldsymbol{\gamma}_{\tau, g} + v_{\tau gi} \quad \text{for } g = 0, 1. \quad (2.4)$$

The vector of explanatory variables, $\mathbf{x}_i$, does not include a female dummy now. After estimating Eq. (2.4) by OLS for both groups, the overall gender difference at the τ -th quantile, $\hat{\Delta}_O^\tau$, can be decomposed as,

$$\hat{\Delta}_O^\tau = \underbrace{\bar{\mathbf{x}}_1(\hat{\boldsymbol{\gamma}}_{\tau, 1} - \hat{\boldsymbol{\gamma}}_{\tau, 0})}_{\hat{\Delta}_U^\tau} + \underbrace{(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_0)\hat{\boldsymbol{\gamma}}_{\tau, 0}}_{\hat{\Delta}_E^\tau}, \quad (2.5)$$

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

If we want to interpret the explained and unexplained part of the decomposition truly as composition and wage structure effect, we need to assume that, at every quantile τ , the error term $v_{\tau g}$ is mean independent of gender, G , i.e.

$$\mathbb{E}[v_{\tau g} | \mathbf{x}, G] = \mathbb{E}[v_{\tau g} | \mathbf{x}] = 0, \quad \forall \tau \in [0, 1], \quad g = 0, 1. \quad (2.6)$$

This requires conditioning on all variables that are correlated with both the gender-specific outcome variables and gender. In the following analyses, we include a broad range of explanatory variables that we have selected based on economic reasoning and prior evidence on the determinants of expected starting salaries among prospective university students. Even though we do not have access to detailed controls for all underlying aspects of economic preferences and beliefs associated with both pay expectations and gender,

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

to satisfy Assumption (2.6) it would suffice to condition on variables that can proxy the underlying facets well enough to achieve conditional independence of group-specific outcomes and group membership. As an example, while we cannot directly control for competitiveness, we know that students who choose a program in Education and indicate that pay is not important to them are similarly competitive and typically less competitive than students majoring in Business who state that pay is important to them.⁷ Moreover, in Section 2.4, we consider a richer set of control variables that includes in addition measures of students' beliefs about their own earnings potential relative to others and the general earnings prospects in their field.

Table 2.3 contains the decomposition results at various unconditional quantiles of the expected own starting salary. While Panel (a) shows the aggregate decomposition results, detailed contributions of the most relevant covariates to the explained and unexplained part are given in Panels (b) and (c), respectively. According to Panel (a) and in line with the results in the previous section, a raw gender gap of 14 to 19 percentage points is present along the entire distribution of the expected own salary.⁸ The part of the total gap that can be explained by gender differences in the covariates considered varies between six and ten percentage points in absolute terms and is always statistically significant. In relative terms, the explained part ranges between 36 percent at the median and 75 percent at the first decile. The unexplained part ranges between three and twelve percentage points and is statistically significant except at the tenth percentile.

Panel (b) of Table 2.3 shows the detailed contributions of selected variables to the explained part of the gender gap in the expected own starting salary. Among the covariates considered, gender differences in field of study explain the largest part of the gender gap in the expected own salary: its contribution fluctuates between four and six percentage points across the quantiles considered and is statistically significant at all but the tenth

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

⁸Note that the total gender gaps in Tables 2.3 and 2.4 differ somewhat from the corresponding raw gender gaps in Table 2.2. The reason for this is that the gaps in Tables 2.3 and 2.4 are computed as the difference between the gender-specific means of the respective RIF that are linear approximations of the unconditional quantiles of male and female expected salaries.

Table 2.3: RIF Decomposition of the Gender Gap in Expected Own Salary

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

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

Table 2.4: RIF Decomposition of the Gender Gap in Expected Average Salary

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

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

percentile. In relative terms, differential sorting into field of study explains between 24 (25th quantile and ninth decile) and 41 (first decile) percent of the total gender gap. Gender differences in views on income importance explain up to two percentage points (seven percent) of the total gender gap. However, the effect is statistically significant only at the 75th and the 90th percentile. The contributions of other covariates (not reported in Table 2.3) are quantitatively less important and not statistically significant.

Table 2.4 shows the corresponding decomposition results at unconditional quantiles of the expected average starting salary. According to Panel (a), the total gender gap

in the expected average salary is with nine to 14 percentage points again smaller than the one in the expected own salary at every quantile considered. Between five and seven percentage points can be explained by gender differences in the included covariates. In relative terms, this corresponds to 36 to 74 percent of the total gap, depending on the quantile considered. The unexplained part lies between two and nine percentage points. While the explained gap is significant at four out of the five quantiles considered, the unexplained gap is significant only at the 25th and 50th percentile. The results of the detailed decomposition of the explained part, shown in Panel (b), reveal that gender differences in field of study contribute the most to explaining the gender gap in expected average salary. In relative terms, they explain between 25 and 54 percent of the total gender gap in the expected average starting salary, which is quantitatively similar to their contribution to the total gap in the expected own salary. Gender differences in views on income importance and other covariates (not reported in Table 2.4) do not contribute significantly to explaining the total gender gap.

All in all, the decomposition results match the patterns documented in the quantile regressions in Table 2.2: a substantial part of the raw gender gap in expected salaries can be attributed to differences in observed characteristics between men and women. Moreover, gender differences in field of study play the by far most important role for explaining the gender gap in expected salaries, which parallels findings in the literature on the gender gap in expected and actual pay (e.g., Collischon, 2019; Kiessling et al., 2019; Fernandes et al., 2020).

2.4 The Role of Biased Beliefs

While field of study or occupation have commonly been documented to be important drivers of the gender pay gap, they represent rather black-box controls for underlying aspects of preferences and beliefs that led to a particular choice of field of study or occupation. In the following we want to open this black box by exploring to what extent

gender differences in prospective students' salary expectations reflect gender differences in biased beliefs about their own relative earnings potential and the general earnings prospects in their field.

2.4.1 Measuring Over- and Underplacement

Our data allow us to compare the expectations of students about their own starting salary, conditional on field of study and intended degree, with their expectation about the average starting salary of other students in the same field of study and with the same intended degree. Specifically, in the survey, we first elicit a respondent's expectation about the average starting salary of graduates from the same field before we elicit the starting salary the respondent expects for herself (see Figure 2.A.1 in the Appendix). In so doing, we explicitly allow respondents to anchor the assessment of their own starting salary to their perception of the monetary success in the labor market of an average graduate with the same field of study and intended degree.

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

Table 2.5: Over- and Underplacement by Gender

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

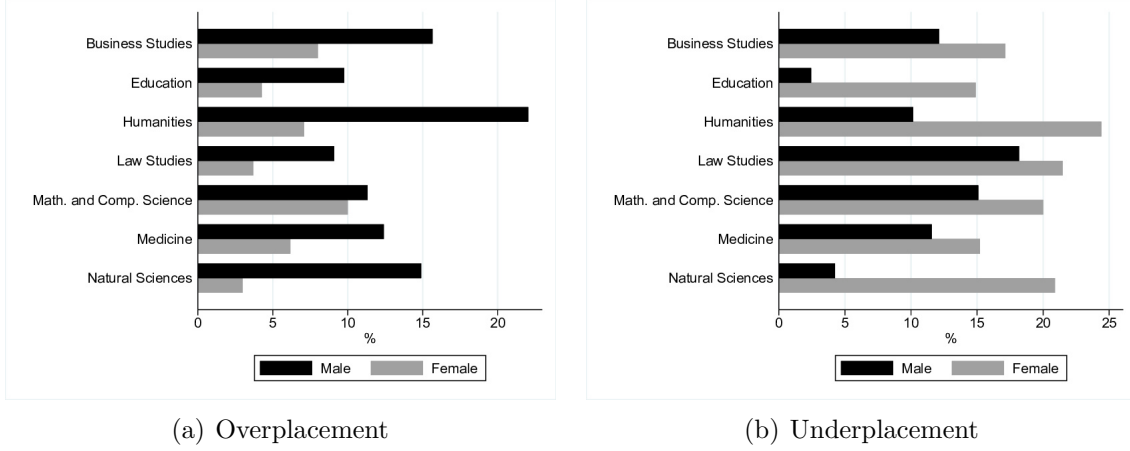
Notes: Means and standard deviations (in parentheses) of dummies for over- and underplacement. The last column shows *t*-values of a test of equality of means between males and females. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. *Source:* Student Survey,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

known by the students themselves, and it is also an important criterion for admission to the different university programs in Germany. To be precise, we define a prospective student as overplacing themselves who expects to earn more than the average graduate in the same program but whose GPA in secondary school is below the median in their chosen program. Analogously, we define a student as underplacing themselves who expects to earn less than average but has a GPA in secondary school above the median in their chosen program.

Table 2.5 displays the means of our over- and underplacement indicators in the full sample, as well as by gender. Female students are with 18 percent as opposed to 12 percent significantly more likely to underplace themselves with respect to their starting salary. Moreover, the share of male students overplacing themselves is with almost 14 percent more than twice as large than the corresponding share of female students with 6 percent. Similar patterns emerge within all fields of study, as can be seen in Figure 2.2. Within all fields, males are more likely to overplace themselves while females are more likely to underplace themselves.

A potential concern with our measurement of over- and underplacement is that, besides ability, there could be other reasons why a student expects to earn more or less than other graduates from the same program earn on average. Prospective students may value financial aspects associated with their study choice differently or desire to work in different industries with differing pay systems. However, in our decomposition of the gender gap,

Figure 2.2: Over- and Underplacement by Field of Study and Gender



Notes: Subfigure (a) shows the share of students who expect to earn more than average and have a final grade in secondary school that is worse than the median grade in the chosen field of study. Subfigure (b) shows the share of students who expect to earn less than average and have a final grade in secondary school that is better than the median grade in the chosen field of study. *Source:* Student Survey, own calculations.

we condition on a broad range of observed characteristics (e.g., income importance, desired branch of business). Hence, this potential issue is not problematic for our identification of the effect of over- and underplacement on the gender gap in expected salaries.

Even though we account for many observed characteristics, there might still be concerns that gender differences in unobserved determinants of wage expectations affect our measures of over- and underplacement. Given that females tend to work fewer hours during their lifetime, one might expect that they intend to work fewer hours after graduation and, consequently, expect lower starting salaries. The previous literature, however, suggests otherwise. Neither is there a substantial gender difference in working hours of recent university graduates (Francesconi and Parey, 2018), nor are female students less likely to expect to be in full-time employment directly after graduation (Fernandes et al., 2020). Moreover, previous research shows that non-cognitive skills and risk preferences do not contribute to the gender gap in expected starting salaries (Kiessling et al., 2019; Reuben et al., 2017). Therefore, we do not think that such factors have a major impact on the gender differences in our indicators for over- and underplacement with respect to starting salaries.

Table 2.6: Over- and Underestimation of Average Salary by Gender

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

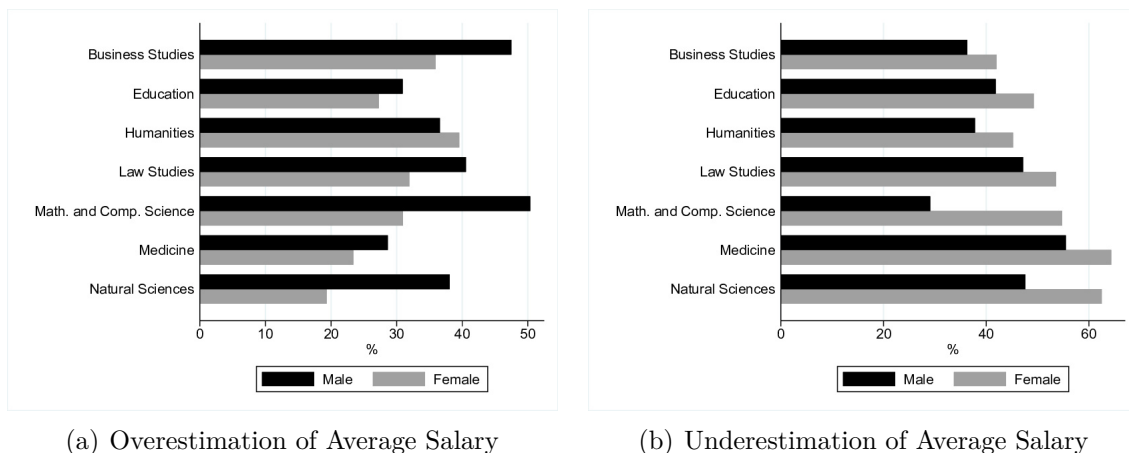
Notes: Means and standard deviations (in parentheses) of dummies for over- and underestimation of average salary. The last column shows *t*-values of a test of equality of means between males and females. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. *Source:*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

2.4.2 Measuring Misperception of Labor Market Prospects

Besides biased beliefs about the own salary in relation to the average salary, prospective students might also have misperceptions about the average starting salary in their chosen field of study. In Section 2.2.2, we have shown that students' wage expectations on average are fairly close to the actual salaries in most study fields. However, at the individual level, there might still be substantial differences between the expected average salary and the actual average salary. Hence, we compare students' expectations of the average starting salary with the actual starting salary in their field of study and analyze whether the incidence of biased beliefs about the general labor market prospects in the chosen field of study differs between male and female students. We construct measures of misperception of actual salaries that indicate whether a student over- or underestimates the average actual starting salary by more than ten percent, respectively.

Table 2.6 displays the means of the indicators for over- and underestimation of the average salary in the full sample, as well as by gender. Firstly, we see that a large number of prospective students either over- or underestimates the average salary by more than ten percent. Male students are with 41 percent as opposed to 30 percent of women significantly more likely to overestimate the average starting salary in their field of study. The share of female students underestimating the average starting salary in their field is with 53 percent twelve percentage points larger than the corresponding share of male students with 41 percent. Moreover, Figure 2.3 displays the share of students who over-

Figure 2.3: Over- and Underestimation by Field of Study and Gender



Notes: Subfigure (a) shows the share of students who overestimate the average salary by more than ten percent. Subfigure (b) shows the share of students who underestimate the average salary by more than ten percent. *Source:* Student Survey, own calculations.

and underestimate the average salary in each field of study. In all fields apart from Humanities male students are more likely to overestimate the average salary than female students. Females, in contrast, are more likely to underestimate the average salary within all fields.

2.4.3 Effects on the Gender Gap in Expected Own Salary

We have shown in Section 2.3 that the gender gap in the expected own salary is larger than the gap in the expected salary for average others. Further, the decomposition reveals that the unexplained part is particularly pronounced for the expected own salary. Thus, we will now study the role of biased beliefs in explaining the larger gap in the expected own salary. To be precise, we decompose the gender gap in the expected own salary along the unconditional distribution, additionally including the indicators for over- and underplacement as well as for over- and underestimation of the average salary as explanatory variables. Table 2.7 reports corresponding results. Since our measures of over- and underplacement are based on GPA in secondary school, we can only perform the analysis

for prospective students who reported their GPA. Therefore, we lose 479 observations.⁹

In line with the complete sample, the gender gap in the expected own starting salary is present at all studied quantiles and varies rather unsystematically: 12 percent at the first decile, 18 percent at the median, and 15 percent at the ninth decile. As can be seen in Panel (b) of Table 2.7, the contribution of overplacement is positive and significantly different from zero at every observed point along the wage distribution, with higher values in the tails of the distribution. Whereas gender differences in overplacement are responsible for 31 percent of the gap at the first decile and for 16 percent of the gap at the ninth decile, they are only responsible for 10 percent of the gap at the median. In contrast to overplacement, the impact of underplacement is significant only at the center of the distribution. At the median and the 75th quantile, the contribution of underplacement is a bit more pronounced than the contribution of overplacement. Thus, while gender differences in overplacement are more important at the bottom and top of the distribution, differences in underplacement play a more decisive role at the center. Overall, our findings are in line with Reuben et al. (2017) who document that gender differences in overconfidence are partly responsible for higher expected salaries of male students.

Furthermore, our results show that misperceptions of the average salary significantly affect the gender gap in the expected own salary. The contribution of gender differences in the propensity to overestimate the average salary is sizable and highly significant at the median and the upper tail of the distribution. Up to 49 percent of the gender gap can be attributed to overestimation of the average salary. Analogously, the contribution of underestimation of the average salary is sizable and highly significant at the lower tail of the distribution up to the median. At the first decile, the contribution of underestimation of the average salary even exceeds the total gender gap. This suggests that, while the stronger tendency of males to overestimate the average salary explains a substantial part of the gender gap in the upper half of the distribution, the stronger tendency of females to underestimate the average salary is responsible for a substantial fraction of the gap in

⁹Baseline results of the conditional and unconditional quantile regressions similar to Table 2.2 but excluding students with missing GPA are shown in Table 2.B.1 in the Appendix.

Table 2.7: RIF Decomposition of the Gender Gap in Expected Own Salary

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

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

the lower half. After including our indicators of biased beliefs, the contribution of field of study to the explained part is only statistically significant at the 75th quantile, while the contribution of income importance to the explained part turns insignificant along the entire distribution. This indicates that these variables are correlated with biased beliefs and have partly absorbed their influence on the total gender gap in our previous decomposition analysis in Section 2.3.2. In Figures 2.2 and 2.3, it can be seen that the incidence of over- and underplacement as well as over- and underestimation of the average salary varies substantially between fields. In line with this argument, other studies that do not account for biased beliefs find significant contributions of differential sorting of male and female students into field of study to the gender gap in wage expectations (Kiessling et al., 2019; Fernandes et al., 2020). Moreover, the unexplained part of the gender gap we find is no longer statistically different from zero at all quantiles considered. This result is in contrast to previous studies examining students' expected salaries, reporting larger and significant unexplained parts of the gender gap (Reuben et al., 2017; Kiessling et al., 2019; Fernandes et al., 2020). Thus, gender differences in biased beliefs seem to be important drivers of the gender gap in wage expectations and should be taken into account.

2.5 Conclusion

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

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

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

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

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

References

Antonczyk, D., Fitzenberger, B., and Sommerfeld, K. (2010). Rising wage inequality, the decline of collective bargaining, and the gender wage gap. *Labour Economics*,

17(5):835–847.

Arcidiacono, P., Hotz, V. J., Maurel, A., and Romano, T. (2020). Ex Ante Returns and Occupational Choice. *Journal of Political Economy*, forthcoming.

Astebro, T., Herz, H., Nanda, R., and Weber, R. A. (2014). Seeking the Roots of Entrepreneurship: Insights from Behavioral Economics. *Journal of Economic Perspectives*, 28(3):49–70.

Baker, R., Bettinger, E., Jacob, B., and Marinescu, I. (2018). The Effect of Labor Market Information on Community College Students’ Major Choice. *Economics of Education Review*, 65:18–30.

Barsky, R., Bound, J., Charles, K. K., and Lupton, J. P. (2002). Accounting for the Black–White Wealth Gap: A Nonparametric Approach. *Journal of the American Statistical Association*, 97(459):663–673.

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.

Bjuggren, C. M. and Elert, N. (2019). Gender differences in optimism. *Applied Economics*, 51(47):5160–5173.

Blau, F. D. and Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3):789–865.

Christofides, L. N., Polycarpou, A., and Vrachimis, K. (2013). Gender Wage Gaps, ‘Sticky Floors’ and ‘Glass Ceilings’ in Europe. *Labour Economics*, 21:86–102.

Collischon, M. (2019). Is There a Glass Ceiling over Germany? *German Economic Review*, 20(4):e329–e359.

Dale, S. and Krueger, A. (2002). Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables. *Quarterly Journal of Economics*, 117(3):1491–1527.

Duttle, K. (2016). Cognitive Skills and Confidence: Interrelations with Overestimation, Overplacement and Overprecision. *Bulletin of Economic Research*, 68(S1):42–55.

Fernandes, A., Huber, M., and Vaccaro, G. (2020). Gender Differences in Wage Expectations. Technical Report arXiv:2003.11496, arXiv.org.

Filippin, A. and Ichino, A. (2005). Gender Wage Gap in Expectations and Realizations. *Labour Economics*, 12(1):125–145.

- Firpo, S. P., Fortin, N. M., and Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, 77(3):953–973.
- Firpo, S. P., Fortin, N. M., and Lemieux, T. (2011). Decomposition Methods in Economics. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4, pages 1–102. Elsevier.
- Firpo, S. P., Fortin, N. M., and Lemieux, T. (2018). Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics*, 6(2).
- Francesconi, M. and Parey, M. (2018). Early gender gaps among university graduates. *European Economic Review*, 109:63 – 82.
- Jacob, B. A. and Wilder, T. (2011). Educational Expectations and Attainment. In Duncan, G. and Murnane, R., editors, *Whither Opportunity? Rising Inequality, Schools, and Children’s Life Chances*, page 133–162. New York: Russell Sage Press.
- Jacobsen, B., Lee, J. B., Marquering, W., and Zhang, C. Y. (2014). Gender differences in optimism and asset allocation. *Journal of Economic Behavior and Organization*, 107:630–651.
- Jensen, R. (2010). The (Perceived) Returns to Education and the Demand for Schooling. *The Quarterly Journal of Economics*, 125(2):515–548.
- Kiessling, L., Pinger, P., Seegers, P., and Bergerhoff, J. (2019). Gender Differences in Wage Expectations: Sorting, Children, and Negotiation Styles. CESifo Working Paper No. 7827, CESifo Munich.
- Klößner, S. and Pfeifer, G. (2019). The Importance of Tax Adjustments when Evaluating Wage Expectations. *The Scandinavian Journal of Economics*, 121(2):578–605.
- Koenker, R. (2005). *Quantile Regression*. Cambridge University Press, Cambridge (MA).
- Koenker, R. and Hallock, K. F. (2001). Quantile Regression. *Journal of Economic Perspectives*, 15(4):143–156.
- Larrick, R. P., Burson, K. A., and Soll, J. B. (2007). Social comparison and confidence: When thinking you’re better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes*, 102(1):76–94.
- Machado, J. A. F. and Mata, J. (2005). Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression. *Journal of Applied Econometrics*, 20(4):445–465.

- Melly, B. (2005). Decomposition of differences in distribution using quantile regression. *Labour Economics*, 12(4):577–590.
- Moore, D. A. and Healy, P. J. (2008). The Trouble with Overconfidence. *Psychological Review*, 115(2):502–517.
- Niederle, M. and Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics*, 122(3):1067–1101.
- Osikominu, A., Grossmann, V., and Osterfeld, M. (2020). Sociocultural Background and Choice of STEM Majors at University. *Oxford Economic Papers*, 72:347–369.
- PersonalMarkt Services GmbH (no date). <http://www.personalmarkt.de/de/>. Accessed on December 24, 2012.
- Reuben, E., Wiswall, M., and Zafar, B. (2017). Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender. *The Economic Journal*, 127(604):2153–2186.
- Ring, P., Neyse, L., David-Barett, T., and Schmidt, U. (2016). Gender Differences in Performance Predictions: Evidence from the Cognitive Reflection Test. *Frontiers in Psychology*, 7:1680.
- Webbink, D. and Hartog, J. (2004). Can Students Predict Starting Salaries? Yes! *Economics of Education Review*, 23(2):103–113.
- Weichselbaumer, D. and Winter-Ebmer, R. (2005). A Meta-Analysis of the International Gender Wage Gap. *Journal of Economic Surveys*, 19(3):479–511.
- Wiswall, M. and Zafar, B. (2015). How Do College Students Respond to Public Information about Earnings? *Journal of Human Capital*, 9(2):117–169.
- Wiswall, M. and Zafar, B. (2018). Preference for the Workplace, Investment in Human Capital, and Gender. *The Quarterly Journal of Economics*, 133(1):457–507.
- Zafar, B. (2013). College Major Choice and the Gender Gap. *The Journal of Human Resources*, 48(3):545–595.

Appendix

A Survey Questions Eliciting Salary Expectations

Figure 2.A.1: Survey Questions To Elicit Expected Starting Salaries

Estimated Starting Salaries

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

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

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

At labor market entry

After five years of work experience

Please estimate your own monthly gross salary after your graduation.

At labor market entry

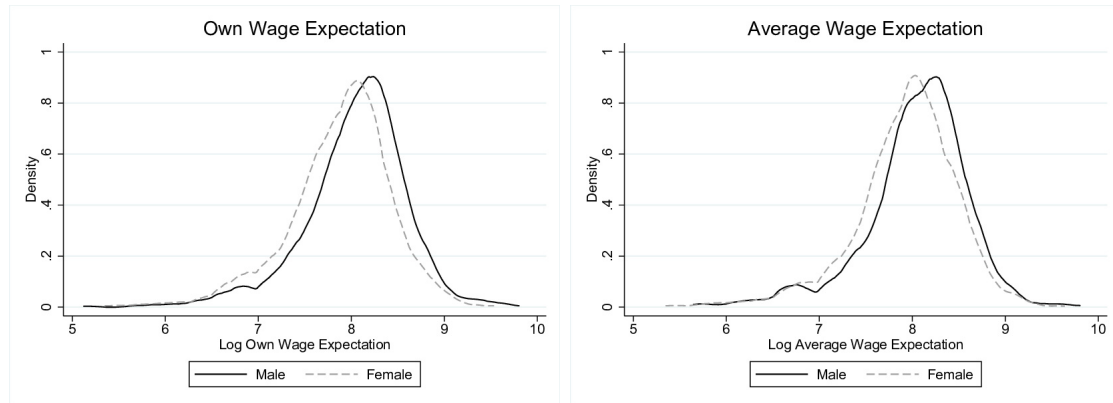
After five years of work experience

[Next](#)

Notes: Extract of the questionnaire the prospective students had to answer during the online survey, which focuses on estimated starting salaries.

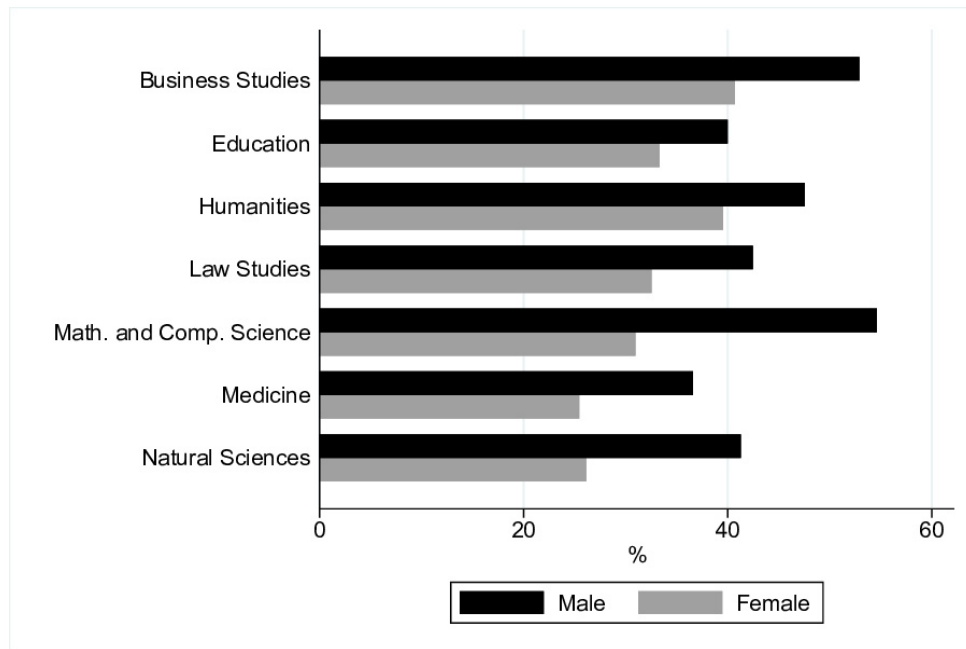
B Additional Empirical Evidence

Figure 2.B.2: Densities of Expected Starting Salaries



Notes: The left subfigure shows kernel density estimates of the logarithm of expected own starting salaries. The right subfigure shows kernel density estimates of the logarithm of expected average starting salaries. In both subfigures, we use an Epanechnikov kernel. *Source:* Student Survey, own calculations.

Figure 2.B.3: Expected Own Starting Salary Above Actual Average Starting Salary



Notes: Share of students for which the expected own starting salary exceeds actual average starting salaries in their chosen field of study. *Source:* Student Survey and PersonalMarkt Services GmbH, own calculations.

Table 2.B.1: Gender Gap in Unconditional and Conditional Quantile Regressions (Excluding Observations with Missing High School GPA)

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

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

C Empirical Support for the Linearity Assumption in the Decomposition Analyses

One potential concern w.r.t the decomposition of the unconditional quantiles is that consistent estimation of the explained and the unexplained part requires that the linearity assumption is satisfied in eq. (2.4) (Barsky et al., 2002; Firpo et al., 2011, 2018). If the relationship is nonlinear, the τ -th quantile of the counterfactual distribution of log expected salaries that women would have, if they formed their expectations in the same way as men, is not equal to $\mathbb{E}[\mathbf{x}|G = 1] \cdot \gamma_{\tau,0}$. Therefore, as an additional robustness check, we apply the reweighted regression approach described by Firpo et al. (2011). The idea of this approach is to reweigh the observations of male students so as to align the distribution of the characteristics of male students to that of female students. The reweighted regression approach allows us to estimate the specification error arising if the regression model is misspecified. A specification error close to zero indicates that the linear model is accurate. In addition, we use the method to calculate the reweighting error. If the reweighting function is consistently estimated, the reweighting error should be close to zero.

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

¹⁰The detailed results are available on request.

Chapter 3

Gender Differences in Reservation Wages: New Evidence for Germany

Sascha Satlukal (University of Hohenheim)

Marina Töpfer (Kore University Enna)

Abstract: Generally, women set lower reservation wages than men what may translate into substantial gender pay gaps in the labor market. Using, for the first time, a rich data set for Germany of non-employed individuals from the German Socio-Economic Panel 2007–2020, we analyze unexplained gender gaps in reservation wages. We compare both parametric and semiparametric estimators based on conventional and data-driven model specifications. The results suggest substantial unexplained gaps in reservation wages in favor of men (up to 8%). In addition, we show that the gaps are larger at the top of the reservation wage distribution as well as among individuals with children and with a high educational attainment. The estimates are robust across the various estimators and model specifications. These findings imply that pronounced unexplained gender gaps in reservation wages exist in Germany. As they are likely to result in actual gender pay gaps, gender gaps in reservation wages should be on the political agenda.

JEL Codes: J7; J16; J31

Keywords: Gender Reservation Wage Gap, Data-Driven Models, Robust Estimation

Acknowledgments: We thank Stephanie Briel and Aderonke Osikominu for helpful comments and suggestions. We are grateful to the participants of the 64th Annual Meeting of the Association of Italian Economists (SIE) 2023, the 38th Conference of the Italian Association of Labor Economists 2023, the 14th International German Socio Economic Panel User Conference 2022, and the IAAEU Workshop on Labour Economics 2022 for useful comments.

Status: Published - Bonaccolto-Töpfer, M., Satlukal, S. Gender differences in reservation wages: New evidence for Germany. *Labour Economics* 91, 102649, (2024). Available at: <https://doi.org/10.1016/j.labeco.2024.102649>.

3.1 Introduction

Considerable attention has been spent on closing gender differences in labor market outcome such as the gender pay gap. Nonetheless, substantial differences in pay persist worldwide (e.g., Blau and Kahn, 2017). An extensive literature concentrates on understanding the sources of gender inequality in realized wages, while a much smaller literature focuses on reservation wages. However, gender gaps do not only exist in realized pay. Previous research shows that females – compared to males – have lower wage expectations (e.g., Reuben et al., 2017; Briel et al., 2021) as well as lower reservation wages (e.g., Brown et al., 2011; Le Barbanchon et al., 2021). Similarly, the literature suggests that biased beliefs and risk attitudes affecting search behavior and gender gaps in reservation wages offer explanations for the gender gap in realized wages (Cortés et al., 2023, Caliendo et al., 2017). Hence, a better understanding of the gender reservation wage differential may help to explain the well-known phenomenon of persistent gender gaps in realized wages.

Typically, unexplained gender pay gaps are of prior interest (Blau and Kahn, 2017; Ñopo, 2008; Goldin, Claudia and Mitchell, Joshua, 2017; Bonaccolto-Töpfer and Briel, 2022, among others). Indeed, according to labor economic theory, gender differences in pay for individuals with similar characteristics should not exist in competitive markets. Nowadays men and women are relatively similar in many traditional characteristics such as education (Goldin, 2014; Blau and Kahn, 2017) and, thus, most part of the wage gap falls on the unexplained part (Bonaccolto-Töpfer and Bonaccolto, 2023). Unexplained gender differences in pay are often interpreted as wage discrimination against women, but they are also related to unobserved determinants of wages. For instance, gender differences in psychological and social factors, such as risk preferences, competitiveness, and bargaining attitudes may lead as well to gender-specific occupational choices and to different career paths (Bertrand, 2011). The above factors apply to both men and women in and out of employment. For instance, Cortés et al. (2023) show that risk preferences and prior beliefs explain a major part of gender gaps in reservation wages as well as early-career wages. Similarly, anticipated discrimination (Caliendo et al., 2017) and gendered

cultural and social norms (McGee and McGee, 2023) have an impact on the unexplained gap in reservation wages.

The problem is that we lack information on all the possible determinants associated with (potential) productivity (Blau and Winkler, 2022) and gender (Bonaccolto-Töpfer and Briel, 2022). Nevertheless, it is important to have an as good as possible estimate of the unexplained gap in reservation wages as it may adversely affect female decisions and behavior reinforcing traditional role models. The latter may result in persistent gender differences in labor market outcomes such as the gender pay gap in realized wages. Strittmatter and Wunsch (2023) suggest that methodological choices are at least as relevant for estimation of the unexplained gender pay gap as considerations about wage determinants. In order to account for both aspects (method and relevant confounders), we estimate the unexplained gap in reservation wages using several estimation methods and a rich data set. The scope of applying different model specifications and various estimators is to obtain more robust and more reliable estimates.

This paper analyzes the gender difference in reservation wages using a sample of non-employed individuals in Germany. We estimate the unexplained gender gap in reservation wages using both parametric and semiparametric estimators such as Linear Regression (LRM), Blinder-Oaxaca (BO) and Augmented Inverse Probability Weighting (AIPW). Apart from using a baseline model inspired by the literature, we apply the Post-Double-Selection (PDS) procedure proposed by Belloni et al. (2014a,b) for data-driven model specifications. The procedure is based on the Least Absolute Shrinkage and Selection Operator (LASSO) introduced by Frank and Friedman (1993) and Tibshirani (1996) which allows for flexible model specifications and offers a convenient way to conduct model selection in case of large numbers of potential controls. Further, we apply the double/debiased machine learning (DML) approach of Chernozhukov et al. (2017, 2018), which combines the AIPW estimator with the use of flexible machine learning methods. For the DML approach we use an ensemble method that optimally combines four machine learning methods (LASSO, Random Forests, Gradient Boosting, Neural Net).

As consistent estimation of the unexplained gender gap relies on the conditional independence assumption, it is important to flexibly control for all relevant confounders. Model specifications that are not flexible enough can lead to biased estimates even if the conditional independence assumption is fulfilled (Ahrens et al., 2024). We use machine learning methods to increase flexibility, as such methods allow us to include a large set of interaction terms and higher-order polynomials of control variables (LASSO) or do not rely on any functional form (Random Forest, Gradient Boosting, Neural Net). Therefore, machine learning methods can help to reduce confounding in the estimation by sufficiently accounting for nonlinear effects of the confounders.

To our best knowledge, this is the first paper that uses machine-learning methods and compares various estimators to estimate gender gaps in reservation wages. In addition, we study potential gender gap heterogeneity across the reservation wage distribution as well as across various subgroups (such as individuals with children or with many years of labor market experience). For the analysis across the distribution, we use linear unconditional quantile regression (Firpo et al., 2009).

Although the literature on the gender gap in reservation wages is relatively sparse compared to the vast empirical literature on the gender pay gap, some studies analyzing the gender reservation wage differential do exist. Brown et al. (2011), for instance, find a substantial gender gap in reservation wages of non-employed individuals using British panel data. Caliendo et al. (2017) show that a substantial part of the gender wage gap can be explained by reservation wages using a sample of newly unemployed individuals in Germany. The papers suggest unexplained gender gaps in reservation wages of 7% for the UK and of 5% for Germany (Brown et al., 2011; Caliendo et al., 2017, respectively). Humpert and Pfeifer (2013) study determinants of male and female reservation wages using German survey data and find that the presence of children in the household increases reservation wages of women but do not affect reservation wages of men. Le Barbanchon et al. (2021) find that women compared to their male counterparts have lower reservation wages and are more sensitive to commuting time than men using French administrative

data of unemployed individuals. McGee and McGee (2023) find *ceteris paribus* gender gaps in reservation wages of 14%-32% in laboratory experiments among Northern American university students and show that women's reservation wages are too low compared to the theoretically optimal values.

Our results, in line with the related literature, suggest a marked unexplained gender gap in reservation wages between 5%-8% (log approximation) in favor of men. The estimates of the gap are statistically different from zero and persist across the various estimators and model specifications. In particular, the LRM and BO estimates are almost insensitive to how the potential reservation wage determinants are included (conventionally by a human researcher or by an algorithm). Comparing the different estimators, we see that the AIPW estimate is largest in case of the baseline model specification, while the BO estimate is most pronounced in case of the data-driven model specification. Hence, conclusions drawn may change depending on the estimation method applied, while estimates are less sensitive to model selection procedures. Economically, our results imply that non-employed women set lower reservation wages than comparable men what may translate into actual gender pay gaps.

A major contribution of our study is that we can draw the above conclusions based on various estimators and different model selection procedures. Our findings from the decomposition and LRM indicate that men and women do not only differ in terms of observable characteristics, but also in terms of returns to these characteristics. For example, males and females differ in their returns to a college degree. Further, there is a gender difference in the effect of being registered as unemployed and receiving unemployment benefits on reservation wages. Our estimates of the unexplained gap across the reservation wage distribution and by subgroups suggest substantial heterogeneity in the gender gap across the distribution, by educational attainment, and between individuals living in a household with and without a child. The unexplained gender gap estimates are larger at the top of the unconditional reservation wage distribution than at the bottom. As a policy implication, our results suggest that particular focus should lie on well-educated mothers.

For instance, subsidised childcare or flexible working arrangements may be powerful tools to augment reservation wages and, presumably, realized wages mitigating gender pay gaps for this group.

The remainder of this paper proceeds as follows. In Section 3.2, we review estimation approaches typically used to estimate gender pay gaps and outline our empirical strategy. Section 3.3 presents the data set as well as descriptive statistics. Section 3.4 shows the main results of our empirical analysis, while Section 3.5 discusses our analysis across the distribution and across selected subgroups. Finally, Section 3.6 concludes.

3.2 Estimation Method

In this section, we outline our estimation approach. First, we define our parameter of interest using the potential outcomes framework (Rubin, 1974) and clarify the relevant identifying assumptions. Second, we describe the parametric and semiparametric estimators that we apply to estimate the unexplained gender gap in reservation wages. We use a similar set of estimators as Strittmatter and Wunsch (2023) use for estimation of the unexplained gender pay gap. Finally, we present our data-driven model specifications based on machine learning.

3.2.1 Parameter of Interest

We aim to estimate the unexplained gender gap in hourly reservation wages, i.e. the reservation wage difference of a women relative to a man holding observable characteristics constant. As pointed out by Fortin et al. (2011), the unexplained part in a wage decomposition may be interpreted as an Average Treatment effect on the Treated (ATT). Considering the group dummy gender as treatment variable, we can define our parameter of interest using the potential outcomes framework.¹

The observed log hourly reservation wage of individual i is y_i and d_{Fi} a gender dummy

¹See, e.g., Rubin (1974) for a description of the potential outcomes framework.

that equals one if the individual is female and zero otherwise. $\mathbf{x}_i$ denotes the vector of observed reservation wage determinants. Moreover, let y_F and y_M denote the potential female (treated) and male (untreated) log hourly reservation wages, respectively. The raw gender gap in reservation wages, $\Delta = \mathbb{E}(y|d_F = 1) - \mathbb{E}(y|d_F = 0)$, can then be expressed as

$$\Delta = \underbrace{\mathbb{E}(y|d_F = 1) - \mathbb{E}_{\mathbf{x}|d_F=1}[\mathbb{E}(y_M|\mathbf{x}, d_F = 1)]}_{\delta \quad \text{Unexplained Part/ATT}} + \underbrace{\mathbb{E}_{\mathbf{x}|d_F=1}[\mathbb{E}(y_M|\mathbf{x}, d_F = 1)] - \mathbb{E}(y|d_F = 0)}_{\eta \quad \text{Explained Part/Selection Bias}} \quad (3.1)$$

The first right-hand term in (3.1) defines the unexplained part of the gender gap in reservation wages and may be interpreted as the ATT of being female. The second right-hand term in (3.1) represents the explained part of the gender gap in the decomposition framework. This term is considered a selection bias resulting from observed confounders in the program evaluation literature (Fortin et al., 2011).

The problem is that y_M represents a counterfactual wage for women. We cannot observe reservation wages of females (i.e. reservation wages of the ‘treated’ group) in a counterfactual world where they are male (untreated). In order to be able to identify the unexplained gender gap, we need to make a conditional independence and common support assumption (e.g., Rosenbaum and Rubin, 1983; Fortin et al., 2011):

$$\mathbb{E}(y_M|\mathbf{x}, d_F) = \mathbb{E}(y_M|\mathbf{x}), \quad (3.2)$$

$$\Pr(d_F = 1|\mathbf{x}) < 1, \quad (3.3)$$

respectively. The conditional mean independence assumption (3.2) implies that the expected counterfactual reservation wage of women is equal to the expected observed reservation wage of men conditional on observed characteristics. That is, it is crucial for identification of the ‘true’ counterfactual reservation wage to condition on all relevant factors that are simultaneously correlated with gender and the male-specific reservation

wage. In the following analysis, we include a broad range of potential reservation wage determinants as explanatory variables. In the vector of control variables, $\mathbf{x}$, we include inter alia information on age, marital status, household income, children, education, labor market experience, last job, last wage, unemployment benefits, job search, and personality traits.

The common support assumption (3.3) implies that for each woman in the population there exists at least one man with the same observed characteristics. Under assumptions (3.2) and (3.3) the unexplained gender gap in reservation wages is identified as

$$\delta = \mathbb{E}(y|d_F = 1) - \mathbb{E}_{\mathbf{x}|d_F=1}[\mu_M(\mathbf{x})] \quad (3.4)$$

where $\mu_M(\mathbf{x}) \equiv \mathbb{E}(y|\mathbf{x}, d_F = 0)$.

One critique of conventional approaches to estimate unexplained pay gaps is that they may lack causal interpretability since they usually neglect potential confounders that are determined prior to birth (Huber, 2015).² Ignoring pre-birth confounders, and, hence, relying merely on controls determined later in life (i.e. after gender) may lead to endogeneity. For instance, variables defined before birth that reflect family background (such as parental education) are potential candidates to confound the gender variable and the explanatory variables. In our case, including parental education controls does not markedly change the estimation results.³

²Huber (2015)'s criticism relates to the BO method as well as to non-parametric (e.g., DiNardo et al., 1996; Nopo, 2008) and quantile decomposition approaches (e.g., Machado and Mata, 2005; Melly, 2005; Firpo et al., 2009; Chernozhukov et al., 2013; Firpo et al., 2018). See Huber and Solovyeva (2020) for an overview.

³Table 3.A.1 displays estimates of the unexplained gender gap in reservation wages with and without controlling for parental education.

3.2.2 Linear Regression

We start by estimating the unexplained gender gap in reservation wages, δ , based on the following linear regression model:

$$y_i = \beta_0 + \delta_{LRM} \cdot d_{Fi} + \mathbf{x}_i \boldsymbol{\beta} + u_i \quad (3.5)$$

where u_i represents an error term. The estimated coefficient of the female dummy, $\hat{\delta}_{LRM}$, is the Linear Regression Model (LRM) estimator of the unexplained gender gap in reservation wages. Note, however, that by estimating a linear regression model we replace the common support assumption (3.3) by an assumption about the functional form of the relationship between observed characteristics and the log hourly reservation wage. Hence, the linear-regression-based estimator is restrictive as the regression model in equation (3.5) does not allow for heterogeneous coefficients of the various reservation wage determinants between women and men. We assume, thus, that women weight the various determinants in the same way as men do, when setting their reservation wage. Moreover, it restricts the gender gap to be unrelated to the values of the observed characteristics. The latter implies, for instance, that the unexplained gender gap in reservation wages does not vary between different age groups. This may be too restrictive as female wage penalties change with educational groups (Gensowski, 2018) or having children (Wilde et al., 2010).

In order to learn more about sources of the gender reservation wage gap, we apply the Gelbach (2016) decomposition to the LRM. We implement the Gelbach decomposition as follows. First, we run a regression of log hourly wages on a female dummy only (what gives the raw gap). Second, we run the baseline specification (this yields the unexplained gap). The idea is to treat the set of baseline controls as omitted in the raw regression and to calculate the contribution of each set of controls (from the baseline specification) when going from the raw to the unexplained gap.⁴ A main advantage of this approach is that – as opposed to sequentially adding different sets of controls – it is robust to sequential-order

⁴The Gelbach (2016) decomposition is based on the formula for omitted variable bias.

bias. Indeed, Gelbach (2016) shows that sequentially adding sets of controls can lead to very different conclusions about how much different covariates contribute depending on their order of addition. The Gelbach decomposition yields an unambiguous quantification of the portion of the gender reservation wage gap due to each control variable or each set of controls of interest.

3.2.3 Blinder-Oaxaca Decomposition

Next, we estimate the unexplained reservation wage gap δ using a Blinder-Oaxaca (BO) decomposition (Blinder, 1973; Oaxaca, 1973), which can be implemented in two steps.⁵ First, we estimate the linear reservation wage model separately for men (M) and women (F):

$$y_{ig} = \beta_{0g} + \mathbf{x}_i \boldsymbol{\beta}_g + u_{ig} \quad \text{for } g = M, F \quad (3.6)$$

Second, after estimating the gender-specific reservation wage models, we can decompose the mean difference in log hourly reservation wages as

$$\hat{\Delta} = \bar{y}_F - \bar{y}_M = \underbrace{(\hat{\beta}_{0F} - \hat{\beta}_{0M}) + \bar{\mathbf{x}}_F(\hat{\boldsymbol{\beta}}_F - \hat{\boldsymbol{\beta}}_M)}_{\hat{\delta}_{BO}} + \underbrace{(\bar{\mathbf{x}}_F - \bar{\mathbf{x}}_M)\hat{\boldsymbol{\beta}}_M}_{\hat{\eta}} \quad (3.7)$$

where $\hat{\beta}_{0g}$ and $\hat{\boldsymbol{\beta}}_g$ are the estimates of the gender-specific intercept and coefficient vector, while $\bar{y}_g$ and $\bar{\mathbf{x}}_g$ represent the average hourly log reservation wages and average characteristics in the male and female subsamples.⁶ The first decomposition term, $\hat{\delta}_{BO}$, captures gender differences in the coefficient vectors and is the BO estimator of the unexplained gender gap in reservation wages. The second decomposition term, $\hat{\eta}$, is the estimated explained part of the gender gap that captures gender differences in observed characteristics. The two decomposition terms can be further decomposed into the contribution of

⁵The BO decomposition can be seen as a particular way to obtain the Gelbach decomposition (see Appendix C of Gelbach (2016) for details).

⁶Note that we use the male coefficient vector as the counterfactual coefficient vector of women in (3.7).

each covariate. Moreover, the unexplained part can alternatively be expressed as

$$\hat{\delta}_{BO} = \frac{1}{N_F} \sum_{i=1}^N d_{Fi}(y_i - \hat{\mu}_M(\mathbf{x}_i)), \quad (3.8)$$

where $N_F = \sum_{i=1}^N d_{Fi}$ and $\hat{\mu}_M(\mathbf{x}_i) \equiv \hat{\beta}_{0M} + \mathbf{x}_i \hat{\beta}_M$. This expression shows, that $\hat{\delta}_{BO}$ is simply the mean difference between the observed reservation wage of women and their counterfactual reservation wage that is predicted based on the estimated male reservation wage model. While the BO estimator does allow for heterogeneous coefficient vectors and heterogeneous unexplained gender gaps, it still imposes the restrictive functional form assumption (3.6). However, as Kline (2011) shows, the BO estimator also constitutes a propensity score reweighting estimator and thus is doubly robust in the sense of Robins et al. (1994). That is, it is consistent if either the propensity score model or the reservation wage model is correctly specified. In contrast to the augmented inverse probability weighting estimator, which we describe below, the BO estimator is based on a linear model for the conditional odds of being treated. Hence, the implicit weights of BO can become negative.

3.2.4 Augmented Inverse Probability Weighting

Finally, we estimate δ using Augmented Inverse Probability Weighting (AIPW), which was first proposed by Robins et al. (1994) and presents an adjustment of the Inverse Probability Weighting estimator (IPW) (Hirano et al., 2003). The major advantage of AIPW over IPW and other estimators is that it is doubly robust, i.e. consistency of AIPW requires only either the propensity score model (represented in equation (3.9)) or the reservation wage model (3.6) to be correctly specified.

The related estimation approach is implemented in three steps. First, we estimate the conditional probability of being female based on the following propensity-score model:

$$p(\mathbf{x}_i) \equiv \Pr(d_{Fi} = 1 | \mathbf{x}_i) = G(\mathbf{x}_i \boldsymbol{\gamma}), \quad (3.9)$$

where we use the logistic distribution function as the binary link function $G(\cdot)$. We then predict $\hat{p}(\mathbf{x}_i)$ for all observations based on model (3.9). Second, we estimate the reservation wage model (3.6) in the male subsample and predict $\hat{\mu}_M(\mathbf{x}_i)$ for all observations. Third, we estimate the unexplained gender gap as

$$\hat{\delta}_{AIPW} = \underbrace{\frac{1}{N_F} \sum_{i=1}^N d_{Fi} \cdot y_i - \sum_{i=1}^N \hat{w}_i^M \cdot y_i}_{\hat{\delta}_{IPW}} - \underbrace{\left(\frac{1}{N_F} \sum_{i=1}^N d_{Fi} \cdot \hat{\mu}_M(\mathbf{x}_i) - \sum_{i=1}^N \hat{w}_i^M \cdot \hat{\mu}_M(\mathbf{x}_i) \right)}_{\text{Augmentation to IPW}} \quad (3.10)$$

where

$$\hat{w}_i^M = \frac{(1 - d_{Fi})\hat{p}(\mathbf{x}_i)}{1 - \hat{p}(\mathbf{x}_i)} \bigg/ \sum_{i=1}^N \frac{(1 - d_{Fi})\hat{p}(\mathbf{x}_i)}{1 - \hat{p}(\mathbf{x}_i)} \quad (3.11)$$

are the normalized IPW weights. Normalization of the IPW weights (denominator in (3.11)) ensures that the sum of the weights equals one in the male subsample and improves finite sample performance (Busso et al., 2014; Imbens, 2004). However, due to instability of the IPW estimator, calculating δ by (3.10) can be problematic when the conditional probability of being a woman is very close to one for some men (Khan and Tamer, 2010). To account for this, we impose a trimming rule and set all propensity score values below 0.01 and above 0.99 to 0.01 and 0.99, respectively. The same trimming rule has also been used by Chernozhukov et al. (2018) in their empirical application.⁷

The AIPW estimator makes a finite sample adjustment to the IPW estimator (first term in (3.10)) by subtracting the difference in the sample mean of $\hat{\mu}_M(\mathbf{x}_i)$ between the female and the re-weighted male sample. If the propensity score model (3.9) is correctly specified the expected value of the augmentation term is zero. Similarly, the AIPW estimator can be expressed as

$$\hat{\delta}_{AIPW} = \underbrace{\frac{1}{N_F} \sum_{i=1}^N d_{Fi}(y_i - \hat{\mu}_M(\mathbf{x}_i))}_{\hat{\delta}_{BO}} - \underbrace{\sum_{i=1}^N \hat{w}_i^M(y_i - \hat{\mu}_M(\mathbf{x}_i))}_{\text{Augmentation to BO}}, \quad (3.12)$$

which shows that it is also an augmented version of the BO estimator. The augmentation

⁷We compute the normalized IPW weights (3.11) after trimming.

term in (3.12) re-weights the observable bias in $\hat{\mu}_M(\mathbf{x}_i)$ and has an expected value of zero if the reservation wage model (3.6) is correctly specified. Furthermore, AIPW has favorable theoretical properties for estimating the so-called nuisance parameters $p(\mathbf{x})$ and $\mu_M(\mathbf{x})$. The latter make the AIPW estimator an attractive tool for applications with flexible machine learning methods (Knaus, 2022).

3.2.5 Data-Driven Model Specifications

As mentioned above, our baseline model specifications for $p(\mathbf{x})$ and $\mu_M(\mathbf{x})$ include a large set of explanatory variables in $\mathbf{x}_i$ that we have selected based on economic reasoning and prior evidence from the related literature on the determinants of reservation wages. Nevertheless, our baseline model specification may still be too parsimonious due to potential nonlinear effects of the confounders $\mathbf{x}_i$ on y_i and d_{Fi} . This may lead to a bias in our estimates of δ . Therefore, we also use data-driven model specifications based on Machine Learning (ML) methods that allow for more model flexibility.

The choice of the ML methods we apply depends on the respective estimator of δ . In case of LRM and BO, we apply the Post-Double-Selection (PDS) procedure proposed by Belloni et al. (2014a,b) that is based on the Least Absolute Shrinkage and Selection Operator (LASSO) introduced by Frank and Friedman (1993) and Tibshirani (1996). The LASSO limits the sum of absolute values of the coefficients by adding a penalty term (L1 regularization) to the OLS minimand. Thereby, it shrinks the OLS coefficients towards zero and sets some of them exactly to zero (Hastie et al., 2009). Thus, it is also a model selection operator. Furthermore, in contrast to OLS, it permits the set of potential explanatory variables to exceed the sample size.

The PDS procedure proceeds in two steps. In a first step, we use the LASSO to select the most important predictors of y_i and d_{Fi} . We estimate two linear regression models with y_i and d_{Fi} as dependent variable, respectively, and a set of potential control variables as explanatory variables using the LASSO.⁸ In comparison to our baseline model

⁸For the post-double-selection procedure we use a slightly modified version of the LASSO first intro-

specification, we expand the set of potential controls by constructing various higher-order polynomials and interaction terms. Whereas our baseline model specification includes 93 control variables, the set of potential control variables for our LASSO model specifications includes more than 2,000 variables. In a second step, we estimate the linear regression model (3.5) by OLS using all control variables that have been selected by LASSO, i.e. have a non-zero coefficient, either as a predictor of y_i or as a predictor of d_{Fi} in the first step. The PDS procedure ensures that we include all variables as controls that are either an important predictor of y_i or of d_{Fi} (Belloni et al., 2014b). For the PDS procedure we choose the penalty parameter in both equations as a function of the sample size and the number of potential control variables, as described in Belloni et al. (2014a,b).

In case of AIPW, we apply the double/debiased machine learning (DML) approach proposed by Chernozhukov et al. (2017, 2018).⁹ The DML approach exploits the so-called Neyman orthogonality of the double robust score of the AIPW estimator, which allows the use of flexible ML methods to estimate the nuisance parameters ($\hat{p}(\mathbf{x}_i)$ and $\hat{\mu}_M(\mathbf{x}_i)$ in Equations (3.10) and (3.11)). Moreover, the DML approach combines the use of ML methods with cross-fitting. Cross-fitting is a form of sample splitting that helps to avoid overfitting and ensures $\sqrt{N}$ convergence of the AIPW estimator (Chernozhukov et al., 2017). We apply 5-fold cross-fitting to estimate the nuisance parameters, i.e. we randomly split our sample into five folds with equal size. We use four of the folds to estimate models for $\hat{p}(\mathbf{x}_i)$ and $\hat{\mu}_M(\mathbf{x}_i)$, while leaving out one fold. We then use the estimated models to predict $\hat{p}(\mathbf{x}_i)$ and $\hat{\mu}_M(\mathbf{x}_i)$ in the left-out fold. This procedure is repeated five times such that each fold is left out once for the estimation of the nuisance parameters.¹⁰ Hence, we use no observations to predict its own nuisance parameters. In the final step, we plug the

duced by Belloni et al. (2012). The modified version includes so called ‘penalty-loadings’ that ensure equivariance of the estimated coefficients to rescaling and address non-normality, heteroskedasticity, as well as clustering of the model errors.

⁹We implement the DML procedure in Stata with the *ddml* package of Ahrens et al. (2024) in combination with the *pystacked* package of Ahrens et al. (2022) that supports a variety of machine learning methods. The *pystacked* package permits to implement multiple machine learning algorithms via Python’s *scikit-learn* of Pedregosa et al. (2011).

¹⁰The cross-fitting algorithm we use is equivalent to the DML2 algorithm described in Chernozhukov et al. (2018). This algorithm is recommended by Chernozhukov et al. (2018) as it performs better in small samples than their alternative algorithm.

cross-fitted nuisance parameters $\hat{p}(\mathbf{x}_i)$ and $\hat{\mu}_M(\mathbf{x}_i)$ into Equation (3.10).

To estimate the nuisance parameters $\hat{p}(\mathbf{x}_i)$ and $\hat{\mu}_M(\mathbf{x}_i)$ we use various ML methods, as it is difficult to choose the best-performing one for a specific application. The reason is that the underlying data generating process is unknown (Ahrens et al., 2024). We apply the two tree-based methods Random Forest (RF) and Gradient Boosting (GB), a neural network method we label as Neural Net (NN) as well as the LASSO.¹¹ The implementation of these methods differs between regression and classification problems (see Pedregosa et al. (2011) for details). We consider the estimation of $\hat{\mu}_M(\mathbf{x}_i)$ as a regression task and the estimation of $\hat{p}(\mathbf{x}_i)$ as a classification task. In case of the RF, we estimate the nuisance parameters with a random forest that averages over 100 trees. For estimation of the nuisance parameters with GB, we perform 100 boosting stages and use a learning rate of 0.1. Our NN estimates are based on a neural network with one hidden layer with 100 neurons, where we choose the rectified linear unit function and the logistic function as activation functions in case of the estimation of $\hat{\mu}_M(\mathbf{x}_i)$ and $\hat{p}(\mathbf{x}_i)$, respectively. For the LASSO we choose the penalty parameter by 5-fold cross validation. In case of the LASSO estimation of $\hat{p}(\mathbf{x}_i)$, we implement a logistic regression with a L1-penalty term. As above, we use the expanded set of potential explanatory variables in case of the LASSO. For all other methods we use our baseline set of explanatory variables.

Finally, we apply an ensemble method called stacking proposed by Wolpert (1992) and Breiman (1996). Stacking optimally combines the four other ML methods we use (base learners) into one meta learner. The weights assigned to the different base learners are chosen such that they maximize the out-of-sample predictive accuracy. By using an ensemble method we avoid the risk of relying on one particularly ill-suited ML method (Ahrens et al., 2024). In fact, the decision which learner is most suitable for a particular application is not trivial as the structure of the underlying data generating process is typically not known by researchers (Ahrens et al., 2024). Hence, by using ensemble learning that trains multiple models to solve the same problem and combines them to get

¹¹This set of ML methods is similar to set used by Chernozhukov et al. (2017) in their empirical applications.

better results, we can obtain more accurate and/or more robust models.

3.3 Data, Sample and Descriptive Statistics

For our empirical analysis we use individual level data from the German Socio-Economic Panel Study (SOEP), which surveys around 15,000 households in Germany on a yearly basis and includes detailed information about socio-economic variables. In particular, the SOEP includes information about reservation wages of non-employed individuals. The SOEP questionnaire first included a question about monthly reservation wages in 1987. Yet, only since 2007, non-employed survey participants are asked on a yearly basis about both their monthly reservation wage and their desired weekly working hours, which allows us to compute hourly reservation wages. Comparing hourly reservation wages instead of monthly reservation wages is important since men and women tend to differ in their intended working time. Therefore, we base our analysis on an unbalanced panel of data from 2007 to 2020. Humpert and Pfeifer (2013) also compute hourly reservation wages using the monthly reservation wage and the desired working hours provided by the SOEP, while the study of Christensen (2001) relies only on the monthly reservation wage information.

3.3.1 Reservation Wage Information and Sample Composition

We use the logarithm of net hourly reservation wages as dependent variable in our analysis. Non-employed individuals in the SOEP are asked about their reservation wage and desired working time in the following way. First, they are asked: “What would your net income have to be for you to accept a position?” Second, they are asked: “How many hours per week would you have to work to earn this net income?” Based on the answers to these two questions we construct the hourly reservation wage of individuals as

$$\text{Hourly Reservation Wage} = \left(\frac{\text{Monthly Reservation Wage}}{4.25 \cdot \text{Desired Weekly Working Hours}} \right) \cdot \frac{1}{\text{CPI}_{2015}}.$$

Table 3.1: Variation in Sample Size after Imposition of Sample Restrictions

Restriction	Sample Size
Asked about reservation wage (non-employed, future employment is not ruled out), age 18 - 65, 2007 - 2020	64,775
+ No missing control variables	13,460
+ Looking for full-time job	7,274
+ Wants to take up work soon	5,914
+ Reservation wage is observed	4,201
+ No extreme reservation wage	4,114 (2,653 unique individuals) = Final sample
	<i>Thereof:</i>
	Female observations: 1,716 (1,174 unique female individuals)

Notes: The table shows how the sample decreased after each sample restriction. A few observations are also omitted due to having a sampling weight of zero. *Source:* German Socio-Economic Panel v37, own calculations.

Since we pool observations across survey years, we deflate the hourly reservation wage using the German Consumer Price Index with base year 2015 (CPI_{2015}).¹²

Table 3.1 shows how our final sample is constructed. Our starting point are all observations of individuals who have been asked about their reservation wage (i.e. non-employed individuals that do not rule out a future employment) and are between 18 and 65 years old over the period 2007 and 2020. Then, we discard all observations with missing control variables. Next, we restrict the sample to individuals searching a full-time job. Restricting the sample to individuals searching for a full-time job results in men and women who are more comparable across observed and unobserved characteristics (Caliendo et al., 2017). This restriction reduces potential selection issues. Following Fedorets and Shupe (2021), we keep only respondents who are available to take up work within the next year. The reason for this choice is that we can hardly consider individuals available to start a new job later than one year ‘active’ job seekers.¹³ Out of the remaining sample, we drop 29%

¹²We obtain the monthly Consumer Price Index from the German Federal Statistic Office (<https://www-genesis.destatis.de/genesis/online>, accessed December, 2022).

¹³As a robustness check, we use a sample that considers individuals who are registered as unemployed

of observations due to missing information on hourly reservation wages. Finally, we omit observations with the one percent smallest and one percent largest values in hourly reservation wages. The final sample consists of 4,114 observations (1,716 female and 2,398 male) stemming from 2,653 individuals (1,174 female and 1,479 male).

Although the imposed sample restrictions aim to provide a comparable sample and, hence, to reduce potential selection bias, we cannot rule it out. In particular, our final sample faces potential sample selection bias due to non-random missing reservation wage information (Fedorets and Shupe, 2021). Table 3.2 provides mean characteristics of reporters and non-reporters of the reservation wage. The share of women and married individuals is statistically significantly higher among the group not reporting their reservation wage. Similarly, these individuals live in households with higher net income and have more often a migration background. Instead, the two groups do not differ statistically significantly with respect to basic or higher education as well as in full- and part-time employment or in past unemployment periods. However, respondents tend to be more likely to be registered as unemployed and to be more actively engaged in job search.

The above cited characteristics may be associated to a stronger labor market attachment. Indeed, by reporting reservation wages individuals may signal their labor market attachment (Brown et al., 2011). As we are interested in individuals who are actually looking for work and thus in individuals that have a relatively high labor market attachment, the group of respondents may be appropriate for our study. Nonetheless, it is important to highlight that the main sample in this paper may not be representative of all non-working job seekers. As a consequence, our estimation results are valid (only) for job seekers who are aged between 18 and 65, look for a full-time job, intend to be employed within one year and report their reservation wage.

at the employment agency but that is not restricted with respect to intended starting time of employment. The estimation results do not change markedly with this alternative sample as shown in Table 3.A.2 and Figure 3.A.1 and 3.A.2 in the Appendix.

Table 3.2: Summary Statistics by (Non-)Missing Reservation Wage

	(1)	(2)	(3)
	Reservation Wage		
	Reported	Not Reported	Difference
Female	0.35 (0.48)	0.42 (0.49)	-0.07***
Age	39.87 (12.22)	40.67 (12.99)	-0.79**
Married	0.31 (0.46)	0.39 (0.49)	-0.09**
Household Net Income (in 1,000 €)	1.73 (1.28)	1.86 (1.33)	-0.13***
Child (under 16) in Household	0.23 (0.42)	0.21 (0.40)	0.02*
Migration Background	0.23 (0.42)	0.32 (0.47)	-0.09***
No School Degree	0.03 (0.17)	0.02 (0.15)	0.00
Other School Degree	0.09 (0.28)	0.16 (0.37)	-0.07***
Basic Secondary Education (<i>Hauptschule</i>)	0.33 (0.47)	0.32 (0.47)	0.01
Secondary Education (<i>Realschule</i>)	0.31 (0.46)	0.23 (0.42)	0.08***
Upper Secondary Education (<i>Abitur</i>)	0.20 (0.40)	0.20 (0.40)	0.01
Vocational Degree	0.70 (0.46)	0.66 (0.47)	0.04***
College Degree	0.14 (0.34)	0.12 (0.33)	0.02
Full-Time Experience	11.32 (11.29)	11.52 (11.32)	-0.20
Part-Time Experience	1.82 (3.08)	1.96 (3.62)	-0.13
Unemployment Experience	4.35 (5.02)	4.27 (4.95)	0.07
Last Net Hourly Wage (in) €	7.01 (4.33)	6.83 (4.16)	0.18
Registered as Unemployed	0.78 (0.42)	0.68 (0.47)	0.09***
Active Job Search in Last Four Weeks	0.66 (0.48)	0.57 (0.50)	0.08***
Observations	4,160	1,690	

Notes: Columns (1) and (2) display mean values of variables for observations without and with missing reservation wage information. Standard deviations are shown in parenthesis. Column (3) shows the mean difference. *, **, and *** indicate statistical significant differences at the 10%-, 5%-, and 1%-level based on a t-test for equal means, respectively. Last wage is CPI adjusted (base year 2015). SOEP sample weights used. *Source:* SOEP v37, own calculations.

3.3.2 Descriptive Statistics

One advantage of using data from the SOEP is that it provides extensive information on a broad range of potential reservation wage determinants that we can use as control variables in our analysis. In our baseline model specification, we include various demographic variables like age, marital status, migration background, household income, and household size, that are all found to be correlated with reservation wages (Brown et al., 2011; Christensen, 2001; Humpert and Pfeifer, 2013). Moreover, we extract information on the employment status of the partner, the age of the youngest child, and whether the individual is engaged with childcare. This choice is inspired by findings from Brown et al. (2011) suggesting that gender differences in variables about household and partner characteristics as well as child-related factors affect the gender gap in reservation wages. Next, we include variables about educational attainment and labor market history, that are also important determinants of reservation wages (Christensen, 2001; Humpert and Pfeifer, 2013) and are related to the corresponding gender gap (Brown et al., 2011). In particular, we control for having a vocational degree, a college degree, labor market experience and unemployment experience.

Additionally, we exploit the panel structure of the SOEP and include information about the last employment. Alongside the last net hourly wage, which we construct analogously to the hourly reservation wage, we control for the occupation and the industry of the last job using the ISCO88 (1-digit) and NACE (Level 1) classifications, respectively. Information on the last job have also been used as explanatory variables for reservation wages in related studies (Brown et al., 2011; Caliendo et al., 2017). We also have information on whether the individual is registered as unemployed as well as on whether the individual receives unemployment benefits (*Arbeitslosengeld* (ALG) and *Arbeitslosengeld II* (ALG II)) and if so, how much.¹⁴ The previous literature has shown that the amount

¹⁴*Arbeitslosengeld* is an unemployment insurance benefit in Germany that is paid upon the onset of unemployment and depending on other conditions. It is limited in time and is normally paid for up to one year, and for up to two years for older unemployed individuals. *Arbeitslosengeld II* is a basic security benefit paid to employable job seekers in Germany that is unlimited in time.

of unemployment benefits is positively associated with reservation wages (Addison et al., 2009).

Further, the SOEP questionnaire asks non-employed individuals about their job search. Thus, we are able to include information on whether an individual has actively searched for a job in the last four weeks, is looking for a full-time job, has difficulty in finding a job, could accept a job immediately, wants to be employed immediately, and thinks a future employment is rather unlikely. Such job search characteristics may affect reservation wages of individuals and may differ between women and men. In fact, for instance Cortés et al. (2023) underline the role of gender differences in job search behavior. Likewise, the SOEP questionnaire contains questions referring to personality traits and risk preferences. It has been shown that risk attitudes are correlated with reservation wages (Pannenberg, 2010; McGee and McGee, 2023; Cortés et al., 2023). The literature on the gender pay gap considers personality traits as potential explanatory factors (e.g., Bonaccolto-Töpfer and Briel, 2022). Therefore, we include the Big Five Personality Traits (openness, conscientiousness, extraversion, agreeableness, and neuroticism) along with the willingness to take risks in our analysis. We consider the latter as control variable given that past research (e.g., Dohmen et al., 2011) has shown that willingness to take risks is a good predictor of actual risk-taking. Lastly, we control for the survey year, the federal state in which an individual lives and the unemployment rate at the federal state level¹⁵ as previous studies highlight that reservation wages are related to local unemployment levels (Brown and Taylor, 2015).

Table 3.3 shows of reservation wages and selected control variables by gender. Columns (1) and (2) display mean values and standard deviations for females and males, respectively, while Column (3) shows the gender-specific mean difference and whether it is statistically significant. The monthly reservation wage of men in our sample is around 138 € higher than that of women on average. However, we also see, that men are willing to work around two hours more per week than women on average, which highlights the impor-

¹⁵We obtain data on the monthly unemployment rate at federal state level from the German Federal Statistic Office (<https://www-genesis.destatis.de/genesis/online>, accessed December, 2022).

tance of comparing hourly reservation wages. When looking at the hourly reservation wage, we find a small but statistically significant difference between females and males in our sample. Women in our sample set on average 0.30 € lower hourly reservation wages compared to men, what corresponds to a gender gap of around 3% in relative terms.

This finding is in line with results from the related literature that typically also finds higher reservation wages for men compared to women (e.g. Brown et al., 2011; Caliendo et al., 2017; Le Barbanchon et al., 2021; McGee and McGee, 2023). However, the raw gender gap we find in our sample is smaller compared to other studies. Brown et al. (2011) and Le Barbanchon et al. (2021) find a raw gender gap in reservation wages of 9% in the UK and France, respectively. Cortés et al. (2023) find an initial gender gap of 15% in a laboratory experiment among US graduates, while Caliendo et al. (2017) document a raw gender gap of 13% among German job seekers.

Panels (b) - (g) of Table 3.3 display summary statistics for selected control variables. Unsurprisingly, we see substantial gender-specific differences in many characteristics. Men in our sample are slightly older on average, while women are more likely to have a partner that is full-time employed, more likely to be engaged in childcare, and to live in a household with a higher monthly net income. There also exist significant gender differences in terms of educational attainment and labor market experience. Females are substantially more likely to have a high-school or college degree than males. However, the average full-time experience of men is more than four years higher than that of women. In addition, men also have higher accumulated unemployment experience. Though, we do not see a significant gender difference in the last observed hourly wage. Interestingly, the last wage is lower than the current reservation wage on average for both men and women. With regard to the unemployment status, we observe that males in our sample are considerably more likely to be registered as unemployed and to receive unemployment benefits.

Moreover, men and women differ in terms of their job search. A higher fraction of men than of women intends to be employed immediately, could accept a job immediately, and has actively searched for a job in the last four weeks. Finally, we also find signifi-

cant gender differences in risk attitudes and personality traits that are in line with the previous literature (e.g. Cortés et al., 2023; Weisberg et al., 2011, respectively). Whereas females score higher in all the Big Five Personality Traits, males have on average a higher willingness to take risks.

Table 3.3: Summary Statistics

	(1) Females	(2) Males	(3) Difference
(a) Reservation Wages			
Monthly Reservation Wage	1389.13 (557.47)	1527.25 (605.86)	-138.12***
Working Hours	37.60 (6.40)	39.76 (4.91)	-2.17***
Hourly Reservation Wage	8.82 (3.53)	9.12 (3.65)	-0.30**
(b) Background Variables			
Age	39.36 (12.06)	40.20 (12.24)	-0.84**
Married	0.31 (0.46)	0.31 (0.46)	0.00
Household Net Income (in 1,000 €)	1.84 (1.32)	1.67 (1.25)	0.17***
Household Size	2.31 (1.28)	2.27 (1.37)	0.03
Partner Works Full-Time	0.25 (0.43)	0.10 (0.31)	0.15***
Partner Works Part-Time	0.01 (0.11)	0.09 (0.29)	-0.08***
Partner Unemployed	0.05 (0.22)	0.09 (0.28)	-0.03***
Youngest Child ≤ 3	0.00 (0.06)	0.01 (0.09)	0.00*
Youngest Child > 3 & ≤ 6	0.03 (0.16)	0.03 (0.17)	0.00
Youngest Child > 6 & ≤ 18	0.15 (0.36)	0.13 (0.34)	0.01
Childcare	0.28 (0.45)	0.19 (0.39)	0.10***
Migration Background	0.23 (0.42)	0.23 (0.42)	-0.01
(c) Education			
No School Degree	0.01 (0.11)	0.04 (0.19)	-0.02***
Other School Degree	0.08 (0.27)	0.09 (0.28)	-0.01
Basic Secondary Education (Hauptschule)	0.26 (0.44)	0.38 (0.49)	-0.12***
Secondary Education (Realschule)	0.34 (0.47)	0.29 (0.45)	0.06***
Upper Secondary Education (Abitur)	0.28 (0.45)	0.16 (0.37)	0.11***
Vocational Degree	0.68 (0.47)	0.72 (0.45)	-0.04**
College Degree	0.20 (0.40)	0.11 (0.31)	0.09***
(d) Labor Market History			
Full-Time Experience	8.72 (9.28)	12.75 (11.98)	-4.03***
Part-Time Experience	3.10 (3.99)	1.13 (2.13)	1.98***
Unemployment Experience	4.04 (5.22)	4.51 (4.90)	-0.46***
Last Net Hourly Wage	6.90 (4.07)	7.06 (4.40)	-0.16
(e) Unemployment			
Registered as Unemployed	0.70 (0.46)	0.82 (0.39)	-0.11***
Receiving Unemployment Benefits I	0.27 (0.44)	0.30 (0.46)	-0.04**
Receiving Unemployment Benefits II	0.37 (0.48)	0.43 (0.49)	-0.05***

Continued on next page

(f) Job Search			
Finding Vacancy: Difficult	0.64 (0.48)	0.62 (0.48)	0.02
Finding Vacancy: Impossible	0.18 (0.38)	0.19 (0.40)	-0.02
Could Accept Job Immediately	0.77 (0.42)	0.84 (0.37)	-0.07***
Active Job Search in Last Four Weeks	0.63 (0.48)	0.67 (0.47)	-0.04**
Employment Intended Immediately	0.70 (0.46)	0.79 (0.41)	-0.09**
Future Employment Rather Unlikely	0.05 (0.22)	0.05 (0.23)	0.00
(g) Personality			
Openness	4.69 (1.27)	4.44 (1.20)	0.25***
Conscientiousness	3.24 (0.88)	2.86 (1.02)	0.37***
Extraversion	2.44 (1.15)	2.12 (1.14)	0.32***
Neuroticism	1.50 (1.21)	1.13 (1.23)	0.37***
Agreeableness	2.88 (0.94)	2.57 (0.98)	0.31***
Risk Tolerance	4.75 (2.47)	5.34 (2.33)	-0.59***
Observations	1,695	2,419	4,114

Notes: Columns (1) and (2) display mean values of variables in female and male subsample, respectively. Standard deviations are shown in parenthesis. Column (3) shows the mean difference between females and males. *, **, and *** indicate statistical significant differences at the 10%-, 5%-, and 1%-level based on a t-test for equal means, respectively. Reservation wages and last wage are CPI adjusted (base year 2015). SOEP sample weights used. *Source:* SOEP v37, own calculations.

3.4 Unexplained Gender Gap in Reservation Wages

In this section, we present the estimation outcome. We start by providing the results for the most widely used methods in the gender gaps literature: standard linear regressions (LRM) and the Blinder-Oaxaca (BO) decomposition based on a baseline model specification. Additionally, we show the Gelbach decomposition of the LRM estimate of the unexplained gap. We conduct decomposition analyses in order to learn more about the role of specific sets of controls and estimators (LRM and BO). Recall that the baseline model specification is inspired by the literature and thus, we expect to find results in line with the related literature. Next, we confront estimates of the unexplained gap based on the various estimators (LRM, BO and AIPW) and model specifications (baseline and data-driven).

3.4.1 Reservation Wage Determinants

Table 3.4 shows the estimation results from conventional OLS regressions.¹⁶ In Column (1) we regress the log hourly reservation wage only on an intercept and the female dummy, while in Column (2) we additionally include our baseline covariates. The raw gender gap in Column (1) amounts to 3% (log approximation) but is not statistically significant. In line with the descriptive statistics from above, the raw reservation wage gap between women and men is negative and thus suggest that females set lower reservation wages on average in our sample. Yet, the gap in reservation wages we find is substantially smaller compared to the gender gap in realized wages in the German working population. According to the German Federal Statistic Office,¹⁷ the raw gender pay gap varied between 18% and 23% in Germany over the time period 2007-2020. As stated, related studies found larger raw gender gaps in reservation wages (e.g. Brown et al., 2011; Caliendo et al., 2017).

The *ceteris paribus* gap in Column (2), however, is statistically significant at a 5% level and amounts to 5% (log approximation) in favor of men. Thus, in a like-with-like comparison, we find a larger gender reservation wage gap. In particular, this *ceteris paribus* gap is comparable to the above cited studies (i.e., Caliendo et al., 2017; Brown et al., 2011). Columns (3) and (4) of Table 3.4 show OLS regressions separately for men and women and reveal substantial differences in determinants of reservation wages for men and women. For instance, having obtained the highest school-leaving degree (*Abitur*) has a significant positive effect on reservation wages of men only. A college degree does positively affect reservation wages of both men and women. Yet, the effect is substantially more pronounced for women. Having young children aged between zero and three or between three and six years affects female reservation wages more strongly. The effect is insignificant in both cases, though. Moreover, women who are registered as being unemployed have on average 9% lower reservation wages than women who are not registered, while

¹⁶Table 3.4 displays only the estimated coefficients of a selection of the included control variables. Full results are shown in Table 3.A.3 (except fixed effects).

¹⁷German Federal Statistic Office (<https://www.destatis.de/DE/Themen/Arbeit/Verdienste/Verdienste-GenderPayGap/Tabellen/ugpg-01-gebietsstand.html>, accessed December, 2022).

the corresponding effect for men is only 6%. In contrast, receiving unemployment benefits (*Arbeitslosengeld*) has a more pronounced negative effect on reservation wages of men. Individuals stating that finding a fitting vacancy is difficult or impossible for them have on average lower reservation wages. The effect is more pronounced for men. Intending to be employed immediately only lowers reservation wages significantly for women. This finding may explain the larger raw gender gap in the study of Caliendo et al. (2017) using a sample of individuals even closer attached to the labor market (only newly unemployed job seekers) than ours.

There are also a number of characteristics that affect reservation wages of men and women similarly. For instance, the effect of being married on reservation wages is comparable between men and women. When the partner works full time, reservation wages of both genders are roughly 5% lower. The higher the amount of unemployment benefits (*Arbeitslosengeld*), the higher reservation wages. Further, higher wages in the last job as well as a higher household income increase reservation wages substantially for both men and women. The effect is slightly stronger for men in case of the last wage and slightly more pronounced for women in case of household income.

The results suggest gender differences with regard to the impact of personality traits on reservation wages. A higher level of extraversion affects only male reservation wages significantly, while higher conscientiousness positively affects only female reservation wages. However, for this set of control variables, all corresponding confidence intervals of the male and female coefficient estimates do overlap. In line with results from laboratory search experiments, risk tolerance is not a major determinant of gender differences in reservation wages (McGee and McGee, 2023).

Table 3.4: Reservation Wage Regressions, Selected Controls

Dependent Variable: Log Hourly Reservation Wage				
	(1)	(2)	(3)	(4)
	Raw	Baseline	Males	Females
Female (Dummy)	-0.0316 (0.0218)	-0.0496** (0.0201)		

Continued on next page

3.4. Unexplained Gender Gap in Reservation Wages

Married (Dummy)	0.0470** (0.0185)	0.0360 (0.0229)	0.0390 (0.0274)
Log Household Monthly Income	0.0531*** (0.0164)	0.0432** (0.0194)	0.0604** (0.0242)
Partner Works Full-Time (Dummy)	-0.0574*** (0.0196)	-0.0527* (0.0285)	-0.0556** (0.0267)
Youngest Child ≤ 3 (Dummy)	0.0095 (0.0441)	0.0074 (0.0512)	0.0379 (0.0895)
Youngest Child > 3 and ≤ 6 (Dummy)	0.0594 (0.0363)	0.0232 (0.0463)	0.0571 (0.0527)
Youngest Child > 6 and ≤ 18 (Dummy)	-0.0250 (0.0211)	-0.0043 (0.0274)	-0.0377 (0.0266)
Basic Secondary Education (Hauptschule) (Dummy)	0.0005 (0.0395)	0.0138 (0.0460)	-0.0287 (0.0558)
Secondary Education (Realschule) (Dummy)	0.0150 (0.0395)	0.0525 (0.0482)	-0.0545 (0.0519)
Upper Secondary Education (Abitur) (Dummy)	0.0721* (0.0429)	0.1168** (0.0538)	-0.0172 (0.0539)
Vocational Degree (Dummy)	0.0190 (0.0174)	0.0192 (0.0222)	-0.0117 (0.0237)
College Degree (Dummy)	0.1198*** (0.0293)	0.1117*** (0.0404)	0.1563*** (0.0333)
Log Hourly Wage Last Job	0.0396*** (0.0136)	0.0413** (0.0163)	0.0333* (0.0187)
Registered as Unemployed (Dummy)	-0.0682*** (0.0224)	-0.0599** (0.0296)	-0.0943*** (0.0298)
Receiving Unemployment Benefits	-0.1162*** (0.0267)	-0.1431*** (0.0343)	-0.0891** (0.0396)
Amount of Unemployment Benefits	0.0001*** (0.0000)	0.0001*** (0.0000)	0.0002*** (0.0000)
Finding Vacancy Difficult (Dummy)	-0.0918*** (0.0205)	-0.1087*** (0.0263)	-0.0476* (0.0268)
Finding Vacancy Impossible (Dummy)	-0.1192*** (0.0246)	-0.1413*** (0.0321)	-0.0507 (0.0347)
Employment Intended Immediately (Dummy)	-0.0490*** (0.0188)	-0.0362 (0.0242)	-0.0546** (0.0263)
Conscientiousness	0.0121 (0.0077)	0.0095 (0.0089)	0.0208* (0.0117)
Extraversion	0.0124* (0.0064)	0.0206*** (0.0077)	0.0053 (0.0096)
Observations	4,114	4,114	2,419
			1,695

Notes: The dependent variable is the log of the hourly reservation wage. The tables only displays the estimated coefficients of a selection of the included control variables. See Table 3.A.3 for the full regression

outcome (except fixed effects). Standard errors are clustered at the individual level and are shown in parenthesis. *, ** and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

3.4.2 Decomposing the Gender Gap

In order to better understand the role of the various sets of controls in explaining the gender gap as well as the role of the two most commonly used estimators in this strand of the literature, i.e. LRM and BO, we decompose the gap.¹⁸ The decomposition analysis is based on Gelbach (2016)'s approach for the LRM estimator and on the standard Blinder (1973) and Oaxaca (1973) decomposition outlined in Section 3.2.

Table 3.5: Reservation Wage Decomposition, Male Reference Category

Estimator	(1)	(2)	(3)	(4)	(5)	(6)
	LRM		BO			
Raw Gender Gap	-0.032 (0.0218)		-0.032 (0.0214)			
Explained Part			0.035 (0.024)			
Unexplained Gap/ Unexplained Part	-0.050** (0.0201)		-0.067*** (0.0213)			
	Difference attributed to in %		Detailed Decomposition Explained Part Unexplained Part in % in %			
Household	0.006** (0.0031)	34	0.005*** (0.0029)	14	0.002 (0.0305)	0.5
Partner	-0.011*** (0.0033)	-61	-0.010** (0.0041)	-28	-0.003 (0.0127)	-1
Children	0.004* (0.0020)	21	0.007** (0.0028)	19	-0.017* (0.0089)	-3
Education	0.019*** (0.0056)	104	0.023*** (0.0072)	65	-0.101 (0.0678)	-20
Labor Market History	-0.014 (0.0108)	-80	-0.004 (0.0153)	-12	0.005 (0.0785)	1
Unemployment	0.005	26	0.006	17	0.020	4

Continued on next page

¹⁸Note that we cannot do similar considerations for AIPW as the contributions of different variables to the gender gap cannot be quantified with the AIPW estimator in a way that does not depend on the ordering of the addition of explanatory variables.

3.4. Unexplained Gender Gap in Reservation Wages

	(0.0035)		(0.0042)		(0.0265)	
Job Search	0.003	16	0.004	11	0.072*	14
	(0.0031)		(0.0039)		(0.0388)	
Personality	0.010**	57	0.011*	30	0.133*	26
	(0.0050)		(0.0057)		(0.0741)	
Remainder	-0.003	-18	-0.006	-17	0.408	79
	(0.0062)		(0.0074)		(0.3232)	
Constant					-0.586*	
					(0.3459)	
	Total difference between raw and unexplained gap:		Sum Explained:		Sum Unexplained except part of constant:	
	0.018	100	0.035	100	0.554**	100
	(0.0193)		(0.0239)		(0.010)	
Observations	4,114		4,114			

Notes: The dependent variable is the log of the hourly reservation wage. The male coefficient vector is used as the reference vector. Explanatory variables are grouped as follows. **Household:** log monthly household income, household size. **Partner:** marital status, employment status of partner. **Children:** age of youngest child, childcare. **Education:** basic secondary education, secondary education, upper secondary education, vocational degree, college degree. **Labor Market History:** full-time experience, part-time experience, unemployment experience, wage last job, industry last job, occupation last job. **Unemployment:** registered as unemployed, receiving unemployment benefits (ALG and ALG II), amount of ALG. **Job Search:** difficulty of finding fitting vacancy, could accept job immediately, active job search in last four weeks, employment intended immediately, future employment rather unlikely. **Personality:** Big Five Personality Traits (openness, conscientiousness, extraversion, neuroticism, agreeableness), risk tolerance. **Remainder:** age, migration background, health, log unemployment rate, federal state, survey year. The percentage contribution for the BO decomposition are calculated relative to the explained or unexplained part excluding the constant (i.e. $(-0.067)-(-0.5859)=0.519$). The standard errors are shown in parenthesis and for the LRM-Gelbach decomposition are clustered at the individual level, while they are calculated as described by Jann (2008) and are clustered at the individual level for the BO decomposition. *, ** and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 3.5 shows the detailed decompositions for the LRM (Columns (1) and (2)) and BO estimator (Column (3)-(6)).¹⁹ Recall that the decompositions are estimated as $\bar{y}_F - \bar{y}_M$. While the well-known BO decomposition presents the raw gap as well as the explained and unexplained parts, the Gelbach approach decomposes the change in coefficient estimates when going from the raw to the unexplained LRM estimate of the gap.

¹⁹The BO decomposition results do not change substantially when using women as reference category. The corresponding (BO) results are shown in Table 3.A.4 in the Appendix.

By construction, the unexplained gap for LRM is identical to the unexplained gap from the baseline specification in Table 3.4 (amounting to -5%). Similarly, the BO decomposition suggests a negative unexplained gap of 7% (log approximation). The marked unexplained parts imply that the gender reservation wage gap cannot be explained based on standard observable characteristics.

Both estimators confirm that differences in education in favor of women (positive sign) substantially contribute to closing the gap. This finding is in line with results from the literature suggesting that women have overtaken men in educational attainment (e.g. Goldin et al., 2006; Goldin, 2014; Blau and Kahn, 2017). As women are more often concerned with children, child-related controls positively contribute to the explained part. Moreover, gender differences in household characteristics and personality traits significantly and positively drive the gap. In terms of the explained part, this implies that women come from better endowed households and exhibit more beneficial personality traits. Thus, differences in personality traits, unemployment and job search between men and women work towards a closing of the differential. In accordance with our results, McGee and McGee (2023) find as well that, despite controlling for factors like personality traits or risk preferences, the gap cannot be entirely explained. Further, these findings are in line with research on gender differences in realized wages (Schäfer and Schwiebert, 2018; Collischon, 2020).

Partner controls, in contrast, have a highly statistically significant and negative effect on the gender gap. This set of controls explains a relevant, though different, fraction of the gap for the LRM and BO estimator (61% and 28%, respectively). This highlights the fact that marriage (Dougherty, 2006) and partnership in general have a sizeable effect on labor market outcomes for both men (positive) and women (negative). The finding may be related to a persistent specialization of men in paid work and of women in domestic work. The result suggests persistent gender norms about appropriate roles in the family and how they impact labor market and family outcomes (Blau and Winkler, 2022). While in case of the LRM about 80% of the gap can be attributed to gender differences in individual's'

labor market history, in case of the BO estimator the same set of controls explains only 12%. Apart from the different relevance of this set of controls across the two estimators, the finding underpins that experience gaps do persist in the labor market (Blau and Kahn, 2017).

In contrast, gender differences in remuneration (the unexplained part in the BO estimator) to past labor market experience account only for 1% of the unexplained BO gap. Despite an overtaking of women relative to men in educational attainment, differences in returns to this set of controls are responsible for 20% of the unexplained gap. For the same estimator, lower remuneration of women for child- and partnership-related controls is attributable to 3% and 1% of the unexplained reservation wage gap, respectively. For all other sets of controls, gender differences in remuneration lead to lower unexplained reservation wage gaps.

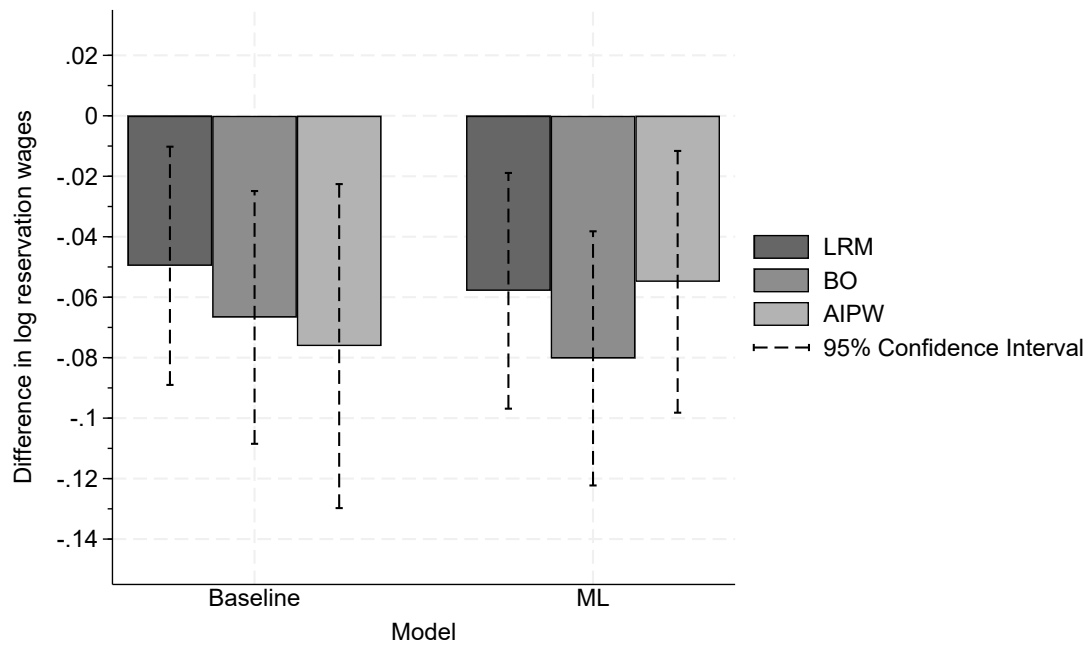
Control variables related to labor market history, unemployment, job search and demographic background characteristics are statistically relatively less significant. Even though the coefficient estimates of most of these sets of controls are small, economically they are mostly relevant. Consider, for instance, the coefficient estimate of the contribution of remaining controls (amounting to 0.003 and 0.006 for the LRM estimator and the explained part of the BO estimator, respectively). Despite the numerically small point estimates, gender differences attributable to this set of controls are responsible for 18% and 17% of the reduction in the gap when going from the raw wage model to the wage model from the baseline specification.

3.4.3 The Role of Various Estimators

Figure 3.1 shows the estimates of the unexplained gender gap in reservation wages using the baseline and the Machine Learning (ML) model specifications based on various estimators.²⁰ Overall, the estimates of the unexplained gap are relatively robust across the different estimation methods and model specifications and range between 5% and

²⁰Table 3.A.5 and 3.A.6 in the Appendix show more detailed results. We show in Figure 3.A.4 and 3.A.5 in the Appendix, the distribution of the propensity score by treatment status (gender).

Figure 3.1: Unexplained Gender Gap in Reservation Wages, Various Estimators



Notes: The figure shows estimates of the unexplained gender gap in reservation wages using linear regressions (LRM), Blinder-Oaxaca decompositions (BO), and augmented inverse probability weighting (AIPW). Baseline model specification and machine learning model specifications (ML) are used. The LRM and BO estimators with the ML specification is based on the post-double-selection procedure by Belloni et al. (2014a,b). The AIPW estimator with the ML specification is based on the double/debiased machine learning approach by Chernozhukov et al. (2017, 2018) and the nuisance parameters (propensity score and predicted male reservation wage) are estimated based on an ensemble method that optimally combines four machine learning methods (Random Forest, Gradient Boosting, Neural Net, and the LASSO). Standard errors are clustered at the individual level. Bootstrapped standard errors (2,000 replications, clustering at the individual level) are used in case of BO with the ML specification. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

8% (log approximations). In particular, the estimated unexplained gaps are all strictly negative and hence are in line with results from the literature of lower *ceteris paribus* or adjusted reservation wages for women compared to men. Comparing our estimates based on the baseline model specification, we find some (small) differences in size of the estimates between the LRM, the BO, and the AIPW estimator. Among the estimates from the baseline model specification, the AIPW estimate is largest and the LRM estimate is smallest. This result is different from findings of Strittmatter and Wunsch (2023) who find the largest gender pay gap estimates with the LRM estimator.

Our parametric LRM estimator is relatively insensitive to how reservation wage determinants are included (ranging from 5.0% to 5.8%). The LRM estimate that is based on the PDS procedure is slightly larger but close to the conventional LRM estimate based on the baseline model specification. This result differs again from findings in the gender-pay-gap literature (Strittmatter and Wunsch, 2023; Bonaccolto-Töpfer and Briel, 2022). The BO estimator, in comparison, does vary slightly more between the two model specifications (ranging from 6.7% to 8.0%), being again larger in case of the ML model specification. The doubly robust AIPW estimator seems to be most sensitive to how reservation wage determinants are included (ranging from 5.5% to 7.7%).²¹ In this case, the estimate based on the ML model specification is considerably smaller.

All in all, given that we control for a wide range of potential reservation wage determinants, our findings provide robust empirical evidence that non-employed women set considerably lower reservation wages than comparable men in Germany. Our results are relatively more sensitive with respect to methodological than to model-specification choices. However, all of our estimates are not statistically significantly different from each other. Using several estimators, thus, adds credibility to the estimation results as well as to the derived policy implications. This is particularly relevant as previous studies on the gender pay gap suggested that gender gap estimates may depend on the estimation

²¹The nuisance parameters for the AIPW estimator with ML model specification in Figure 3.A.1 are estimated with the ensemble method described in Section 3.2.5. Figure 3.A.3 in the Appendix displays the AIPW estimates that only use one of the different base learners (Random Forrest, Gradient Boosting, Neural Net, LASSO), respectively, to estimate the nuisance parameters.

method applied. In terms of magnitude, our estimates of the unexplained gender gap in reservation wages are comparable with the findings of previous studies. Caliendo et al. (2017) find an unexplained gender gap of 5% in the German labor market. For France, Le Barbanchon et al. (2021), controlling for confounding factors, obtained a gender gap of 4% between men and women, while Brown et al. (2011) find an unexplained gender gap of 7% in the UK.

3.5 Heterogeneity in the Gender Gap

In this section we study how the unexplained gender gap differs across different points of the reservation wage distribution and by different subgroups. This analysis may be helpful on two dimensions. First, detecting particularly pronounced unexplained gender gaps at specific points of the distribution or for specific subgroups may help to better understand the drivers of the gender gap. Second, by exactly matching on dimensions like marital status and presence of children we may mitigate concerns that our estimates are biased. For instance, we know from the literature that men are more optimistic about their abilities compared to women (e.g., Niederle, 2017). If this behavior is prevalent only in specific subsamples such as higher-educated individuals or at specific points of the reservation wage distribution, we may find heterogeneity in the unexplained reservation wage gap across different types of individuals.

3.5.1 Gender Reservation Wage Gaps across the Distribution

To evaluate the gender gap across the unconditional reservation wage distribution we use unconditional quantile regressions proposed by Firpo et al. (2009). However, since we pool observations over different survey years, we first have to adjust our outcome variable for potential changes in the reservation wage distribution (DiNardo et al., 1996). To do so, we regress the log hourly reservation wage on the set of survey-year dummies and use the residual of this regression as the outcome variable. Hence, we compute Recentered

Table 3.6: Estimates of Gender Gap in Reservation Wages across the Distribution, Baseline Model

	(1)	(2)	(3)	(4)	(5)
RIF at Percentile	10%	25%	50%	75%	90%
Raw Gender Gap	0.0084 (0.0339)	-0.0158 (0.0195)	-0.0395 (0.0256)	-0.0105 (0.0253)	-0.0813** (0.0413)
Unexplained Gap					
RIF-OLS/LRM	0.0319 (0.0390)	-0.0143 (0.0225)	-0.0488** (0.0246)	-0.0386 (0.0275)	-0.1284** (0.0531)
BO	0.0560 (0.0430)	-0.0120 (0.0275)	-0.0529* (0.0283)	-0.0523 (0.0333)	-0.1862*** (0.0644)
Observations	4,114	4,114	4,114	4,114	4,114

Notes: The dependent variable is the survey year adjusted log hourly reservation wage. The raw gap is calculated as $RIF(q_\tau y_F) - RIF(q_\tau, y_M)$, where $RIF(q_\tau, y_G)$ refers the RIF of G at percentile q_τ with $G = F, M$. Bootstrapped standard errors (2,000 replications) are clustered at the individual level and are shown in parenthesis. *, ** and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Influence Functions (RIF) of various percentiles of the survey-year adjusted reservation wage distribution and regress them on our set of baseline control variables in the male and female subsamples, respectively (Firpo et al., 2009).

Table 3.6 shows the unexplained gender gap at various percentiles based on unconditional quantile regressions including a female dummy. We find a raw gender gap in favor of men between the first and the third quartile of the distribution that is similar in size to the raw gender gap at the mean. Yet, the raw gap is again not statistically significant in this range of the distribution. In contrast, the results suggest a tiny gender gap of 0.8% (log approximation) in favor of women at the first decile of the distribution, that is statistically and economically insignificant though. At the ninth decile, however, the raw gender gap amounts to 8% in favor of men and is statistically significant at the 5%-level. These findings suggest that the raw gender gap in reservation wages changes across the distribution. Also the unexplained gap varies substantially across the distribution. In case of the decomposition (BO), the unexplained part decreases from 6% (in favor of women) at the first decile to -19% (in favor of men) at the ninth decile. Similarly, the

unexplained gender gap estimated with a simple regression (LRM) decreases from 3% at the first decile to -13% at the ninth decile.

Thus, our findings suggest that the unexplained gender gap is larger at the top of the reservation wage distribution compared to the middle and the bottom. Interestingly, our results mimic the glass ceiling effect of the gender gap in realized wages. The phenomenon consists of particularly pronounced gender pay gaps at the upper tail of the wage distribution and has been documented for Germany (e.g. Collischon, 2019). Using German SOEP data but considering the period 2010-2017, Fedorets and Shupe (2021) find the largest *ceteris paribus* gender reservation wage gaps at the median and top (i.e. at the seventh decile in their case). However, in contrast to our results, the analysis of Fedorets and Shupe (2021) suggests relative constant *ceteris paribus* gender differences in reservation wages in favor of men between 5% and 9%.²² Further, our findings differ from the results of Brown et al. (2022) for Italy who found that the gender gap in reservation wages is particularly pronounced at the bottom of the distribution.

We also use the ML model specifications instead of our baseline model specification for the unconditional quantile regressions. Results are shown in Table 3.A.7 in the Appendix, but do not differ substantially from our results in Table 3.6. All estimates suggest a larger unexplained gender gap at the upper tail of the reservation wage distribution. Hence, the pattern of the unexplained reservation wage gap across the distribution seems to be robust to different model specifications. This finding is similar to our result at the mean. Table 3.A.7 also shows the aggregate BO decomposition across the distribution (for more details see Fortin et al., 2011; Firpo et al., 2018). As at the mean, the explained and unexplained part point in opposite directions at most evaluated points of the distribution. The explained part increases from -5% (in favor of men) at the first decile to 11% (in favor of women) at the ninth decile.

²²Figure 3.A.7 shows the pattern of point estimates when using a similar specification (panel (a)) as in Fedorets and Shupe (2021) or our baseline specification (panel (b)) over the period 2010-2017. We consider for estimation the first to seventh decile of the reservation wage distribution as in Fedorets and Shupe (2021). The pattern of point estimates is highly comparable.

3.5.2 Gender Reservation Wage Gaps for Selected Subgroups

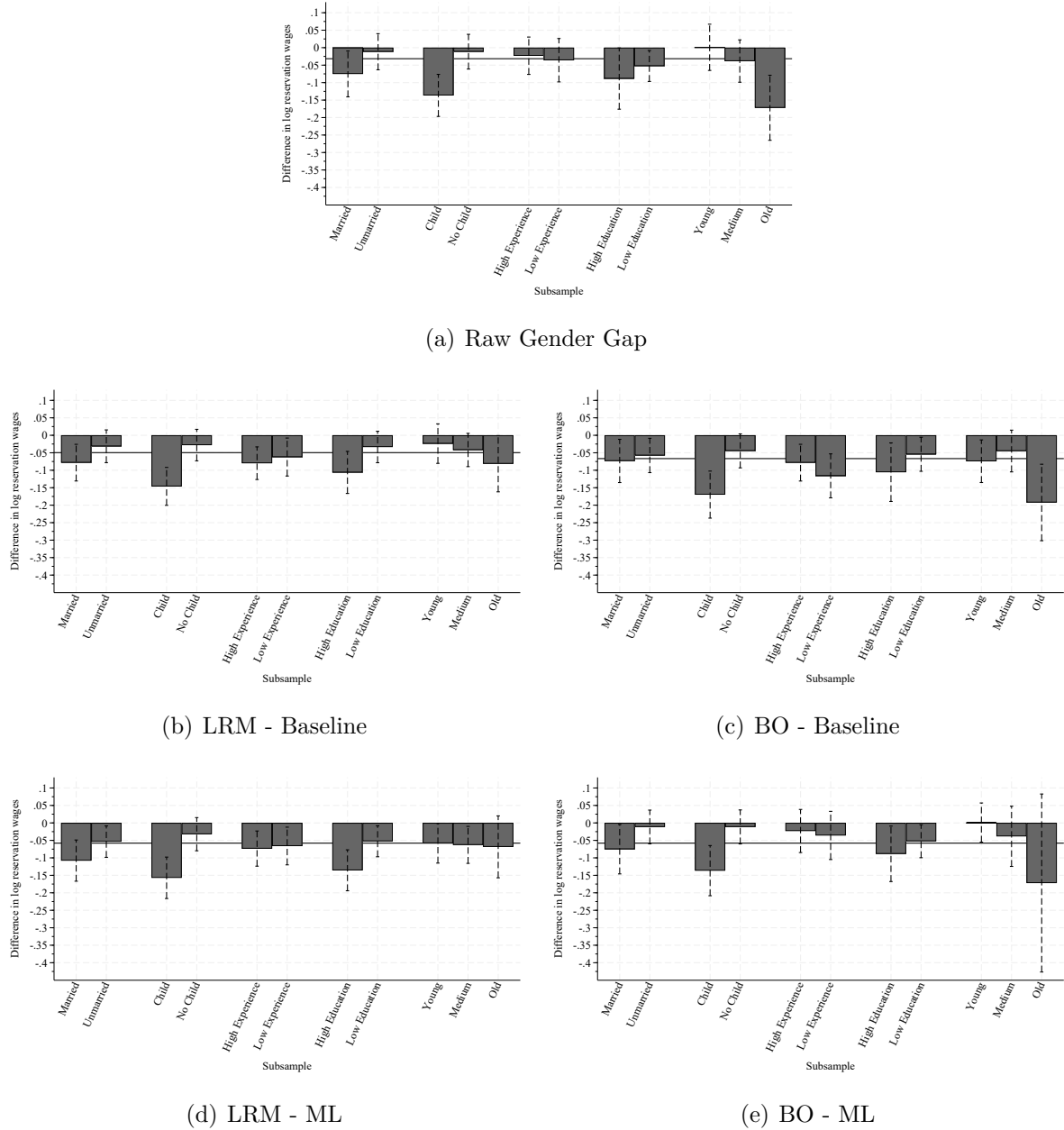
In this subsection, we compare the unexplained gender gap between various subgroups that are defined based on marital status, children, labor market experience, education, and age. To estimate the unexplained gender gap in those subgroups we use the LRM and BO estimators. We do not apply the AIPW estimator as some of the subgroups are fairly small which leads to very large standard errors in case of the AIPW. Apart from our baseline model specification, we again use also the ML model specifications. Our results from the subgroup analysis are represented in Figure 3.2.²³

First, we compare the gender gap between married and unmarried individuals. As the raw gender gaps, all unexplained gender gap estimates suggest that the differential is larger among married individuals. In case of the BO estimate with the baseline model specification, however, the difference between the two groups is very small. The results are not very sensitive to model specifications, while they change relatively more due to methodological choices (LRM and BO). Overall, the result of more pronounced gaps among married individuals may suggest that married men tend to be more confident in setting (higher) reservation wages compared to married women. The latter may be considered evidence for the persistent male breadwinner model. In line with this, Le Barbanchon et al. (2021) show that reservation wage disparities between males and females are higher among married individuals. Moreover, the finding is comparable to higher realized wages of married men (e.g., Antonovics and Town, 2004; Arntz et al., 2022; Bonilla et al., 2022).

Next, we differentiate between individuals with and without a child (below age 16). Our results suggest substantial differences in the size of the raw and unexplained gender gap between these two groups. Across all estimators and model specifications the gender gap estimates are considerably larger among individuals with a child. In fact, all estimates among individuals with a child range between 14% and 17% (in favor of men) and are at least twice as large as the respective aggregate estimates in the full sample.

²³Table 3.A.8 in the Appendix shows more detailed results.

Figure 3.2: Raw and Unexplained Gender Gaps in Reservation Wage within Subgroups



Notes: The bars display the raw gender gap (a) or unexplained gender gap estimates (b - e) within the respective subgroups. The dashed vertical lines display 95%-confidence intervals. The solid horizontal lines display the respective estimate in the full sample. Linear regression (LRM), Blinder-Oaxaca decomposition (BO), and Augmented Inverse Probability Weighting (AIPW) are used. Baseline model specification and machine learning model specifications (ML), are used. Standard errors are clustered at the individual level. Bootstrapped standard errors (2,000 replications, clustering at the individual level) are used in case of BO with ML model specifications. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Among individuals without a child, in contrast, the gender gap estimates are lower than in the full sample and are in most cases insignificant or only weakly significant. Similarly, Le Barbanchon et al. (2021) find higher gender reservation wage gaps for parents and only small gaps in favor of men for childless individuals. Higher gaps for parents may be linked to the child penalty of mothers that has been documented in the literature (e.g., Kleven et al., 2019; Fitzenberger et al., 2013). Mothers may anticipate lower wage offers after giving birth and consequently set lower reservation wages. This finding differs from the results of Brown et al. (2011) who find that the unexplained gender gap in the UK is larger for individuals without a child. Higher reservation wages among mothers may be related to their higher propensity to stay at home. Instead, to counteract this, fathers may have to search for jobs with higher intensity and thus may reduce their reservation wages (Browning, 1992). Both factors may lead to smaller gaps among parents. Our results of larger gaps among parents compared to non parents suggest, similar to Humpert and Pfeifer (2013), that the male breadwinner model and hence the traditional family division of labor plays a dominant role in labor supply and job search considerations.

Further, we group individuals based on their labor market experience.²⁴ If the unexplained gender gap in reservation wages is mainly due to anticipated discrimination against women, we may expect that the gender gap is lower among individuals with lower labor market experience since they are less likely to have experienced discrimination before (Caliendo et al., 2017). Similar to Caliendo et al. (2017) finding a larger gender gap for individuals with a high labor market experience, the LRM estimates of the gender gap are larger in the high experience group.²⁵ The corresponding BO estimates are larger in the low experience group, instead. The reason may be that men are especially confident in their wage requests at the beginning of their careers compared to their female colleagues. Evidence from the literature shows that the gender gap between men and women in the likelihood of negotiating wages is larger when starting a job (Biasi and Sarsons, 2021).

²⁴In order to group individuals into a low and high experience group, we follow Caliendo et al. (2017) and use the median of the ratio of labor market experience and age in years minus 18 to disentangle the age and experience effect.

²⁵Note that Caliendo et al. (2017) use linear regressions in their subgroup analysis.

This may counteract the smaller anticipated discrimination effect of females. The example shows that, depending on the estimator chosen, implications may change. However, in our case, differences in point estimates are negligible in three out of four estimators. The only exception is the baseline BO estimate, where we find a slightly more pronounced difference in estimated gaps among the two groups. Hence, overall, our results suggest no marked pattern of reservation wage differences among individuals with high and low past labor market experience.

Next, we group individuals into a low and high education group based on whether they have obtained the highest school-leaving degree in Germany (*Abitur*) or not. The raw gender gap and all estimates of the unexplained gender gap are larger among individuals with a higher educational attainment. While the estimates in the low education group only range between 3% and 8%, the estimates in the high education group range between 9% and 14%. Lastly, we analyze potential gender gap heterogeneity with respect to age. We divide our sample into a young (18 - 40), medium (41 - 55) and old (56 - 65) age group. All estimates of the unexplained gender gap are highest in the old age group. Yet, the results are very sensitive to the choice of estimator and model specification. Some estimates suggest rather small differences between the three age groups, while others suggest pronounced differences.

All in all, we find unexplained gender gaps in favor of men in all subgroups strengthening our conclusion from the results in the full sample. Moreover, we see that the gender gap is considerably larger among individuals with a child. The result indicates that unobserved factors related to having a child may be major drivers of the gender gap in reservation wages. Further, parents making their decisions based on traditional role models could explain this result. This explanation is also supported by our finding that the gender gap is larger among married individuals. In addition, our results suggest a considerable higher gender gap among the highly educated, which is in line with our finding that the gap is larger at the top of the reservation wage distribution. For the other subgroup comparisons, we either do not see substantial differences or the result

depends on the estimator and/or model specification. Yet, in line with the results from the main analysis, the sensitivity with respect to the model specification procedure is less pronounced.

3.6 Conclusion

This paper analyzes gender gaps in reservation wages in Germany using conventional and data-driven model specifications as well as various estimators. Reducing gender differences in reservation wages may alter traditional gender roles and close, at least in part, the gender wage gap. We use SOEP data offering a large set of information on individual characteristics and thus allowing us to control for a broad range of potential reservation wage determinants. The latter makes the conditional independence assumption plausible. We find a relatively small raw gender gap in log hourly reservation wages in our sample (3%) in favor of men. However, once we control for various characteristics of individuals, the results reveal a statistically significant unexplained gender gap in reservation wages of up to 8% between men and women. These estimates are comparable to those from the related literature.

We focus on the unexplained gender gap in reservation wages that may reflect discriminatory behavior against women, traditional role models and gender differences in social factors. As this gap may translate into persistent differences in actual pay, it is of interest to have a robust and as unbiased as possible estimate of this indicator. To achieve this, both methodological choices and model specifications may be important (Strittmatter and Wunsch, 2023). Our estimate of the unexplained gender gap in reservation wages is relatively robust across different model specifications, while it is more sensitive to the choice of the estimator. This finding underlines the importance of not relying on only one estimation method. The regression results further suggest that males and females differ with regard to reservation wage returns of several characteristics. For instance, having a college degree, being registered as unemployed, and receiving unemployment benefits

affect women differently compared to men when setting their reservation wage. Our BO decomposition results indicate that the explained and unexplained part of the gender gap point in opposite directions suggesting that women in our sample outperform men in terms of observable reservation wage determinants, while men enjoy higher returns given identical characteristics. This pattern persists across the reservation wage distribution.

Moreover, our findings from the analysis across the reservation wage distribution show that the unexplained gender gap is largest at the top of the distribution and thus mimics the glass ceiling effect in realized wages. As typically well-educated individuals are placed at this part of the distribution, (anticipated) unequal treatment of women may be particularly pronounced for high-educated women, which is supported by our results from the subgroup analysis. Furthermore, we find that the unexplained difference between women and men in their reservation wage is particularly large among individuals living in a household with a child underlining the relevance of children for gender differences in reservation wages and the related consumption-leisure trade-offs of women. A caveat of this study is that we cannot entirely rule out sample selection bias. Therefore, our estimation results are only interpretable conditional on the considered sample.

Overall, our findings provide strong empirical evidence that German female job searchers looking for a full-time job and intending to be employed within one year set considerably lower reservation wages than comparable males. As a consequence and in order to circumvent that gender gaps in reservation wages translate into actual gender pay gaps, fighting gender differences in reservation wages should be on the political agenda. Particular emphasis should be on high-educated women as well as on mothers and on their (re-)integration in the labor market. Flexible working arrangements, childcare services and parental leave programs that distribute child care duties more equally between men and women may be a powerful tool to achieve this goal. Further research focusing on this group may provide helpful insights and may allow (more precisely) tailored policy recommendations to fight gender gaps in reservation wages.

References

- Addison, J. T., Centeno, M., and Portugal, P. (2009). Do Reservation Wages Really Decline? Some International Evidence on the Determinants of Reservation Wages. *Journal of Labor Research*, 30(1):1–8.
- Ahrens, A., Hansen, C. B., and Schaffer, M. E. (2022). pystacked: Stacking generalization and machine learning in Stata. Papers 2208.10896, arXiv.org.
- Ahrens, A., Hansen, C. B., Schaffer, M. E., and Wiemann, T. (2024). ddml: Double/debiased machine learning in Stata. *The Stata Journal*, 24(1):3–45.
- Antonovics, K. and Town, R. (2004). Are all the good men married? Uncovering the sources of the marital wage premium. *American Economic Review*, 94(2):317–321.
- Arntz, M., Ben Yahmed, S., and Berlingieri, F. (2022). Working from home, hours worked and wages: Heterogeneity by gender and parenthood. *Labour Economics*, 76:102169.
- Belloni, A., Chen, D., Chernozhukov, V., and Hansen, C. (2012). Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain. *Econometrica*, 80(6):2369–2429.
- Belloni, A., Chernozhukov, V., and Hansen, C. (2014a). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2):29–50.
- Belloni, A., Chernozhukov, V., and Hansen, C. (2014b). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2):608–650.
- Bertrand, M. (2011). Chapter 17 - New Perspectives on Gender. volume 4 of *Handbook of Labor Economics*, pages 1543–1590. Elsevier.
- Biasi, B. and Sarsons, H. (2021). Information, confidence, and the gender gap in bargaining. In *AEA Papers and Proceedings*, volume 111, pages 174–178. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.
- Blau, F. D. and Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3):789–865.
- Blau, F. D. and Winkler, A. E. (2022). *The economics of women men and work (9th ed.)*. Oxford University Press, Oxford (UK).

- Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources*, 8(4):436–455.
- Bonaccolto-Töpfer, M. and Bonaccolto, G. (2023). Gender wage inequality: new evidence from penalized expectile regression. *The Journal of Economic Inequality*, 21(3):511–535.
- Bonaccolto-Töpfer, M. and Briel, S. (2022). The gender pay gap revisited: Does machine learning offer new insights? *Labour Economics*, 78:102223.
- Bonilla, R., Malo, M. ., and Pinto, F. (2022). Marriage wage premium with contract type heterogeneity. *Labour Economics*, 78:102198.
- Breiman, L. (1996). Stacked regressions. *Machine Learning*, 24(1):49–64.
- Briel, S., Osikominu, A., Pfeifer, G., Reutter, M., and Satlukal, S. (2021). Gender differences in wage expectations: the role of biased beliefs. *Empirical Economics*, forthcoming.
- Brown, S., Popli, G., and Sasso, A. (2022). Decomposing the gender reservation wage gap in Italy: A regional perspective. *Journal of Regional Science*, 62(2):499–540.
- Brown, S., Roberts, J., and Taylor, K. (2011). The gender reservation wage gap: Evidence from British Panel data. *Economics Letters*, 113(1):88–91.
- Brown, S. and Taylor, K. (2015). The reservation wage curve: Evidence from the UK. *Economics Letters*, 126:22–24.
- Browning, M. (1992). Children and household economic behavior. *Journal of Economic Literature*, 30(3):1434–1475.
- Busso, M., DiNardo, J., and McCrary, J. (2014). New Evidence on the Finite Sample Properties of Propensity Score Reweighting and Matching Estimators. *The Review of Economics and Statistics*, 96(5):885–897.
- 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*, 136:161–173.
- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., and Newey, W. (2017). Double/Debiased/Neyman Machine Learning of Treatment Effects. *American Economic Review*, 107(5):261–65.

- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68.
- Chernozhukov, V., Fernández-Val, I., and Melly, B. (2013). Inference on counterfactual distributions. *Econometrica*, 81(6):2205–2268.
- Christensen, B. (2001). The Determinants of Reservation Wages in Germany Does a Motivation Gap Exist? (No. 1024). Technical report, Kiel Working Paper No. 1024, Kiel Institute of World Economics.
- Collischon, M. (2019). Is There a Glass Ceiling over Germany? *German Economic Review*, 20(4):e329–e359.
- Collischon, M. (2020). The Returns to Personality Traits Across the Wage Distribution. *LABOUR*, 34(1):48–79.
- Cortés, P., Pan, J., Pilossoph, L., Reuben, E., and Zafar, B. (2023). Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab*. *The Quarterly Journal of Economics*, 138(4):2069–2126.
- DiNardo, J., Fortin, N. M., and Lemieux, T. (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica*, 64(5):1001–1044.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the european economic association*, 9(3):522–550.
- Dougherty, C. (2006). The marriage earnings premium as a distributed fixed effect. *Journal of Human Resources*, 41(2):433–443.
- Fedorets, A. and Shupe, C. (2021). Great expectations: Reservation wages and minimum wage reform. *Journal of Economic Behavior & Organization*, 183:397–419.
- Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, 77(3):953–973.
- Firpo, S., Fortin, N. M., and Lemieux, T. (2018). Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics*, 6(2).
- Fitzenberger, B., Sommerfeld, K., and Steffes, S. (2013). Causal effects on employment after first birth — A dynamic treatment approach. *Labour Economics*, 25:49–62.

- Fortin, N., Lemieux, T., and Firpo, S. (2011). Chapter 1 - Decomposition Methods in Economics. volume 4 of *Handbook of Labor Economics*, pages 1–102. Elsevier.
- Frank, L. E. and Friedman, J. H. (1993). A Statistical View of Some Chemometrics Regression Tools. *Technometrics*, 35(2):109–135.
- Gelbach, J. B. (2016). When do covariates matter? And which ones, and how much? *Journal of Labor Economics*, 34(2):509–543.
- Gensowski, M. (2018). Personality, IQ, and lifetime earnings. *Labour Economics*, 51:170–183.
- Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119.
- Goldin, C., Katz, L. F., and Kuziemko, I. (2006). The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic Perspectives*, 20(4):133–156.
- Goldin, Claudia and Mitchell, Joshua (2017). The New Life Cycle of Women’s Employment: Disappearing Humps, Sagging Middles, Expanding Tops. *Journal of Economic Perspectives*, 31(1):161–82.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). Linear methods for regression. In *The elements of statistical learning*, pages 43–99. Springer.
- Hirano, K., Imbens, G. W., and Ridder, G. (2003). Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score. *Econometrica*, 71(4):1161–1189.
- Huber, M. (2015). Causal pitfalls in the decomposition of wage gaps. *Journal of Business & Economic Statistics*, 33(2):179–191.
- Huber, M. and Solovyeva, A. (2020). On the sensitivity of wage gap decompositions. *Journal of Labor Research*, 41:1–33.
- Humpert, S. and Pfeifer, C. (2013). Explaining age and gender differences in employment rates: a labor supply-side perspective. *Journal for Labour Market Research*, 46(1):1–17.
- Imbens, G. W. (2004). Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. *The Review of Economics and Statistics*, 86(1):4–29.
- Jann, B. (2008). The Blinder–Oaxaca Decomposition for Linear Regression Models. *Stata Journal*, 8(4):453–479.

-
- Khan, S. and Tamer, E. (2010). Irregular Identification, Support Conditions, and Inverse Weight Estimation. *Econometrica*, 78(6):2021–2042.
- Kleven, H., Landais, C., and Søgaaard, J. E. (2019). Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209.
- Kline, P. (2011). Oaxaca-Blinder as a Reweighting Estimator. *American Economic Review*, 101(3):532–37.
- Knaus, M. C. (2022). Double machine learning-based programme evaluation under unconfoundedness. *The Econometrics Journal*, 25(3):602–627.
- Le Barbanchon, T., Rathelot, R., and Roulet, A. (2021). Gender Differences in Job Search: Trading off Commute against Wage. *The Quarterly Journal of Economics*, 136(1):381–426.
- Machado, J. A. and Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, 20(4):445–465.
- McGee, A. and McGee, P. (2023). Gender Differences in Reservation Wages in Search Experiments. Technical report, IZA Discussion Paper No. 16577, Institute of Labor Economics (IZA).
- Melly, B. (2005). Decomposition of differences in distribution using quantile regression. *Labour Economics*, 12(4):577–590.
- Niederle, M. (2017). Chapter 8 - Gender. volume 2 of *The Handbook of Experimental Economics*, pages 481–562. Princeton University Press, Princeton.
- Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14(3):693–709.
- Ñopo, H. (2008). Matching as a Tool to Decompose Wage Gaps. *The Review of Economics and Statistics*, 90(2):290–299.
- Pannenberg, M. (2010). Risk attitudes and reservation wages of unemployed workers: Evidence from panel data. *Economics Letters*, 106(3):223–226.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

- Reuben, E., Wiswall, M., and Zafar, B. (2017). Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender. *The Economic Journal*, 127(604):2153–2186.
- Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994). Estimation of Regression Coefficients When Some Regressors are not Always Observed. *Journal of the American Statistical Association*, 89(427):846–866.
- Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688.
- Schäfer, K. C. and Schwiebert, J. (2018). The Impact of Personality Traits on Wage Growth and the Gender Wage Gap. *Bulletin of Economic Research*, 70(1):20–34.
- Strittmatter, A. and Wunsch, C. (2023). Labor Market Sorting and the Gender Pay Gap Revisited. Mimeo, available at <http://dx.doi.org/10.2139/ssrn.4624731>.
- Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.
- Weisberg, Y., DeYoung, C., and Hirsh, J. (2011). Gender Differences in Personality across the Ten Aspects of the Big Five. *Frontiers in Psychology*, 2.
- Wilde, E. T., Batchelder, L., and Ellwood, D. T. (2010). The mommy track divides: The impact of childbearing on wages of women of differing skill levels. Technical report, National Bureau of Economic Research.
- Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, 5(2):241–259.

Appendix

Table 3.A.1: Estimates of Unexplained Gender Gap in Reservation Wages with and without controlling for Family Background

	LRM		BO	
	(1)	(2)	(3)	(3)
Unexplained Gender Gap	-0.0496** (0.0201)	-0.0481** (0.202)	-0.0667*** (0.0213)	-0.0601*** (0.0215)
Family Background		✓		✓
Observations	4,114	4,114	4,114	4,114

Notes: The table shows estimates of the unexplained gender difference in log hourly reservation wages with different model specifications (LRM and BO estimator). While the estimates in Column (1) and (2) use linear regressions (LRM), the estimates in Column (3) and (4) use Blinder-Oaxaca decompositions (BO). The estimates in Column (1) and (3) rely on the baseline model specification specified above. In Column (2) and (4) family background controls are added to the baseline model specification. Family background controls include dummies for the highest level of parental education: basic secondary education, secondary education, upper secondary education, college of the mother and father, respectively. Standard errors are clustered at the individual level. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 3.A.2: Summary Statistics (Registered as Unemployed Sample)

	(1)	(2)	(3)
	Females	Males	Difference
Monthly Reservation Wage	1320.17 (449.97)	1447.95 (521.85)	-127.78***
Working Hours	37.95 (5.74)	39.73 (4.46)	-1.78***
Hourly Reservation Wage	8.33 (3.13)	8.65 (3.22)	-0.32***
Observations	1,221	2,000	3,221

Notes: Columns (1) and (2) display mean values of variables in female and male subsample, respectively. Standard deviations are shown in parenthesis. Column (3) shows the mean difference between females and males. *, **, and *** indicate statistical significant differences at the 10%-, 5%-, and 1%-level based on a t-test for equal means, respectively. Reservation wages are CPI adjusted (base year 2015). SOEP sample weights used. *Source:* SOEP v37, own calculations.

Table 3.A.3: Reservation Wage Regressions

Dependent Variable: Log Hourly Reservation Wage				
	(1)	(2)	(3)	(4)
	Raw	Controls	Males	Females

Continued on next page

Female (Dummy)	-0.0316 (0.0218)	-0.0496** (0.0201)		
Age		-0.0016 (0.0014)	-0.0000 (0.0024)	-0.0028 (0.0018)
Married (Dummy)		0.0470** (0.0185)	0.0360 (0.0229)	0.0390 (0.0274)
Log Household Monthly Income		0.0531*** (0.0164)	0.0432** (0.0194)	0.0604** (0.0242)
Household Size		-0.0097 (0.0082)	-0.0067 (0.0101)	-0.0086 (0.0118)
Partner Works Full-Time (Dummy)		-0.0574*** (0.0196)	-0.0527* (0.0285)	-0.0556** (0.0267)
Partner Works Part-Time (Dummy)		0.0084 (0.0259)	-0.0022 (0.0276)	0.0707 (0.0740)
Partner Unemployed (Dummy)		0.0480** (0.0224)	0.0635** (0.0256)	-0.0120 (0.0394)
Youngest Child ≤ 3 (Dummy)		0.0095 (0.0441)	0.0074 (0.0512)	0.0379 (0.0895)
Youngest Child ≥ 3 and ≤ 6 (Dummy)		0.0594 (0.0363)	0.0232 (0.0463)	0.0571 (0.0527)
Youngest Child ≥ 6 and ≤ 18 (Dummy)		-0.0250 (0.0211)	-0.0043 (0.0274)	-0.0377 (0.0266)
Childcare (Dummy)		0.0450*** (0.0169)	0.0691*** (0.0230)	0.0221 (0.0223)
Medium Health (Dummy)		-0.0024 (0.0158)	-0.0230 (0.0202)	0.0162 (0.0220)
Good Health (Dummy)		0.0275 (0.0168)	0.0186 (0.0204)	0.0409* (0.0224)
Migration Background (Dummy)		-0.0037 (0.0231)	-0.0001 (0.0291)	-0.0041 (0.0283)
No Degree (Dummy)		-0.0118 (0.0469)	0.0096 (0.0544)	-0.0520 (0.0828)
Other Degree (Dummy)		0.0250 (0.0483)	0.0247 (0.0563)	-0.0289 (0.0685)
Basic Secondary Education (Hauptschule) (Dummy)		0.0005 (0.0395)	0.0138 (0.0460)	-0.0287 (0.0558)
Secondary Education (Realschule) (Dummy)		0.0150 (0.0395)	0.0525 (0.0482)	-0.0545 (0.0519)
Upper Secondary Education (Abitur) (Dummy)		0.0721* (0.0429)	0.1168** (0.0538)	-0.0172 (0.0539)
Vocational Degree		0.0190 (0.0174)	0.0192 (0.0222)	-0.0117 (0.0237)
College Degree		0.1198*** (0.0293)	0.1117*** (0.0404)	0.1563*** (0.0333)

Continued on next page

Full-Time Experience	0.0032** (0.0014)	0.0020 (0.0023)	0.0033 (0.0022)
Part-Time Experience	-0.0012 (0.0025)	-0.0046 (0.0045)	0.0020 (0.0029)
Unemployment Experience	0.0014 (0.0021)	-0.0018 (0.0028)	0.0042 (0.0034)
Log Hourly Wage Last Job	0.0396*** (0.0136)	0.0413** (0.0163)	0.0333* (0.0187)
Registered as Unemployed (Dummy)	-0.0682*** (0.0224)	-0.0599** (0.0296)	-0.0943*** (0.0298)
Receiving ALG (Dummy)	-0.1162*** (0.0267)	-0.1431*** (0.0343)	-0.0891** (0.0396)
Receiving ALG II (Dummy)	-0.0064 (0.0196)	-0.0321 (0.0254)	0.0374 (0.0255)
Amount of ALG (in €)	0.0001*** (0.0000)	0.0001*** (0.0000)	0.0002*** (0.0000)
Finding Vacancy Difficult (Dummy)	-0.0918*** (0.0205)	-0.1087*** (0.0263)	-0.0476* (0.0268)
Finding Vacancy Impossible (Dummy)	-0.1192*** (0.0246)	-0.1413*** (0.0321)	-0.0507 (0.0347)
Could Accept Job Immediately (Dummy)	0.0104 (0.0231)	-0.0263 (0.0292)	0.0284 (0.0337)
Active Job Search in Last Four Weeks (Dummy)	0.0140 (0.0154)	0.0243 (0.0173)	0.0091 (0.0232)
Employment Intended Immediately (Dummy)	-0.0490*** (0.0188)	-0.0362 (0.0242)	-0.0546** (0.0263)
Future Employment Rather Unlikely	0.0127 (0.0275)	0.0388 (0.0339)	-0.0193 (0.0439)
Openness	0.0037 (0.0058)	-0.0059 (0.0078)	0.0151 (0.0096)
Conscientiousness	0.0121 (0.0077)	0.0095 (0.0089)	0.0208* (0.0117)
Extraversion	0.0124* (0.0064)	0.0206*** (0.0077)	0.0053 (0.0096)
Neuroticism	0.0036 (0.0059)	0.0091 (0.0074)	-0.0021 (0.0081)
Agreeableness	0.0027 (0.0067)	-0.0034 (0.0081)	0.0122 (0.0100)
Risk Tolerance	0.0021 (0.0030)	0.0007 (0.0035)	0.0022 (0.0045)
Log Unemployment Rate	-0.0281 (0.0600)	-0.0922 (0.0823)	0.1000 (0.0816)
Industry Last Job (ISCO88 1-digit)	✓	✓	✓
Occupation Last Job (NACE Level 1)	✓	✓	✓

Continued on next page

Federal State Fixed Effects		✓	✓	✓
Survey Year Fixed Effects		✓	✓	✓
Observations	4,114	4,114	2,419	1,695

Notes: The dependent variable is the log of the hourly reservation wage. Standard errors are clustered at the individual level and are shown in parenthesis. *, ** and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 3.A.4: Reservation Wage Decomposition, Female Reference Category

(a) Aggregate Decomposition		
Raw Gender Gap	-0.0316	
	(0.0214)	
Explained Part	-0.0022	
	(0.0247)	
Unexplained Part	-0.0294	
	(0.0246)	
(b) Detailed Decomposition		
Set of Control	Contribution	
	Explained Part	Unexplained Part
Household	0.0070*	0.0004
	(0.0039)	(0.0310)
Partner	-0.0133**	0.0006
	(0.0064)	(0.0127)
Children	0.0013	-0.0121*
	(0.0022)	(0.0064)
Education	0.0145**	-0.0929
	(0.0063)	(0.0677)
Labor Market History	-0.0343*	0.0354
	(0.0176)	(0.0830)
Unemployment	0.0032	0.0227
	(0.0043)	(0.0306)
Job Search	0.0023	0.0737*
	(0.0026)	(0.0405)
Personality	0.0149*	0.1288*
	(0.0084)	(0.0676)
Remainder	0.0023	0.4000
	(0.0068)	(0.3248)
Constant		-0.5859*
		(0.3459)
Observations	7,159	

Notes: The dependent variable is the log of the hourly reservation wage. The female coefficient vector is used as the reference vector. Explanatory variables are grouped as follows. **Partner:** marital status, employment status of partner. **Household:** log monthly household income, household size. **Children:** age of youngest child, childcare. **Education:** basic secondary education, secondary education, upper secondary education, vocational degree, college degree. **Labor Market History:** full-time experience, part-time experience, unemployment experience, wage last job, industry last job, occupation last job. **Unemployment:** registered as unemployed, receiving unemployment benefits (ALG and ALG II), amount of ALG. **Job Search:** difficulty of finding fitting vacancy, could accept job immediately, active job search in last four weeks, employment intended immediately, future employment rather unlikely. **Personality:** Big Five Personality Traits (openness, conscientiousness, extraversion, neuroticism, agreeableness), risk tolerance. **Remainder:** age, migration background, health, log unemployment rate, federal state, survey year. The standard errors are calculated as described by Jann (2008), are clustered at the individual level, and are shown in parenthesis. *, ** and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 3.A.5: Estimates of Unexplained Gender Gap in Reservation Wages, Baseline Model Specification

	(1)	(2)	(3)
	LRM	BO	AIPW
Unexplained Gender Gap	-0.0496** (0.0201)	-0.0667*** (0.0213)	-0.0762*** (0.0273)
Observations	4,114	4,114	4,114
Female Observations	1,695	1,695	1,695
Number of Controls	92	92	92

Notes: The table shows different estimates of the unexplained gender difference in log hourly reservation wages. Column (1) linear regression. Column (2) Blinder-Oaxaca decomposition. Column (3) Augmented Inverse Probability Weighting. Standard errors are clustered at the individual level. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 3.A.6: Estimates of Unexplained Gender Gap in Reservation Wages, ML Model Specifications

	(1)	(2)	(3)
	LRM	BO	AIPW
Unexplained Gender Gap	-0.0579*** (0.0199)	-0.0802*** (0.0214)	-0.0549** (0.0221)
Observations	4,114	4,114	4,114
Female Observations	1,695	1,695	1,695
Selected Controls	71	46	

Notes: The table shows different estimates of the unexplained gender difference in log hourly reservation wages based on machine learning model specifications. Column (1) linear regression. Column (2) Blinder-Oaxaca decomposition. Column (3) Augmented Inverse Probability Weighting. The LRM and BO estimators are based on the post-double-selection procedure by Belloni et al. (2014a,b). The AIPW estimator is based on the double/debiased machine learning approach by Chernozhukov et al. (2017, 2018) and the nuisance parameters (propensity score and predicted male reservation wage) are estimated based on an ensemble method that optimally combines four machine learning methods (Random Forest, Gradient Boosting, Neural Net, and the LASSO). Standard errors are clustered at the individual level. Bootstrapped standard errors (2,000 replications, clustering at the individual level) are used in case of BO. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 3.A.7: Estimates of the Gender Gap in Reservation Wages (ML Model) and of the Aggregate RIF-Decomposition (Baseline Model) across the Distribution

	(1)	(2)	(3)	(4)	(5)
RIF at Percentile	10%	25%	50%	75%	90%
(a) Aggregate BO RIF decomposition (Baseline Model)					
Unexplained Part	0.0560 (0.0430)	-0.0120 (0.0275)	-0.0529* (0.0283)	-0.0523 (0.0333)	-0.1862*** (0.0644)
Explained Part	-0.0476 (0.0338)	-0.0038 (0.0227)	0.0134 (0.0264)	0.0418 (0.0301)	0.1049* (0.0545)
(b) Unexplained Gender Gap Estimates (ML Model)					
RIF-OLS (LRM)	-0.0255 (0.0283)	-0.0589*** (0.0180)	-0.0371* (0.0190)	-0.0285 (0.0191)	-0.0678* (0.0350)
BO	0.0520 (0.0318)	-0.0242*** (0.0219)	-0.0678** (0.0265)	-0.0558* (0.0315)	-0.1879*** (0.0581)
Observations	4,114	4,114	4,114	4,114	4,114

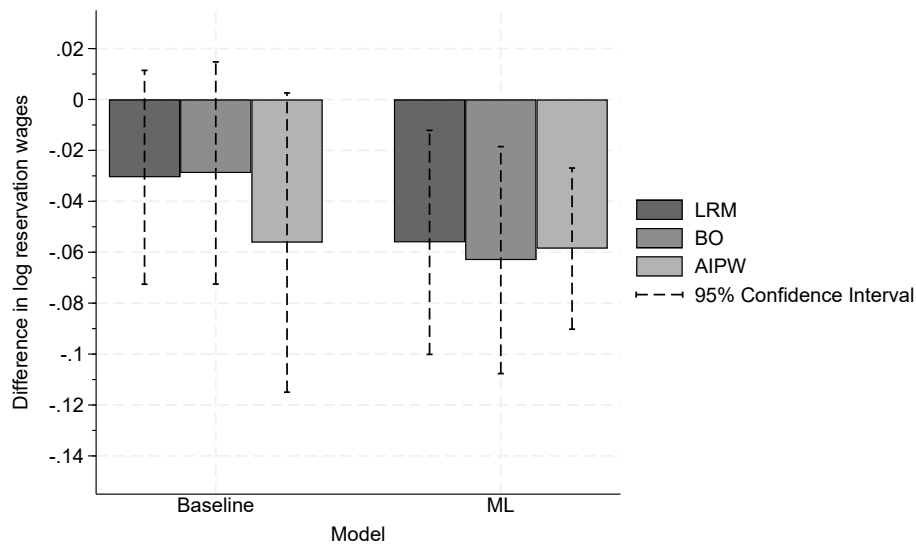
Notes: The dependent variable is the survey year adjusted log hourly reservation wage. The raw gap is calculated as $RIF(q_\tau y_F) - RIF(q_\tau, y_M)$, where $RIF(q_\tau, y_G)$ refers the RIF of G at percentile q_τ with $G = F, M$. Male reference category used for RIF-Decompositions. Bootstrapped standard errors (2,000 replications) are clustered at the individual level and are shown in parenthesis. *, ** and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 3.A.8: Estimates of Gender Gap in Reservation Wages within Subgroups

	(1)	(2)	(3)	(4)	(5)	(6)
		Unexplained Gender Gap Estimates				
		Baseline		ML		
	RGG	LRM	BO	LRM	BO	Obs.
(a) Marital Status						
Married	-0.0751** (0.0273)	-0.0783*** (0.0267)	-0.0734** (0.0316)	-0.1079*** (0.0301)	-0.0751** (0.0361)	1,701
Unmarried	-0.0116 (0.0317)	-0.0314 (0.0239)	-0.0577** (0.0251)	-0.0535** (0.0230)	-0.0116 (0.0246)	2,413
(b) Child in Household						
Child	-0.1366*** (0.0307)	-0.1461*** (0.0277)	-0.1697*** (0.0343)	-0.1571*** (0.0302)	-0.1366*** (0.0366)	1,643
No Child	-0.0115 (0.0253)	-0.0280 (0.0230)	-0.0446* (0.0248)	-0.0322 (0.0242)	-0.0115 (0.0248)	2,430
(c) Labor Market Experience						
High	-0.0231 (0.0273)	-0.0798*** (0.0238)	-0.0783*** (0.0268)	-0.0735*** (0.0257)	-0.0231 (0.0315)	2,054
Low	-0.0358 (0.0317)	-0.0622** (0.0277)	-0.1161*** (0.0321)	-0.0661** (0.0275)	-0.0358 (0.0351)	2,060
(d) Education						
High	-0.0884** (0.0447)	-0.1063*** (0.0307)	-0.1055** (0.0427)	-0.1356*** (0.0299)	-0.0884** (0.0405)	828
Low	-0.0524** (0.0226)	-0.0335 (0.0230)	-0.0544** (0.0248)	-0.0526** (0.0226)	-0.0524** (0.0241)	3,286
(e) Age						
Young	0.0010 (0.0337)	-0.0238 (0.0288)	-0.0742** (0.0311)	-0.0581** (0.0290)	0.0010 (0.0286)	2,159
Medium	-0.0383 (0.0309)	-0.0421* (0.0243)	-0.0450 (0.0394)	-0.0629** (0.0270)	-0.0383 (0.0440)	1,428
Old	-0.1720*** (0.0475)	-0.0813** (0.0410)	-0.1922*** (0.0558)	-0.0685 (0.0452)	-0.1720 (0.1299)	527

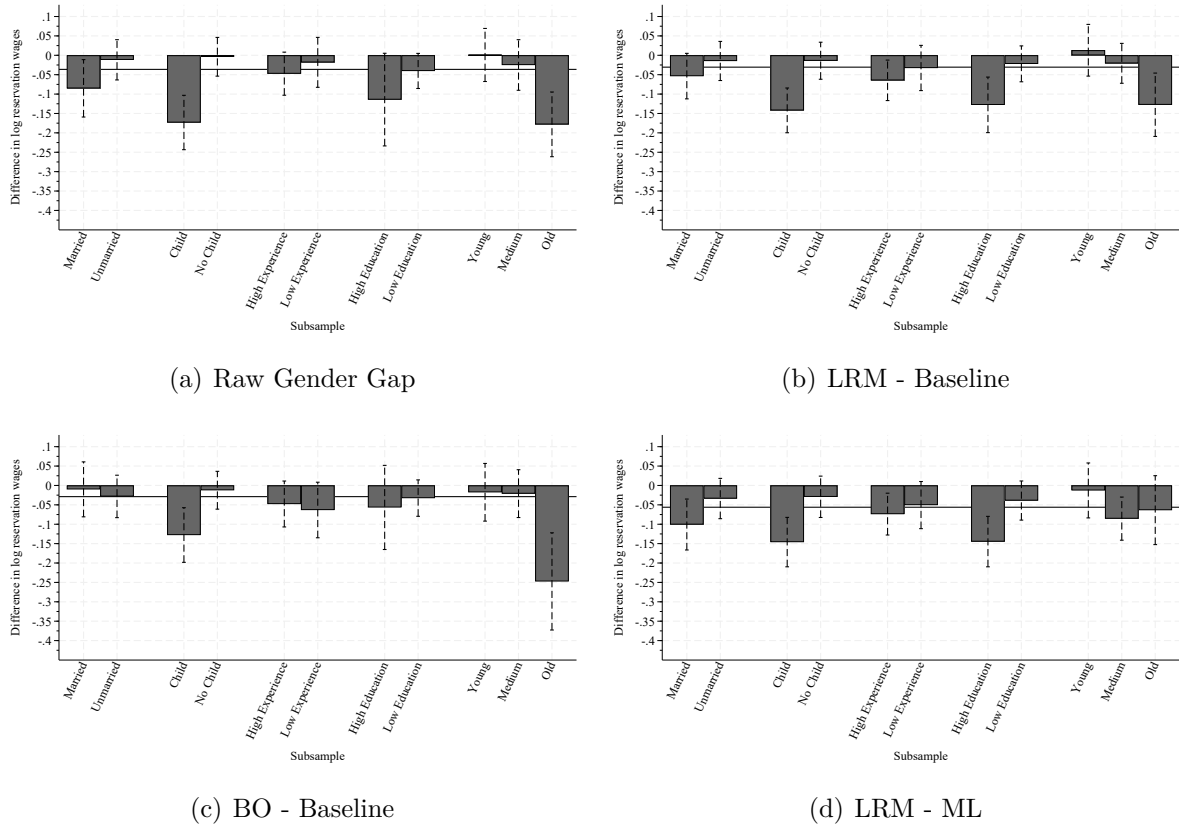
Notes: The table displays raw gender gaps (RGG) as well as linear regression (LRM) and Blinder-Oaxaca decomposition (BO) estimates of the unexplained gender gap within the respective subgroup. Baseline and machine learning (ML) model specifications are used. Standard errors are clustered at the individual level. Bootstrapped standard errors (2,000 replications, clustering at the individual level) are used in case of BO with ML model specifications. *, ** and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 3.A.1: Unexplained Gender Gap in Reservation Wages, Various Estimators (Registered as Unemployed Sample)



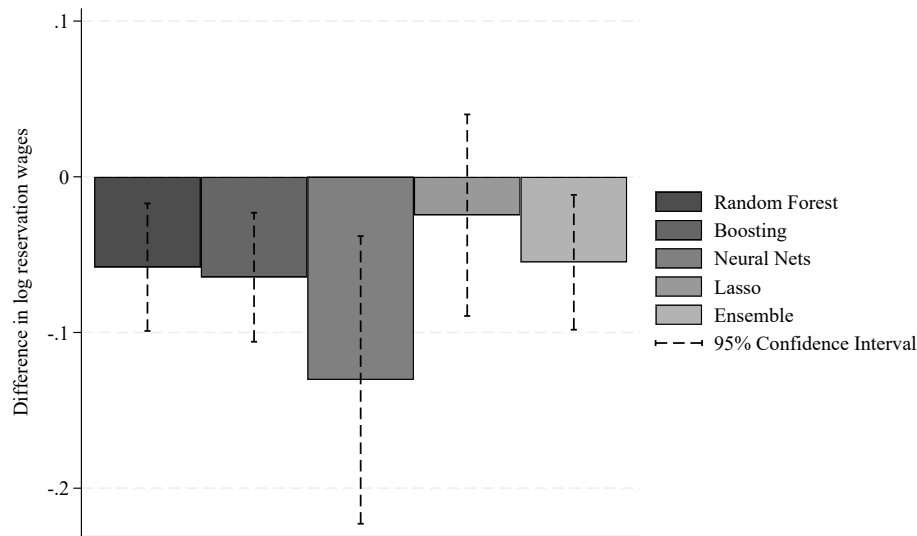
Notes: The figure shows estimates of the unexplained gender gap in reservation wages using an alternative sample that only includes individuals who are registered as unemployed but that is not restricted with respect to intended starting time of employment. Baseline model specification and machine learning model specifications (ML) are used. Standard errors are clustered at the individual level. Bootstrapped standard errors (2,000 replications, clustering at the individual level) are used in case of BO with the ML specification. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 3.A.2: Raw and Unexplained Gender Gaps in Reservation Wage within Subgroups (Registered as Unemployed Sample)



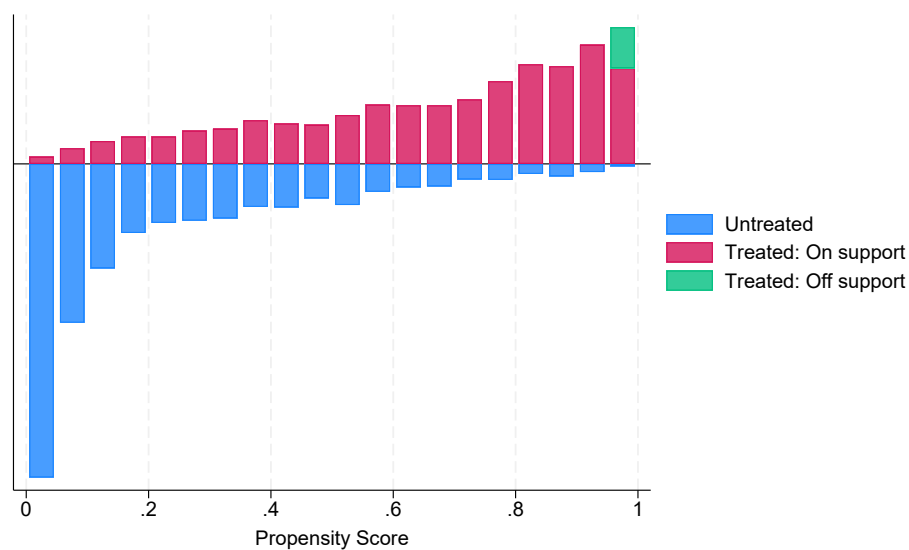
Notes: The results are based on an alternative sample that only includes individuals who are registered as unemployed but that is not restricted with respect to intended starting time of employment. The bars display the raw gender gap (a) or the unexplained gender gap estimates (b - d) within the respective subgroups. The dashed vertical lines display 95%-confidence intervals. The solid horizontal lines display the respective estimate in the full sample. Standard errors are clustered at the individual level. We do not show the BO estimator with the machine learning model specification as some of the subgroups are fairly small which leads to very large standard errors in this case. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 3.A.3: Unexplained Gender Reservation Wage Gap AIPW Estimates With Various ML Methods



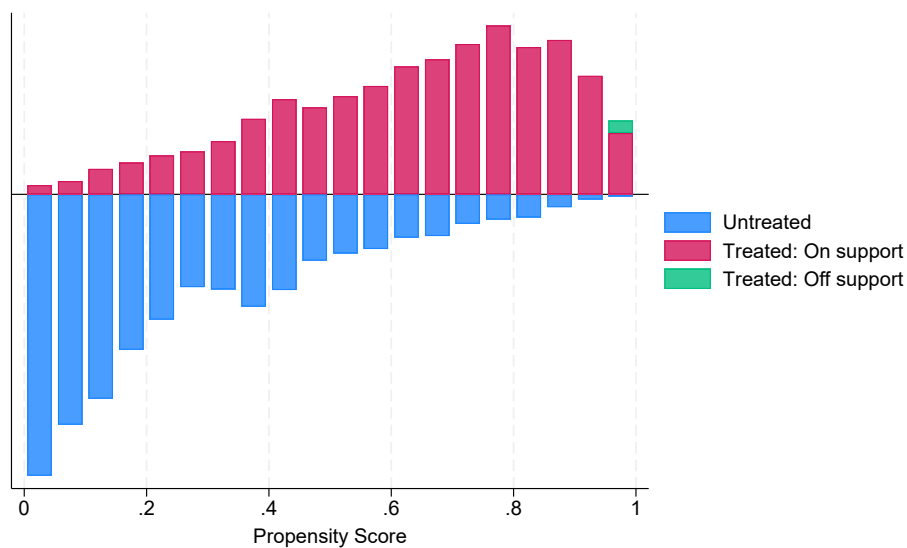
Notes: The figure shows estimates of the unexplained gender gap in reservation wages using augmented inverse probability weighting (AIPW) with different machine learning methods. The nuisance functions (propensity score and predicted male reservation wage) are estimated with Random Forest, Boosting, Neural Net, LASSO, and an Ensemble method that optimally combines the four other machine learning methods. Standard errors are clustered at the individual level. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 3.A.4: Distribution of Propensity Score with Baseline Model



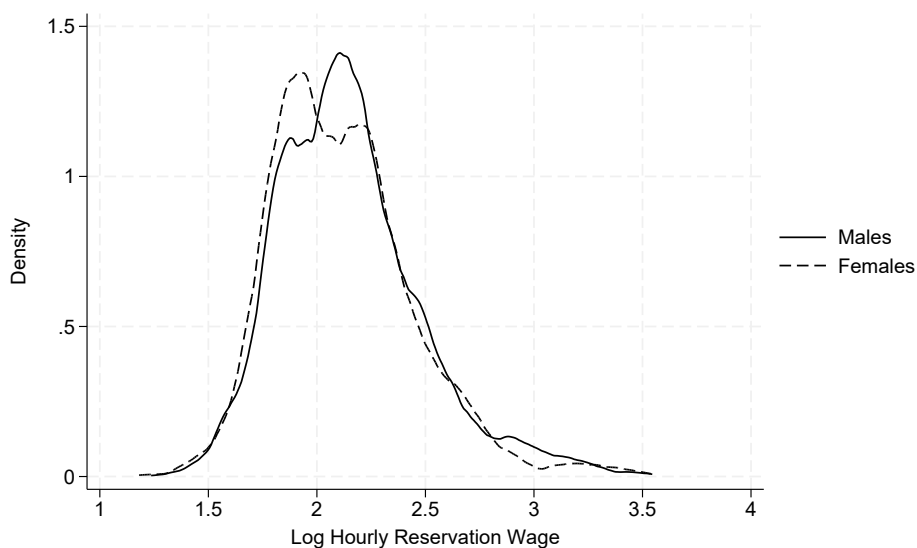
Notes: The figure shows the histogram of propensity scores estimated with logit model for females (treated) and males (untreated). *Source:* German Socio-Economic Panel v37, own calculations.

Figure 3.A.5: Distribution of Propensity Score with ML



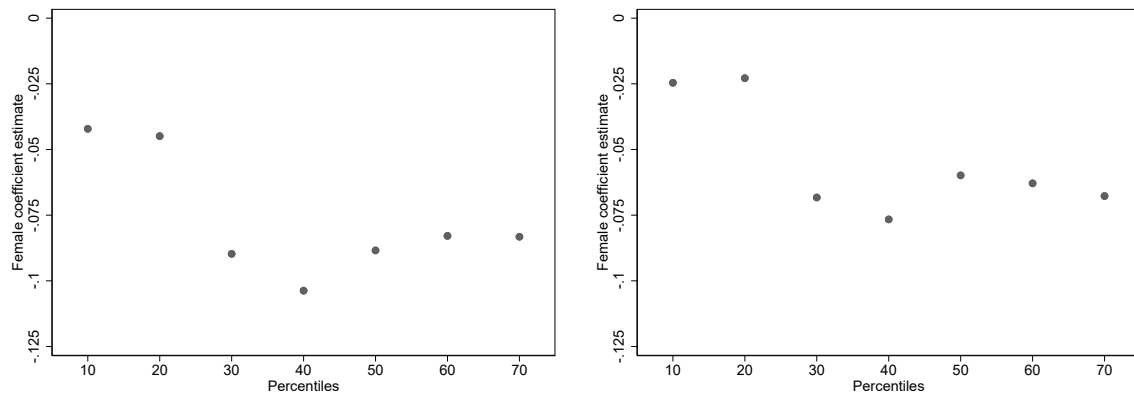
Notes: The figure shows the histogram of propensity scores estimated with Ensemble Machine Learning (ML) method for females (treated) and males (untreated). *Source:* German Socio-Economic Panel v37, own calculations.

Figure 3.A.6: Distribution of Hourly Reservation Wage by Gender



Notes: The figure shows kernel density estimates of log hourly reservation wages by gender. Hourly reservation wages are CPI adjusted (base year 2015). SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 3.A.7: Pattern of the Unexplained Gender Reservation Wage Gap across the Distribution, LRM



(a) Specification similar to Fedorets and Shupe (2021)

(b) Our Baseline Specification

Notes: The figure shows coefficient estimates of a female dummy from RIF regression at selected percentiles τ , where $\tau = 10, 20, 30, \dots, 70$. Control variables in the specification in panel (a) include age, past full-time and part-time experience as well as past unemployment experience. Further, we add dummies for secondary education, basic secondary education, upper secondary education, vocational degree, college degree, migration background, children under 16 in the household and being married. Baseline specification in panel (b) as specified above. *Source:* German Socio-Economic Panel v37, own calculations over the period 2010-2017.

Chapter 4

Gender Differences in Labor Market Expectations and Wages

Sascha Satlukal (University of Hohenheim)

Abstract: This article investigates gender differences in individual labor market expectations and their impact on gender gaps in wages and reservation wages. For this analysis, I use survey data from the German Socio-Economic Panel that include elicited beliefs about job loss and job finding probabilities of respondents. To assess the accuracy of these subjective beliefs, I confront them with objective probabilities that are predicted based on a rich set of observed individual characteristics. Comparing the biases in expectations between women and men, I find that women are more pessimistic about keeping or finding a job. Further, I find (i) a negative relationship between perceived job loss probabilities and wages among employees and (ii) a positive relationship between perceived job finding probabilities and reservation wages among unemployed job seekers. In the final step, decomposition results show that differences in beliefs about job insecurity significantly contribute to the gender wage gap. Additionally, I provide evidence that the effect of job loss expectations on wages is larger among workers without a collectively agreed wage and among college graduates.

JEL Codes: J16; J31; D84

Keywords: Biased Beliefs, Gender Wage Gap, Job Loss Expectations, Job Finding Expectations

Acknowledgments: I am grateful for valuable remarks from Aderonke Osikominu, Nadja Dwenger, and Steffen Altmann as well as from seminar and conference participants at the University of Hohenheim and at the 15th International German Socio-Economic Panel User Conference 2024.

Status: Not submitted

4.1 Introduction

Although gender equality in the labor market has improved considerably in the last decades, gender differences in earnings persist worldwide (e.g., Blau and Kahn, 2017; Goldin, 2014). In Germany, women’s hourly earnings are on average 18% lower than those of men (Eurostat, 2021). Since standard variables like labor market experience, education, and occupational choices cannot fully explain the gender disparity in wages, recent literature has focused on what other factors account for the remainder of the gender gap. In this context, one strand of the literature studies the role of biased beliefs about ability and labor market prospects. It has been documented that women are on average less confident about their ability and future wages than men (e.g., Briel et al., 2022; Reuben et al., 2017). Moreover, previous work has shown that these differences in beliefs can help explain gender gaps in earnings (Cortés et al., 2023).

Gender differences in beliefs could also occur in relation to workers’ expectations of losing or finding a job, as workers are found to misjudge their risk of losing a job as well as their chances of finding a job (Mueller and Spinnewijn, 2023a). Whereas workers tend to be too optimistic of finding a job, they tend to be too pessimistic of losing their job (Mueller et al., 2021; Balleer et al., 2023a). These misperceptions matter because they can impact individuals’ labor market choices and outcomes. Empirically, previous studies have found a negative association between perceived job insecurity and wages on the one hand and a positive association between perceived job finding chances and reservation wages on the other hand (Campbell et al., 2007; Balleer et al., 2023a). Thus, if men’s and women’s expectations of losing or finding a job diverge, this may affect the gender wage differential.

This paper analyzes gender differences in job loss and job finding expectations and their effect on the gender wage gap using German survey data. To do so, I evaluate elicited beliefs on job insecurity in a sample of employees and elicited beliefs on job finding chances in a sample of unemployed job seekers. In a first step, I compare the subjective beliefs to objective predictions of job loss and job finding probabilities that

are obtained by applying machine learning methods and using a large set of observed individual characteristics as predictors. I define the difference between expectations and predictions as bias measures at the individual level and compare the resulting average values of male and female respondents. In the next step, I estimate the relationships between the biased beliefs and wages and reservation wages, respectively. To this end, I regress the hourly wages of employees on the bias in job loss expectations and various control variables including detailed information on employees' current job. Similarly, I regress the monthly reservation wages of job seekers on the bias in job finding expectations controlling for job search characteristics and other variables. In the last step, I decompose the gender gaps in wages and reservation wages using the well known method of Blinder (1973) and Oaxaca (1973). This allows me to quantify the contribution of differences in job loss and job finding expectations to the gender gaps in wages and reservation wages, respectively. In addition, I investigate whether the effect of the labor market expectations and the corresponding contributions to the gender gaps differ between specific groups. First, I compare the effect of the perceived job insecurity on wages between workers with and without a collectively agreed wage. Next, I study heterogeneity with respect to age and educational attainment. Finally, I explore differences across the wage and reservation wage distribution by estimating unconditional quantile regressions and decompositions (Firpo et al., 2009, 2018).

With this work, I mainly relate to three research streams. First, my study speaks to the growing literature analyzing biases in workers' expectations regarding their labor market prospects (see Mueller and Spinnewijn (2023a) for an extensive review). For instance, Spinnewijn (2015) shows that unemployed job seekers considerably underestimate their remaining unemployment duration. Likewise, the findings of Mueller et al. (2021) and Balleer et al. (2023a) suggest that job seekers overestimate their chances of finding a job. Moreover, Dickerson and Green (2012) and Balleer et al. (2023a) demonstrate that workers' expectations of a potential job loss are too pessimistic. Jäger et al. (2024) document that employees also have biased beliefs about their outside options with other employers.

Particularly, they find that workers employed at low-paying firms tend to underestimate the wages they could earn at other firms.

Second, I contribute to the literature on gender differences in beliefs about one's own abilities and labor market prospects. A number of previous studies in relation to overconfidence finds that men are more confident than women about their absolute and relative abilities (e.g., Adamecz-Völgyi and Shure, 2022; Niederle and Vesterlund, 2007; Reuben et al., 2017). For example, experimental evidence by Reuben et al. (2017) shows that men are twice as likely as women to overestimate their relative abilities. Further, gender differences in beliefs do also appear in direct relation to labor market prospects. Relating their experimental overconfidence measure to wage expectations, Reuben et al. (2017) find that overconfidence can partly explain the gender gap in students' wage expectations. Similarly, Briel et al. (2022) show that female students are less confident than their male counterparts about their own earnings potential relative to other students in their field, which substantially affects the gender gap in wage expectations. Dawson (2017) demonstrate that female job seekers are less optimistic about their future earnings, by comparing expected wages with predicted market wages.

Third, my analysis is related to studies assessing the effect of differences in beliefs on gender gaps in labor market outcomes. Adamecz-Völgyi and Shure (2022) look at the role of overconfidence in explaining the gender difference in the probability of being in a top job in mid-career. They measure overconfidence by comparing individuals' subjective estimate of their ability to an objective measure of cognitive ability in adolescence. According to their estimates the overconfidence measure can explain up to 11 percent of the gender gap in top job employment. Cortés et al. (2023) analyze how biased beliefs affect the job search process and the subsequent early career wages of college graduates. They provide field and experimental evidence that females exhibit a lower level of overconfidence leading to lower reservation earnings compared to males. As a consequence, women accept jobs earlier and earn less than men when entering the labor market. Thus, their findings indicate that biased beliefs contribute to the early career gender wage gap of college graduates.

To the best of my knowledge, this paper is first to assess gender differences in job loss and job finding expectations and the resulting effect on wages and reservation wages. Balleer et al. (2023a) use the same data source to examine the effect of misperceptions of job loss and job finding probabilities on wages, but they do not focus on gender differences. Moreover, I estimate the objective job loss and job finding probabilities with machine learning methods to obtain more accurate predictions. In contrast to the studies of Adamecz-Völgyi and Shure (2022) and Cortés et al. (2023), I do not only focus on a specific career stage or skill group when evaluating the effect of beliefs on gender gaps in labor market outcomes. My analysis includes subjects of various ages, skill levels and occupations. Further, I study the role of biased beliefs among both employed and unemployed individuals. In addition, I contribute to the literature by comparing the effect of biased beliefs on wages and the corresponding gender gap between different groups and across the wage distribution.

My results suggest that women are on average more pessimistic about keeping or finding a job than comparable men. Whereas both men and women overestimate the probability of losing their job, women do so to a larger extent. The probability of finding a job is only overestimated by men, while women's perceptions are almost unbiased on average. Further, I find a negative relationship between perceived job insecurity and wages and a positive relationship between perceived job finding chances and reservation wages. When decomposing the gender gaps in wages, I see that differences in beliefs positively contribute to the gaps in wages and reservation wages. However, the contribution is only significant in case of the job loss bias. Additionally, I show that the perceived job insecurity mainly affects wages of workers without a collectively agreed wage. Lastly, I find that the effect of the job loss expectations on wages and the gender wage gap is particularly pronounced for college graduates.

The remainder of the paper is structured as follows: In Section 4.2, I describe the survey data and relevant variables that are used in my analysis. Section 4.3 provides descriptive statistics of the elicited beliefs, shows how the bias measures are constructed

and compares them between men and women. Section 4.4 presents results of the wage and reservation wage regressions and the corresponding decompositions of the gender gap. In Section 4.5, I report estimation results within specific subgroups and across the wage distribution. Section 4.6 concludes.

4.2 Data

For my empirical analysis, I use data from the German Socio-Economic Panel Study (SOEP), which is a longitudinal, nationally representative household survey. The survey is conducted annually (since 1984) and covers around 15,000 households and 30,000 individuals in Germany. It includes detailed individual information about a wide range of socio-economic variables. Conveniently, survey participants are also asked about individual labor market expectations, which makes the SOEP a well-suited data source for my analysis. In particular, the SOEP includes questions about the perceived job loss and job finding probabilities within two years of employed and unemployed individuals, respectively. Employed individuals in the SOEP are asked: “How likely is it that you will lose your job within the next two years?” Equivalently, unemployed job seekers are asked: “How likely is it that you take a paid job within the next two years?” In both cases respondents are requested to answer on a probability scale from 0 to 100% in steps of 10%.

Due to their panel structure, the SOEP data also allow one to confront these perceived job loss and job finding probabilities with actual outcomes. That is, I can observe whether individuals have actually lost their job or found a new employment within the given time horizon. Survey respondents are regularly asked whether they have left their job since the beginning of the last calendar year, and if so, for what reason and in what month. As reasons for the termination of the employment relationship, the questionnaire offers respondents an extensive list of options including both voluntary and involuntary causes. I classify “My place of work or office closed”, “I was dismissed by employer”,

“Mutual agreement with employer” or “I completed a temporary job” as job losses.¹ While the last two options might not be job losses in the strict sense, they might still be considered by respondents when evaluating their job loss probabilities (Balleer et al., 2023a). Unfortunately, the questionnaire does not include a similar question with respect to job finding. But, the SOEP offers an activity calendar that encompasses various activity spells of individuals over the life-cycle including different employment spells (full-time, part-time, and minor employment as well as unemployment). The data set contains the start and end month of any given activity spell, which allows me to extract information on whether a job seeker has found a job within the next two years or not. As a successful job finding, I classify all transitions from unemployment into some form of employment.²

The job loss and job finding questions were asked bi-annually between 1999 and 2015 as well as in 2018 and 2020.³ However, in 2018 and 2020 the given time horizon overlaps with the COVID-19 pandemic. Therefore, I base my analysis only on the nine waves between 1999 and 2015 that include the job loss and job finding questions. For the information on actual job loss and job finding I use yearly data from 1999 to 2017 and the activity calendar. As the perceived job loss and job finding probabilities are elicited from different sets of respondents (employed/unemployed), I use two different samples. In the first sample, I include all individuals that are employed in some form (full-time employed, part-time employed, marginally employed, in vocational training) at the time of the interview and have stated their perceived job loss probability.⁴ But, I discard all individuals above the age of 65. In the second sample, I include all individuals that are registered as unemployed at the time of the interview, would like to take up a job at least within the coming year, and have stated their perceived job finding probability. As in the first sample, I restrict the sample to individuals aged 65 or younger. In both samples I

¹Balleer et al. (2023a) use the same definition of a job loss as well as a more narrow and a more wide definition. Their results are robust the use of different job loss definitions.

²I follow Balleer et al. (2023a) and consider the following activities as employment: full-time, part-time and marginal employment, short-time work, second job, minijob, vocational training, and apprenticeship.

³In 2011, only the job finding question was asked.

⁴The sample also includes civil servants and employees in the public sector who enjoy a high degree of job security in Germany. But, I include indicators for being a civil servant and working in the public sector as predictors for losing a job and as control variables in my main analysis.

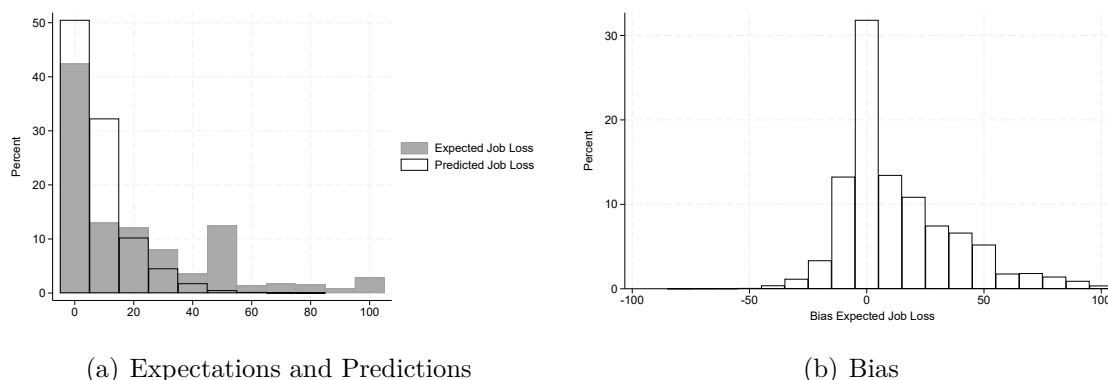
pool observations across the different survey years. Whereas the ‘employed’ sample consist of over 85,000 observations stemming from around 31,600 individuals, the ‘unemployed’ sample only consists of over 6,700 observations stemming from around 4,400 individuals. Around 48% of all observations in both samples are female.

In addition to individual labor market expectations, the SOEP includes information on wages and reservation wages that I use as dependent variables in my main analysis. Employed respondents are asked about their monthly net labor income as well as their weekly working hours. I combine both variables to compute hourly wages of employees. Similarly, unemployed respondents are asked to state their monthly net reservation wage and their desired weekly working hours. But, as the information on the desired working time is not available for the entire time period (only available since 2007), I focus on monthly rather than hourly reservation wages.⁵

Furthermore, the SOEP provides extensive individual information on various topics such as education, employment, job search, and personality, that are very useful for my analysis. On the one hand, I can use these variables as predictors of the job loss and job finding probabilities. On the other hand, I can use them as control variables when analyzing the effect of biased labor market expectations on wages and reservation wages. First and foremost, I consider demographic variables like age, marital status, children, migration background, household income, and health. Next, I extract information on educational attainment and labor market history. For employees, I also include various characteristics of their current employment. For instance, I consider their tenure, employer size, type of contract (limited or unlimited), and whether they work full- or part-time. In addition, I use information on the occupation and industry they work in based on the ISCO88 (1-digit) and NACE (Level 1) classifications, respectively. For unemployed job seekers, the SOEP provides information such as whether they have actively searched for a job in the last four weeks, whether they are seeking a full- or part-time employment, and whether

⁵Since I pool observations across survey years, I deflate wages and reservation wages using the German Consumer Price Index with base year 2015. I obtain the monthly Consumer Price Index from the German Federal Statistics Office (<https://www-genesis.destatis.de/genesis/online>, accessed December, 2022).

Figure 4.1: Distribution of Expected and Predicted Job Loss Probabilities



Notes: (a) displays histogram of expected and predicted job loss probabilities. (b) displays histogram of bias in job loss expectations (difference between expectations and predictions). Predictions are rounded to the closest decile on the probability scale. *Source:* German Socio-Economic Panel v37, own calculations.

they receive unemployment benefits. Furthermore, I use the activity calendar to extract information on the duration of job seekers' current unemployment spell. Moreover, the SOEP questionnaire contains questions referring to numerous personality traits. I include the Big Five Personality Traits (openness, conscientiousness, extraversion, agreeableness, and neuroticism), external and internal locus of control as well as positive and negative reciprocity in my analysis. Likewise, respondents are asked about their willingness to take risks. Lastly, I use information on which federal state respondents live in.

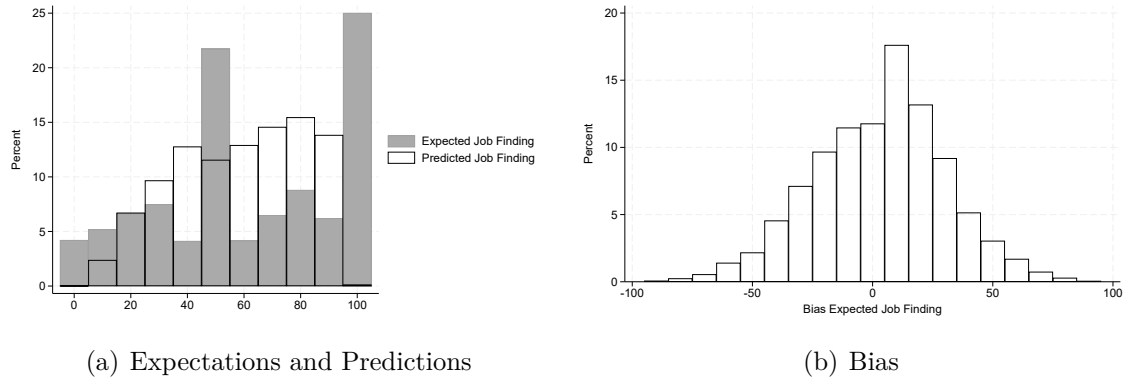
4.3 Bias in Labor Market Expectations

4.3.1 Job Loss and Job Finding Expectations

As described above, I observe the subjective perceptions of respondents about their job loss and job finding chances. Subfigure (a) of Figures 4.1 and 4.2 display the distribution of expected job loss and job finding probabilities.⁶ The respective mean values and standard deviations in the full sample and by gender are shown in Table 4.1. Overall, the distribution of expected job loss probabilities is quite dispersed. The stated probabilities cover all possible choices between 0 and 100%. Yet, 0% is by far the most frequent answer.

⁶Figures 4.A.1 to 4.A.4 in Appendix A display distributions by gender.

Figure 4.2: Distribution of Expected and Predicted Job Finding Probabilities



Notes: (a) displays histogram of expected and predicted job finding probabilities. (b) displays histogram of bias in job finding expectations (difference between expectations and predictions). Predictions are rounded to the closest decile on the probability scale. *Source:* German Socio-Economic Panel v37, own calculations.

Over 40 percent of all employees believe that there is no chance that they will lose their job within two years (or at least think the probability is closer to 0 than 10%). There also seems to be some bunching at 50%. On average, employees in my sample expect to lose their job within two years with a probability of 21.6%. The average expected job loss probability of women (22%) is similar to, but slightly higher than, that of men (21.2%). The gender difference is statistically significant at the 5%-level.

The stated job finding probabilities are more evenly distributed than the job loss probabilities and cover again all possible values. However, there is some bunching at 50% and 100%. A quarter of all respondents is certain that they will find a job within the next two years (or at least thinks the probability is closer to 100 than 90%). Job seekers on average state a job finding probability of 60.6%. Males (63.5%) are on average more optimistic to find a job than females (56.9%), with the difference being highly statistically significant.

4.3.2 Predicting Job Loss and Job Finding Probabilities

As described in Section 4.2, I observe whether employed and unemployed respondents have actually lost or found a job, respectively. However, I cannot directly compare the

Table 4.1: Summary Statistics of Job Loss and Job Finding Probabilities

	(1)	(2)	(3)	(4)
	Total	Females	Males	Difference
(a) Job Loss Probability				
Expectations	21.56 (26.02)	22.02 (26.66)	21.18 (25.47)	0.84**
Predictions	7.85 (10.21)	7.75 (9.80)	7.94 (10.54)	-0.18
Bias	13.71 (24.64)	14.27 (25.14)	13.24 (24.21)	1.02***
Observations	85,157	40,710	44,447	
(b) Job Finding Probability				
Expectations	60.63 (31.58)	56.93 (31.09)	63.48 (31.65)	-6.56***
Predictions	58.00 (22.83)	56.35 (22.20)	59.27 (23.23)	-2.92***
Bias	2.63 (27.58)	0.58 (28.21)	4.21 (26.98)	-3.64***
Observations	6,728	3,240	3,488	

Notes: Columns (1) - (3) display mean values of respective variable as well as respective standard deviations in parentheses. Column (4) displays difference in means between female and male subsample. *, **, and *** indicate statistical significance of the difference at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

actual outcomes to the expectations at the individual level, since the expectations are stated as probabilities. By comparing average expectations with average outcomes, I can only identify a bias at the aggregate level. Hence, I follow Balleer et al. (2023a) and compare the perceived job loss and job finding probabilities to predicted probabilities. In doing so, I can confront the perceived probabilities with objective measures that are based on the extensive individual information that the SOEP provides.

To predict the occurrence of job loss and job finding as accurately as possible, I apply machine learning methods and use a large set of predictors. As predictors of job loss, I include, first of all, demographic characteristics like age, migration background, marital status, and whether the respondents lives in a household with a child (under the age of 16).⁷ Next, I use information on respondents' school leaving degree and whether they have completed a vocational or college degree as well as their working (full- and part-time) and unemployment experience. Furthermore, I incorporate information on respondents' current job, such as the tenure at their current job, employer size, working time (full- or

⁷The full list of predictors is shown in Table 4.C.1 in Appendix C.

part-time), whether the working contract is limited, whether they working in their learned occupation, and whether they are employed in the public sector and are civil servants. I also use occupation and industry fixed effects. Finally, I include federal state and year fixed effects. For the prediction of the job finding probabilities, I use the same variables except for the job characteristics and occupation and industry fixed effects. Instead, I additionally include information on subjective health, whether individuals have actively searched for a job in the last four weeks, whether they are looking for a full- or part-time employment. Moreover, I include information on the duration of individuals' current unemployment spell and whether they receive unemployment benefits.

Predicting the occurrence of job loss and job finding are binary classification tasks. I use a combination of various machine learning methods for this task, since it is *ex ante* difficult to predict which method performs best in any given application.⁸ Moreover, averaging over multiple methods may outperform the use of the best-performing individual method (Ahrens et al., 2023). To combine the different machine learning methods into one final prediction I use *stacking*, which was first proposed by Wolpert (1992) and Breiman (1996). Stacking is implemented in two steps.⁹ In the first step, I compute multiple cross-validated predicted probabilities for each observation based on different machine learning methods (base learners).¹⁰ The cross-validation ensures that more flexible base learners that overfit are not weighted too heavily by the stacking algorithm. In the second step, the different predictions are combined into one final predicted probability (final learner). To do so, the observed outcome (dummy for job loss or job finding within two years) of individuals is regressed on the different cross-validated predicted probabilities (without intercept) using nonnegative least squares with the restriction that the coefficients (weights) sum to one. The final prediction is then obtained by re-fitting the base learners in the full sample (i.e. without cross-validation) and plugging the corresponding predicted probabilities into the estimated regression function (Ahrens et al., 2023; Hastie

⁸Balleer et al. (2023a) use a simple probit model to estimate job loss and job finding probabilities.

⁹I implement stacking in Stata with the *pystacked* package of Ahrens et al. (2023), which allows one to execute multiple machine learning algorithms via Python's *scikit-learn* of Pedregosa et al. (2011).

¹⁰I apply 5-fold cross-validation.

et al., 2009).¹¹ Hence, stacking can be interpreted as a weighted average of base learners, with the weights being chosen such that they maximize out-of-sample predictive accuracy. It has been shown by van der Laan et al. (2007) that stacking performs asymptotically at least as well as the best-performing individual learner (when the set of base learners is not too large).

As base learners for the stacking, I choose logistic regression, the two tree based methods random forest and gradient boosting, as well as a neural net.¹² All of these base learners are suited for classification tasks and are also probabilistic classifiers, i.e. they are not only able to predict the most likely class but also provide predicted probabilities, which is required for my analysis. In case of the logit regression, I enlarge the set of predictors by including flexible functions of continuous variables. For instance, I use second order polynomials of age and the labor market experience variables. Moreover, I include flexible functions of tenure and the duration of the current unemployment spell for the prediction of job loss and job finding probabilities, respectively.¹³ In addition, I use industry $\times$ year fixed effects when predicting the probability of a job loss to account more flexibly for potential business cycle effects. For the estimation with random forest, I set the number of individual trees equal to 400. The gradient boosting estimates are based on 100 boosting stages and a learning rate of 0.1. Lastly, in case of neural net, I use a network with one hidden layer consisting of 50 neurons and choose the logistic function as activation function.

After predicting the job loss and job finding probabilities, I round the predictions

¹¹The predicted probabilities that are estimated in the first step and are used to estimate the regression function (the weights for the different base learners) are cross-validated. Hence, no observation is used to fit the learner for the prediction of its own probability at this stage. However, the predicted probabilities that are plugged into the estimated regression function in the final step are not cross-validated, i.e. they are fitted in the entire sample.

¹²For details on the implementation of the machine learning algorithms, see Pedregosa et al. (2011).

¹³Previous research (e.g., Mueller and Spinnewijn, 2023b) shows that the duration of the unemployment spell nonlinearly affects job seekers' chances of finding a job. I include dummy variables for an unemployment duration of 0 - 2 months, 2 - 6 months, 6 - 12 months, 12 - 24 months, 24 - 36 months, and more than 36 months. Since for some observations unemployment spells are left-censored, I use having a left-censored unemployment spell as baseline category. Similarly, the probability to lose a job may depend on tenure in a complex way. Thus, I include dummy variables for a tenure of 1 - 2 years, 2 - 3 years, 3 - 5 years, 5 - 10 years, 10 - 20 years, and more than 20 years (less than 1 year is baseline category).

to the closest decile on the probability scale to make them comparable to the perceived probabilities.¹⁴ The distribution of the (rounded) predicted probabilities is displayed by the white bars in Subfigure (a) of Figures 4.1 and 4.2, respectively. Summary statistics in the full sample and by gender are shown in Table 4.1.¹⁵ For the majority of all employees, the probability of losing their job in the next two years is rather low. Over 80 percent of all respondents are predicted to lose their job with 0 or 10%. On average, employees face a job loss probability of 7.9%. The average predicted job loss probability is slightly higher for men (7.9%) than for women (7.8%). But, the difference is statistically insignificant. With respect to job finding, the predictions are more evenly distributed across the probability scale. Yet, almost none of the job seekers is predicted to find a job with 0 or 100%. The average predicted job finding probability is 58%. For females the average probability is significantly lower than for males. The gender difference amounts to around 3%-points.

4.3.3 Bias in Expected Probabilities

Having obtained predicted probabilities, I can compare them with the perceived ones and compute bias measures at the individual level. As bias measures, I simply take the difference between the perceived and predicted probabilities of respondents. Subfigures (b) of Figures 4.1 and 4.2 depict the distribution of the respective bias measures, while summary statistics are shown in Table 4.1.¹⁶

In terms of a possible job loss, respondents are overly pessimistic. Employees in my sample overestimate their job loss probability on average by 13.7%-points. This can also be seen in Figure 4.1, which shows that the distribution of the job loss bias is clearly right skewed. While both women and men tend to be too pessimistic with respect to losing their job, women overestimate the probability to a larger extent. The gender difference

¹⁴My findings are robust to the use of exact rather than rounded predictions. Estimation results without rounding the predicted probabilities are shown in Appendix B.

¹⁵The stacking weights for the predictions are shown in Table 4.A.1 in Appendix A. The distribution of exact predicted probabilities based on the different base learners and the stacking algorithm are displayed in Figures 4.A.9 and 4.A.10 in Appendix A.

¹⁶Figures 4.A.7 and 4.A.8 in Appendix A show mean values of the bias measures within various subgroups.

of 1%-points is highly statistically significant. Unemployed job seekers, in contrast, are on average too optimistic of finding a job. However, the bias is far less pronounced in this case. The mean value of the job finding bias is only 2.6%-points. Furthermore, the distribution is almost symmetrically, even though it peaks at 10 rather than 0%. With respect to the gender difference, I see again that men are significantly more optimistic than women. Whereas women are nearly unbiased in their job finding expectations on average, men overestimate their job finding chances by 4.2%-points.

4.3.4 Predictive Power of Expectations

The observed misjudgments of labor market prospects, particularly in case of the perceived job loss probability, raise the question whether the elicited beliefs are nevertheless meaningful and possess some predictive power for actual outcomes. Given that there is no monetary incentive involved, one might be concerned that respondents do not give much thought to their stated expectations. However, a number of previous studies have documented that elicited beliefs about individual labor market prospects are strongly correlated with realizations (e.g., Hendren, 2017; Mueller et al., 2021; Dickerson and Green, 2012).

To see whether the stated expectations are predictive for realizations in my sample, I first regress the indicators for an actual job loss or job finding on the respective expected probabilities. The corresponding results are shown in Table 4.2. As can be seen in Columns (1) and (3), the elicited beliefs are highly correlated with actual outcomes. The correlation, though, is substantially higher between the perceived job finding chances and actual job finding. In Columns (2) and (4), I condition on the predicted probabilities that are based on observed characteristics. But, the coefficients of the expected probabilities remain positive and highly significant in both cases. Hence, the elicited beliefs do possess predictive power for realizations, even conditional on the observed characteristics.

Furthermore, I analyze the plausibility of the elicited beliefs by comparing the average expectations with the average realizations across several subgroups. Figures 4.A.5

Table 4.2: Predictive Power of Labor Market Expectations

	Job Loss		Job Finding	
	(1)	(2)	(3)	(4)
Expectations	0.2515*** (0.0073)	0.0751*** (0.0065)	0.5669*** (0.0259)	0.1627*** (0.0275)
Predictions		1.3963*** (0.0189)		1.0447*** (0.0350)
Observations	87,939	81,768	6,060	6,025
R^2	0.054	0.283	0.132	0.303

Notes: Table displays linear probability model estimates. Dependent variable in Columns (1) and (2) is an indicator for a job loss within two years. Dependent variable in Columns (3) and (4) is an indicator for a job finding within two years. Standard errors are clustered at the individual level and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

and 4.A.6 in Appendix A plot the average realizations, predictions, and expectations across subgroups based on characteristics like age, education or tenure. Comparing the average perceived job loss probabilities with realizations, it stands out that the average perceptions exceed the realizations in almost all subgroups. Thus, the upward bias in the perceived job loss probabilities is universally observed in all subgroups of respondents. Yet, with regard to most characteristics, the ranking between the different categories is identical for the expectations and realizations. For instance, just as the actual job loss probability, the average job loss expectation is highest among low-skilled employees and lowest among high-skilled employees. With respect to the job finding expectations, the upward bias is not observed in all considered subgroups. But, in most cases, the ranking of average expectations between categories again matches with the corresponding ranking of realizations. This demonstrates that, even though the elicited beliefs are biased on average, they are still reasonable to a certain extent and are not completely irrational.

4.3.5 Gender Differences

The comparison of the average bias measures suggests that women are less optimistic than men about keeping their job or finding a new one. However, the observed gender

Table 4.3: Gender Difference in Job Loss Expectations

Dependent Variable: Job Loss Bias						
	(1)	(2)	(3)	(4)	(5)	(6)
Female	1.0226*** (0.3784)	0.9923*** (0.3596)	0.8859** (0.3880)	1.6209*** (0.3880)	1.4651*** (0.3863)	1.8915*** (0.5411)
Predicted Prob.		-0.1656*** (0.0194)	-0.1099*** (0.0219)	-0.3588*** (0.0277)	-0.3754*** (0.0278)	-0.3687*** (0.0431)
Demographics			✓	✓	✓	✓
Education			✓	✓	✓	✓
Experience			✓	✓	✓	✓
Job Controls				✓	✓	✓
Federal State FE					✓	✓
Year FE					✓	✓
Personality						✓
Observations	85,157	85,157	85,157	82,539	82,539	43,025
R^2	0.000	0.005	0.019	0.107	0.120	0.138

Notes: Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

differences may also be related to other characteristics that are correlated with gender and the individual bias in labor market expectations. To see whether the gender differences persist or change conditional on observed characteristics, I run linear regressions. I regress the respective bias on a female dummy, the predicted probabilities, and on other control variables.¹⁷ Controlling for the predicted job loss and job finding probabilities, respectively, is important, as the possible values of the bias measure depend on the predictions by construction. For instance, it is not possible for employees with predicted job loss probabilities of 0% to underestimate the chances of losing their job. Further, the predicted probabilities differ significantly between women and men.

Results with respect to the job loss bias are displayed in Table 4.3. Controlling for the predicted job loss probabilities does not change the gender gap in the job loss bias significantly, as can be seen in Column (2). Moreover, Columns (3) - (5) show that the gender difference does not disappear when I gradually add controls for demographic character-

¹⁷The full list of control variables is shown in Table 4.C.1 in Appendix C.

Table 4.4: Gender Difference in Job Finding Expectations

Dependent Variable: Job Finding Bias						
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-3.6362** (1.5058)	-4.4524*** (1.2033)	-4.4434*** (1.0723)	-4.5096*** (1.1494)	-3.2321*** (1.1231)	-3.4855** (1.6117)
Predicted Prob.		-0.2794*** (0.0241)	-0.5037*** (0.0391)	-0.4108*** (0.0579)	-0.4299*** (0.0600)	-0.5042*** (0.0862)
Demographics			✓	✓	✓	✓
Education			✓	✓	✓	✓
Experience			✓	✓	✓	✓
Job Search				✓	✓	✓
Unemployment				✓	✓	✓
Federal State FE					✓	✓
Year FE					✓	✓
Personality						✓
Observations	6,728	6,728	6,484	6,484	6,484	3,091
R^2	0.004	0.058	0.130	0.151	0.185	0.230

Notes: Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

istics, education, labor market experience, job characteristics, federal state, and survey year. Quite in the contrary, the gender difference does even get larger. In Column (6), I additionally control for different personality traits (Big Five, locus of control, and reciprocity) and risk tolerance. Since the personality traits are not available in all survey years the sample is reduced to halve its size in this case. Based on this regression, the average job loss bias is 1.9%-points higher for women, *ceteris paribus*. Thus, the estimated gender difference is almost twice as high as the unconditional gender gap (displayed in Column (1)). The result also indicates that the gender difference in beliefs is not due to differences in personality or any other observed characteristics.

Table 4.4 displays the corresponding results for the job finding bias. Column (2) shows that the gender gap in the job finding bias increases to 4.5%-points when conditioning on the predicted job finding probabilities. Controlling for demographic characteristics, education, labor market history, job search characteristics, and unemployment duration

does not affect the size of the estimated gender difference substantially. Adding controls for the federal state and the survey year, however, does reduce the gap just below the unconditional difference (displayed in Column (1)). Yet, the inclusion of personality traits and risk tolerance in Column (6) increases the difference again. According to this estimate, which is again based on a reduced sample, the conditional gender gap in the job finding bias amounts to 3.5%-points. Hence, the differential between women and men can once more not be explained by differences in observed characteristics.

Overall, I see that women tend to be more pessimistic about their labor market prospects, i.e. they perceive a job loss as more and a job finding as less likely than comparable men on average. This finding is in line with previous literature showing that women are less optimistic about their future wages (Briel et al., 2022; Cortés et al., 2023). Furthermore, my results suggest that these gender differences in beliefs are not driven by differences in observed personality traits or risk preferences.

4.4 Labor Market Expectations and Wages

Having seen that women are less confident about keeping or finding a job than men, I next analyze whether the biased beliefs affect wages and the associated gender gaps. Higher perceived job insecurity may reduce employees' bargaining power in wage negotiations, and thus, lead to lower wages, as is argued by Campbell et al. (2007). Similarly, lower perceived job finding chances may reduce the reservation wages of job seekers. In two recent studies Balleer et al. (2023b) and Balleer et al. (2023a) examine the implications of biased beliefs for wage bargaining in a formal way. They formulate a search-and-matching model of the labor market, where wages are determined by generalized Nash bargaining between workers and firms. According to their model, an increase in the perceived job loss probability of workers reduces their wages. Intuitively, this is because pessimistic workers accept lower wages as they heavily discount the effect of future wages on the job value. For a decrease in the perceived job finding probabilities of job seekers, the model predicts

a decrease of reservation wages. The reason being that optimistic workers overestimate the value of their outside option when bargaining with the firm, and hence, they demand higher wages. If the theoretical predictions are correct, the observed gender differences in beliefs may affect the gender gaps in wages and reservation wages, which I will explore in the following.

4.4.1 Relating Bias in Labor Market Expectations to Wages

In the first step, I relate the subjective beliefs about job loss and job finding to wages and reservation wages, respectively, to see whether the empirical relationships are in line with the theoretical predictions in my sample. To examine the respective relationships I estimate the wage equations,

$$\ln(y)_i = \alpha + \gamma bias_i + \rho pred_i + \delta df_i + \mathbf{x}_i\boldsymbol{\beta} + u_i, \quad (4.1)$$

where the dependent variable, $\ln(y)_i$, is the logarithm of the hourly wage or of the monthly reservation wages of individual i . $bias_i$ and $pred_i$ denote the bias and the predicted probability with regard to job loss or job finding. df_i is a female dummy variable and $\mathbf{x}_i$ is the vector of control variables. u_i represents the error term. The theoretical predictions would suggest a negative coefficient, γ , in case of the job loss bias and a positive γ in case of the job finding bias.

Relating Job Loss Bias to Wages

For the estimation of the relationship between the bias in job loss expectations and hourly wages, I include a similar set of variables in $\mathbf{x}_i$, as was used as predictors for the occurrence of a job loss. Hence, I control for a wide range of variables about demographics, education, labor market experience and job characteristics.¹⁸ To allow for nonlinear effects of continuous variables, I again use second order polynomials of age and labor market ex-

¹⁸The full list of control variables is shown in Table 4.C.1 in Appendix C.

Table 4.5: Bias in Job Loss Expectation and Wages

Dependent Variable: Log Hourly Wage					
	(1)	(2)	(3)	(4)	(5)
Job Loss Bias	-0.0176*** (0.0033)	-0.0221*** (0.0027)	-0.0169*** (0.0025)	-0.0098*** (0.0024)	-0.0073** (0.0035)
Predicted Prob.	-0.0235*** (0.0005)	-0.0129*** (0.0004)	-0.0085*** (0.0004)	-0.0070*** (0.0004)	-0.0076*** (0.0006)
Female		-0.1496*** (0.0067)	-0.1116*** (0.0070)	-0.1027*** (0.0068)	-0.0747*** (0.0101)
Demographics		✓	✓	✓	✓
Education		✓	✓	✓	✓
Experience		✓	✓	✓	✓
Job Characteristics			✓	✓	✓
Federal State FE				✓	✓
Industry $\times$ Year FE				✓	✓
Personality					✓
Observations	77,981	77,981	77,981	77,981	40,344
R^2	0.198	0.400	0.478	0.506	0.507

Notes: Job Loss Bias is standardized. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

perience as well as a flexible function of tenure. Furthermore, I also include federal state and industry $\times$ year fixed effects. In some specifications I additionally incorporate the observed personality traits and risk preferences.

Estimation results are shown in Table 4.5.¹⁹ The estimated coefficients in Column (1) are based on a regression where I only include the job loss bias and the predicted job loss probability as explanatory variables. The coefficient of the job loss bias is negative, highly statistically significant, and suggests that a one standard deviation increase in the job loss bias is associated with 1.8 percent (log approximations) lower wages on average. In Column (2), I add a female dummy and control variables for demographics, education and labor market experience, which leads to a slightly larger negative effect of the job loss bias. Next, in Column (3), I additionally include job characteristics,

¹⁹Table 4.A.2 in Appendix A shows estimation results by gender and for full-time employees only.

which reduces the size of the effect again. Adding federal state and industry $\times$ year fixed effects in Column (4) further reduces the size of the effect. However, the estimated coefficient remains sizable and highly statistically significant. According to the estimate in Column (4), a one standard deviation increase in the job loss bias lowers wages on average, *ceteris paribus*, by 1 percent. Lastly, in Column (5), I add the observed personality traits and risk preferences to the regression model, which reduces the sample size substantially. Yet, the effect of the job loss bias remains close to the estimate in Column (4) and is statistically significant at the 5%-level. Hence, I find a negative relationship between the expected job loss probability and wages, that does also hold conditional on various observed individual characteristics including character skills. This finding is in line with the theoretical predictions. Moreover, my results confirm findings of previous literature that has looked at the relationship between perceived job insecurity and wages empirically (Campbell et al., 2007; Balleer et al., 2023a).

Relating Job Finding Bias to Reservation Wages

In case of the reservation wage equation, I include a similar set of variables in the vector of control variables, $\mathbf{x}_i$, as I have used to predict the job finding probabilities.²⁰ That is, I control for various information on demographics, education, labor market experience, job search characteristics, and the current unemployment spell. Once more, I include second order polynomials of age and labor market experience as well as a flexible function of unemployment duration. Additionally, I again control for personality traits and risk preferences in some specifications.

Table 4.6 displays my results of the reservation wage regressions.²¹ When conditioning only on the predicted job finding probability, the estimated coefficient of the job loss bias is positive and suggests that a one standard deviation increase in the job finding bias is associated with a 1.8 percent (log approximations) increase in reservation wages on average,

²⁰The full list of control variables is shown in Table 4.C.1 in Appendix C.

²¹Table 4.A.3 in Appendix A shows estimation results by gender and for individuals looking for a full-time job only.

Table 4.6: Bias in Job Finding Expectation and Reservation Wages

Dependent Variable: Log Monthly Reservation Wage					
	(1)	(2)	(3)	(4)	(5)
Job Finding Bias	0.0177 (0.0110)	0.0152 (0.0099)	0.0199** (0.0094)	0.0098 (0.0094)	0.0050 (0.0147)
Predicted Job Finding	0.0030*** (0.0005)	0.0037*** (0.0008)	-0.0006 (0.0010)	-0.0008 (0.0010)	-0.0016 (0.0016)
Female		-0.2181*** (0.0215)	-0.1522*** (0.0201)	-0.1399*** (0.0197)	-0.1300*** (0.0275)
Demographics		✓	✓	✓	✓
Education		✓	✓	✓	✓
Experience		✓	✓	✓	✓
Job Search			✓	✓	✓
Unemployment			✓	✓	✓
Federal State FE				✓	✓
Year FE				✓	✓
Personality					✓
Observations	4,401	4,269	4,269	4,269	2,055
R^2	0.025	0.179	0.258	0.283	0.304

Notes: Job Finding Bias is standardized. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

as shown in Column (1). However, the coefficient is not statistically significant. But this is probably due to the much smaller sample compared to the wage regressions. In terms of size, the coefficient is almost identical to the coefficient in Column (1) of Table 4.5. In Column (2) I add the female dummy and controls for demographics, education, and labor market experience. Yet, this does not affect the size of the coefficient of the job finding bias substantially. The addition of the job search characteristics and information on the current unemployment spell in Column (3) slightly increases the effect and it becomes statistically significant at the 5%-level. However, adding the federal state and survey year fixed effects in Column (4) decreases the effect somewhat. Based on this estimate, a one standard deviation increase in the job finding bias increases reservation wages on average by 1 percent, *ceteris paribus*. But, the estimated coefficient is not statistically significant in this case. In Column (5), I again incorporate the personality traits and risk preferences, which shrinks the sample size by half. In this case, the estimated effect of the job loss bias remains positive but is somewhat smaller. Although the estimated coefficients are not statistically significant in Columns (4) and (5), the effect of the perceived job finding chances on reservation wages conditional on observed characteristics is comparable in size to the corresponding effect in case of the perceived job insecurity. In any case, the sign of the effect matches with the predictions of the theoretical model. In all columns, the estimates indicate a positive relationship between the perceived job finding probability and reservation wages.

4.4.2 Contribution of Labor Market Expectations to Gender Gaps in Wages

As the perceived job insecurity and job finding chances seem to affect wages and reservation wages, respectively, I analyze whether the differences in beliefs contribute to the corresponding gender gaps. To do so, I apply the well-known decomposition method proposed by Blinder (1973) and Oaxaca (1973), which allows one to break down the difference in means between two groups in an explained and unexplained part. In a first

step, I estimate the gender specific wage equations,

$$\ln(y)_i = \alpha_g + \gamma_g \text{bias}_i + \rho_g \text{pred}_i + \mathbf{x}_i \boldsymbol{\beta}_g + u_i, \quad \text{for } g = F, M. \quad (4.2)$$

In a second step, I can then decompose the gender gap in wages as,

$$\begin{aligned} \hat{\Delta}_O &= \overline{\ln(y)}_F - \overline{\ln(y)}_M \\ &= \underbrace{(\hat{\alpha}_F - \hat{\alpha}_M) + (\hat{\gamma}_F - \hat{\gamma}_M) \overline{\text{bias}}_F + (\hat{\rho}_F - \hat{\rho}_M) \overline{\text{pred}}_F + \bar{\mathbf{x}}_F (\hat{\boldsymbol{\beta}}_F - \hat{\boldsymbol{\beta}}_M)}_{\hat{\Delta}_U} \\ &\quad + \underbrace{\hat{\gamma}_M (\overline{\text{bias}}_F - \overline{\text{bias}}_M) + \hat{\rho}_M (\overline{\text{pred}}_F - \overline{\text{pred}}_M) + (\bar{\mathbf{x}}_F - \bar{\mathbf{x}}_M) \hat{\boldsymbol{\beta}}_M}_{\hat{\Delta}_E}, \end{aligned} \quad (4.3)$$

where $\overline{\ln(y)}_g$, $\overline{\text{bias}}_g$, $\overline{\text{pred}}_g$, and $\bar{\mathbf{x}}_g$ denote the gender specific sample means. The first term, $\hat{\Delta}_U$, is the unexplained part of the gender gap capturing gender differences in the coefficients and intercept. The second term, $\hat{\Delta}_E$, is the explained part capturing gender differences in observed characteristics. In addition to the aggregate decomposition, both parts can be further broken down into the contributions of single regressors or groups of regressors. This allows me to quantify the contribution of gender differences in beliefs about job loss and job finding to the gender gap in wages.

Contribution of Job Loss Bias to Gender Gap in Wages

The decomposition results for the gender gap in hourly wages are shown in Table 4.7. The estimates in the first two columns are based on the full sample of employees, while the estimates in the last two columns are based only on full-time employees. In Columns (2) and (4), I additionally include the personality traits as explanatory variables, which reduces the respective sample size.

In the full sample, the hourly wages of women are on average 21.5 percent (log approximation) lower than those of men, as can be seen in Panel (a). 16.7 percentage points (77.7 percent) of this wage differential can be explained by gender differences in observed

Table 4.7: Wage Decomposition

	Full Sample		Full-Time Only	
	(1)	(2)	(3)	(4)
Aggregate Decomposition				
Gender Gap	0.2147*** (0.0078)	0.2011*** (0.0106)	0.2106*** (0.0080)	0.1963*** (0.0107)
Explained Part	0.1668*** (0.0149)	0.1920*** (0.0203)	0.0817*** (0.0078)	0.0980*** (0.0117)
Unexplained Part	0.0479*** (0.0154)	0.0091 (0.0208)	0.1289*** (0.0075)	0.0983*** (0.0111)
Contribution To Explained Part				
Job Loss Bias	0.0007** (0.0003)	0.0006 (0.0004)	0.0006 (0.0005)	0.0001 (0.0004)
Predicted Prob.	-0.0014 (0.0011)	-0.0017 (0.0015)	-0.0004 (0.0007)	-0.0006 (0.0008)
Demographics	0.0081*** (0.0025)	0.0048 (0.0035)	0.0401*** (0.0033)	0.0383*** (0.0045)
Education	0.0007 (0.0019)	0.0014 (0.0023)	-0.0113*** (0.0025)	-0.0104*** (0.0031)
Experience	-0.0305** (0.0132)	-0.0319* (0.0181)	0.0239*** (0.0058)	0.0359*** (0.0088)
Job Characteristics	0.1632*** (0.0122)	0.2108*** (0.0168)	0.0112 (0.0078)	0.0148 (0.0090)
Federal state	0.0009 (0.0018)	-0.0008 (0.0022)	0.0105*** (0.0022)	0.0072*** (0.0027)
Industry $\times$ Year	0.0252*** (0.0082)	-0.0090 (0.0076)	0.0071 (0.0063)	-0.0013 (0.0059)
Personality		0.0178*** (0.0037)		0.0141*** (0.0029)
Observations	77,981	40,344	54,301	27,662

Notes: The dependent variable is the log hourly wage. The male coefficient vector is used as reference vector. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

characteristics. 4.8 percentage points (22.3 percent) remain unexplained. Panel (b) of the table shows that differences in beliefs about job insecurity do positively and significantly contribute to the gender wage gap. The contribution of the job loss bias measure to the explained part is around 0.1 percentage points, which amounts to 0.4 percent of the explained part and 0.3 percent of the raw gender gap. Thus, although the contribution is statistically significant at the 5%-level, it is rather small in comparison to the contribution of other factors. The decomposition result looks similar when I additionally include the personality traits and risk preferences as covariates. The contribution of the job loss bias remains positive and similar in size. Yet, the contribution is now no longer statistically significant. But, this is probably mostly due to the reduced sample size.

When focusing on full-time employed individuals only, the aggregate decomposition result changes. While the raw gender gap is almost identical to the gap in the full-sample, the explained part is substantially smaller. Only 8.2 percentage points (38.8 percent) of the gender wage gap among full-time employees can be explained by differences in observed characteristics. The contribution of the job loss bias, though, remains similar in absolute terms. Hence, it makes up a larger relative share of the explained part in this case. 0.7 percent of the explained part are now attributed to differences in beliefs about job insecurity. After adding the personality traits and risk preferences, however, the contribution becomes substantially smaller.

Contribution of Job Finding Bias to Gender Gap in Reservation Wages

Table 4.8 displays the results of decomposing the gender reservation wage differential. The first two columns refer to the entire sample of unemployed job seekers. In the last two columns, I only consider individuals that are looking for a full-time employment.

The raw gender gap in monthly reservation wages amounts to 26.2 percent (log approximation) when considering all job seekers. The fraction of this gender gap that can be explained by observed characteristics is rather small. Based on the aggregate decomposition, 10 percentage points (38.3 percent) of the gap can be explained, but 16.2 percentage

Table 4.8: Reservation Wage Decomposition

	Full Sample		Looking for Full-Time	
	(1)	(2)	(3)	(4)
Aggregate Decomposition				
Gender Gap	0.2621*** (0.0193)	0.2563*** (0.0296)	0.1804*** (0.0194)	0.1745*** (0.0289)
Explained Part	0.1004*** (0.0230)	0.1036*** (0.0340)	0.0318** (0.0150)	0.0144 (0.0303)
Unexplained Part	0.1617*** (0.0293)	0.1528*** (0.0408)	0.1486*** (0.0212)	0.1601*** (0.0373)
Contribution To Explained Part				
Job Finding Bias	0.0018 (0.0021)	0.0016 (0.0045)	0.0041 (0.0036)	0.0034 (0.0073)
Predicted Prob.	0.0003 (0.0045)	-0.0033 (0.0084)	0.0019 (0.0031)	-0.0008 (0.0072)
Demographics	-0.0218*** (0.0078)	-0.0256* (0.0147)	-0.0106 (0.0071)	-0.0153 (0.0112)
Education	0.0000 (0.0044)	-0.0117 (0.0095)	0.0022 (0.0048)	-0.0088 (0.0091)
Experience	0.0306* (0.0162)	0.0481* (0.0257)	0.0104 (0.0091)	0.0136 (0.0198)
Job Search	0.0779*** (0.0176)	0.0877*** (0.0192)	-0.0008 (0.0018)	0.0009 (0.0023)
Unemployment	0.0026 (0.0039)	0.0075 (0.0075)	0.0023 (0.0039)	0.0042 (0.0078)
Federal State	0.0068 (0.0054)	0.0030 (0.0092)	0.0181*** (0.0069)	0.0113 (0.0115)
Year	0.0022 (0.0021)	0.0055 (0.0041)	0.0043 (0.0027)	0.0128** (0.0056)
Personality		-0.0091 (0.0135)		-0.0068 (0.0144)
Observations	4,269	2,055	3,531	1,668

Notes: The dependent variable is the log monthly reservation wage. The male coefficient vector is used as reference vector. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

points (61.7 percent) remain unexplained. The largest contribution to the explained part comes from gender differences in the intention to work full- or part time. This is not surprising, since females are more likely to work part-time and the outcome variable is monthly rather than hourly reservation wages. Gender differences in the perceived job finding chances do also positively contribute to the explained part of the gap. 0.2 percentage points, i.e. 1.8 percent of the explained part and 0.7 percent of the raw gender gap, are attributed to the job finding bias. Thus, numerically the contribution is larger than the contribution of the job loss bias to the wage gap. However, it is not statistically significant in this case, as the sample of unemployed individuals is much smaller. When adding the personality traits and risk preferences in Column (2), the contribution of the job finding bias remains very similar.

In the restricted sample, that only includes job seekers looking for a full-time employment, the raw gender gap in reservation wages is with 18 percent considerably smaller. Furthermore, the explained part of the gender gap is just 3.2 percentage points (17.6 percent) and is only statistically significant at the 5%-level. Yet, the contribution of the job finding bias is now larger in absolute and relative terms. 0.4 percentage points of the raw gender gap is due to differences in the perceived job finding chances in this case. In relative terms, this amounts to 12.9 percent of the explained part and 2.3 percent of the raw gender gap. The inclusion of the personality traits and risk preferences reduces the size of the contribution a bit in the restricted sample, but it still remains positive and comparable in size.

4.5 Effect Heterogeneity

4.5.1 Collective or Individual Wage Agreement

The theoretical predictions of Balleer et al. (2023a) and Balleer et al. (2023b) on the relationship between biased beliefs and wages are based on a model where wages are determined by bargaining between workers and firms. Hence, according to the model, in-

dividual beliefs about job loss and job finding chances are related to wages because they affect the bargaining position of workers. However, in many industries and occupations in Germany wages are determined by collective wage agreements between trade unions and employer associations or individual employers. Workers whose wages are based on a collective wage agreement do not negotiate directly with employers about their pay. For them, individual beliefs are therefore unlikely to affect wages through this channel. For that reason, I study whether the effect of biased beliefs about job insecurity on earnings differs between employees with and without a collectively agreed wage. A considerable smaller effect for employees with collectively agreed wage would support the theoretical reasoning for the relationship. Employed survey participants in the SOEP are asked whether they are paid according to a collectively agreed wage agreement or not. Unfortunately though, this question has only been part of the questionnaire since 2014. Thus, I observe whether wages are collectively agreed or not only in the latest wave of the SOEP that is included in my analysis (2015). Just over half of all individuals in my sample in that year has been paid according to a collective wage agreement.

Table 4.9 displays the estimated effect of the perceived job insecurity on wages within the two subgroups.²² The results in the first three rows are based on the full sample, while the last three rows are based only on full-time employees. Since the information on the wage agreement is only available in 2015, I first estimate the effects in Columns (1) and (4) across both subgroups to compare the results in this reduced sample to my previous results in the full sample. Yet, the estimated effects in 2015 are very similar to the effects in the full sample (Column (4) of Table 4.5 and Column (4) of Table 4.A.2). When I compare the effect between workers with and without a collectively agreed wage, however, I see remarkable differences. For employees who are paid according to a collective wage agreement the effect is almost zero and statistically insignificant. For all other employees, in contrast, the effect is negative, statistically significant, and more pronounced than in the full sample. A one standard deviation increase in the job loss bias reduces wages

²²Table 4.A.4 in Appendix A displays results when additionally controlling for personality traits.

Table 4.9: Collective vs. Individual Wage Agreement

Dependent Variable: Log Hourly Wage						
	Full Sample			Full-Time Only		
	All	Collective Wage		All	Collective Wage	
	(1)	Yes	No	(4)	Yes	No
	(1)	(2)	(3)	(4)	(5)	(6)
Job Loss Bias	-0.0098 (0.0066)	-0.0002 (0.0079)	-0.0168* (0.0090)	-0.0234*** (0.0065)	-0.0136 (0.0091)	-0.0336*** (0.0085)
Predicted Prob.	-0.0080*** (0.0012)	-0.0088*** (0.0016)	-0.0079*** (0.0016)	-0.0038*** (0.0013)	-0.0050*** (0.0018)	-0.0028* (0.0017)
Female	-0.0468*** (0.0137)	-0.0408** (0.0168)	-0.0469** (0.0215)	-0.1111*** (0.0126)	-0.0980*** (0.0159)	-0.1143*** (0.0192)
Observations	13,389	6,929	6,406	8,344	4,484	3,824
R^2	0.524	0.543	0.520	0.554	0.536	0.587

Notes: Results are only based on the survey year 2015. All columns include the same control variables as in Column (4) of Table 4.5, except the year fixed effects. Job Loss Bias is standardized. Standard errors are bootstrapped (400 replications) and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

on average by 1.7 percent among workers without a collectively agreed wage. This stark contrast in the effect is also observed when only considering full-time employees. Although the effect in this case is also larger for individuals with a collectively agreed wage, it remains statistically insignificant and far less pronounced than in the other subgroup. Among full-time employees without a collectively agreed wage, a one standard deviation increase in the job loss bias is associated with 3.4 percent lower wages on average. Hence, my results suggest that the negative effect of the perceived job insecurity on wages is indeed mostly driven by workers whose pay is not subject to a collective wage agreement. This finding bolsters the assertion that a higher perceived job insecurity by employees leads to lower wages because it weakens their bargaining position.

4.5.2 Heterogeneity Across Age and Educational Attainment

Next, I analyze whether the effects of the perceived job loss and job finding probabilities on wages and reservation wages varies between different age and educational groups.

Table 4.10: Wage Effect by Age and Education

Dependent Variable: Log Hourly Wage						
	Full Sample			Full-Time Only		
	(1)	(2)	(3)	(4)	(5)	(6)
(a) Age						
	Young	Medium	Old	Young	Medium	Old
Job Loss Bias	-0.0014 (0.0038)	-0.0133*** (0.0035)	-0.0191*** (0.0046)	-0.0164*** (0.0041)	-0.0163*** (0.0037)	-0.0205*** (0.0048)
Observations	24,142	34,213	19,626	15,681	24,400	14,220
R^2	0.529	0.456	0.467	0.422	0.469	0.497
(b) Education						
	Low	Medium	High	Low	Medium	High
Job Loss Bias	0.0044 (0.0061)	-0.0132*** (0.0028)	-0.0201*** (0.0054)	-0.0113 (0.0070)	-0.0155*** (0.0029)	-0.0307*** (0.0059)
Observations	11,210	47,394	19,377	4,699	34,389	15,213
R^2	0.490	0.402	0.450	0.385	0.433	0.450

Notes: All columns include the same control variables as in Column (4) of Table 4.5. Job Loss Bias is standardized. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Moreover, I inspect the corresponding contributions to the gender gaps across these subgroups. With respect to age, I group individuals into a young (35 or younger), medium (36 - 50), and old (above 50) age group. Similarly, with respect to educational attainment, I define a low-skilled, medium-skilled, and high-skilled subgroup. The low-skilled group includes all respondents who have neither obtained a vocational nor a college degree, the medium-skilled group includes all respondents who have obtained a vocational degree but no college degree, and the high-skilled group includes all respondents with a college degree.

Table 4.10 displays the effects of the job loss bias on hourly wages across the different subgroups, while Table 4.11 shows the contribution to the gender wage gap.²³ In terms of age, the effect is most pronounced in the old age group. Among young employees,

²³Tables 4.A.5 and 4.A.6 in Appendix A display results when controlling for personality traits.

Table 4.11: Wage Decomposition by Age and Education

	Full Sample			Full-Time Only		
	(1)	(2)	(3)	(4)	(5)	(6)
(a) Age						
	Young	Medium	Old	Young	Medium	Old
Gender Gap	0.0996*** (0.0132)	0.2754*** (0.0107)	0.2659*** (0.0137)	0.1338*** (0.0107)	0.2315*** (0.0115)	0.2155*** (0.0158)
Explained Part	0.0821*** (0.0129)	0.1631*** (0.0212)	0.2695*** (0.0370)	0.0546*** (0.0099)	0.0873*** (0.0112)	0.0862*** (0.0206)
Unexplained Part	0.0175 (0.0134)	0.1123*** (0.0224)	-0.0035 (0.0385)	0.0792*** (0.0102)	0.1442*** (0.0115)	0.1292*** (0.0206)
Contribution of						
Job Loss Bias	0.0011* (0.0006)	-0.0002 (0.0004)	0.0011 (0.0007)	0.0015** (0.0007)	-0.0000 (0.0006)	0.0004 (0.0008)
to Explained Part						
Observations	24,142	34,213	19,626	15,681	24,400	14,220
(b) Education						
	Low	Medium	High	Low	Medium	High
Gender Gap	0.0560** (0.0219)	0.2177*** (0.0084)	0.2601*** (0.0152)	0.2189*** (0.0208)	0.1996*** (0.0088)	0.2692*** (0.0165)
Explained Part	0.0528 (0.0344)	0.1912*** (0.0312)	0.1586*** (0.0163)	0.0386 (0.0265)	0.0876*** (0.0106)	0.1471*** (0.0144)
Unexplained Part	0.0031 (0.0358)	0.0265 (0.0318)	0.1015*** (0.0166)	0.1803*** (0.0279)	0.1119*** (0.0109)	0.1221*** (0.0135)
Contribution of						
Job Loss Bias	0.0002 (0.0009)	0.0006 (0.0005)	0.0011 (0.0008)	0.0009 (0.0014)	0.0005 (0.0005)	0.0007 (0.0010)
to Explained Part						
Observations	11,210	47,394	19,377	4,699	34,389	15,213

Notes: The dependent variable is the log hourly wage. The male coefficient vector is used as reference vector. All columns include the same explanatory variables as in Table 4.7. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 4.12: Reservation Wage Effect by Age and Education

Dependent Variable: Log Monthly Reservation Wage						
	Age			Education		
	Young	Medium	Old	Low	Medium	High
	(1)	(2)	(3)	(4)	(5)	(6)
Job Finding Bias	0.0091 (0.0136)	0.0045 (0.0127)	0.0252 (0.0204)	0.0014 (0.0181)	0.0050 (0.0101)	0.0593 (0.0439)
Observations	1,577	1,679	1,013	988	2,798	483
R^2	0.325	0.392	0.300	0.339	0.319	0.327

Notes: All columns include the same control variables as in Column (4) of Table 4.6. Job Finding Bias is standardized. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

in contrast, the effect is much smaller and statistically insignificant. But, when only considering full-time employees, the effect in the young age group is also significant and similar to the effect in the medium age group. With respect to the contribution to the gender wage gap, though, the picture looks different. The contribution of the job loss bias is largest among young and old employees, but it is only significant in the young age group. Among full-time employees, the contribution is substantially larger in the young age group than in the other age groups. Comparing the effect of the perceived job insecurity on wages across educational groups, I see that it is strongest in the high-skilled and weakest in the low-skilled group. Among low-skilled employees the effect is even slightly positive and statistically insignificant. The contribution of the job loss bias to the gender wage gap is also largest in the high-skilled and smallest in the low-skilled group. Yet, the estimated contribution is not statistically significant in any educational group. In the sample of full-time employees, in contrast, the contribution is largest in the low-skilled group (but still insignificant).

The analogous results in case of the job finding bias are displayed in Table 4.12 and 4.13.²⁴ The comparison between the age groups shows that the effect of the perceived job finding chances on reservation wages is largest in the old and smallest in the medium

²⁴Tables 4.A.7 and 4.A.8 in Appendix A display results when controlling for personality traits.

Table 4.13: Reservation Wage Decomposition by Age and Education

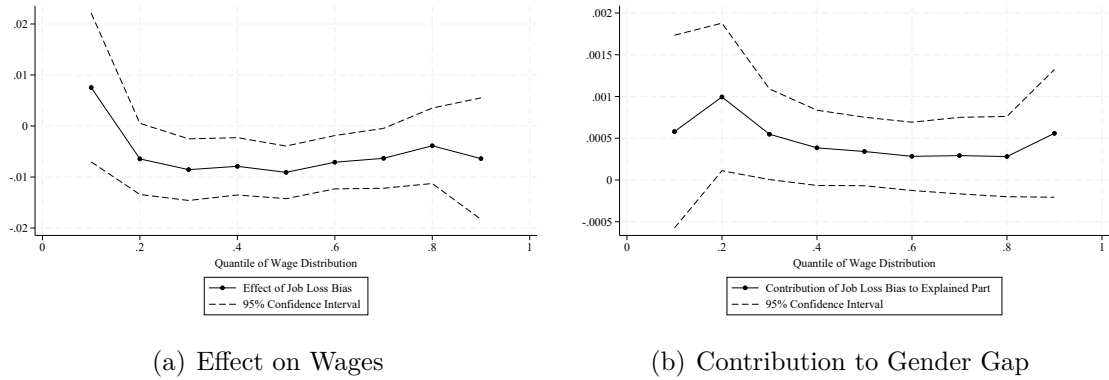
	Age			Education		
	Young (1)	Medium (2)	Old (3)	Low (4)	Medium (5)	High (6)
Gender Gap	0.2357*** (0.0301)	0.2560*** (0.0260)	0.3163*** (0.0438)	0.3251*** (0.0417)	0.2753*** (0.0223)	0.0945 (0.0737)
Explained Part	0.0600 (0.0380)	0.0654** (0.0326)	0.1256 (0.0974)	0.0196 (0.0515)	0.1416*** (0.0395)	0.2502* (0.1458)
Unexplained Part	0.1757*** (0.0452)	0.1905*** (0.0415)	0.1907* (0.1039)	0.3055*** (0.0607)	0.1336*** (0.0452)	-0.1557 (0.1325)
Contribution of						
Job Finding Bias	0.0003 (0.0019)	-0.0006 (0.0025)	0.0050 (0.0063)	0.0010 (0.0045)	0.0007 (0.0019)	-0.0001 (0.0133)
to Explained Part						
Observations	1,577	1,679	1,013	988	2,798	483

Notes: The dependent variable is the log monthly reservation wage. The male coefficient vector is used as reference vector. All columns include the same explanatory variables as in Table 4.8. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

age group. Similarly, the estimated contribution to the gender gap in reservation wages is largest among old employees. Across the educational groups, I see again the largest effect in the high-skilled group. But, the contribution to the gender gap is largest in the low-skilled group. However, as before in the full sample, the estimated effects and contributions to the gender gap are always statistically insignificant in case of the job finding bias.

All in all, my results suggest that the effect of individual labor market expectations on wages and reservation wages as well as the contribution to the gender wage gap are particularly pronounced among college graduates. A possible reason for this finding may be that wage bargaining tends to be more relevant for high-skilled workers and thus individual beliefs are more important for wages in this group. The results with respect to age are less conclusive, but do indicate that individual beliefs about job loss and job finding are more relevant for the gender gaps at a younger or older age than in between.

Figure 4.3: Effect of Job Loss Bias Across Wage Distribution



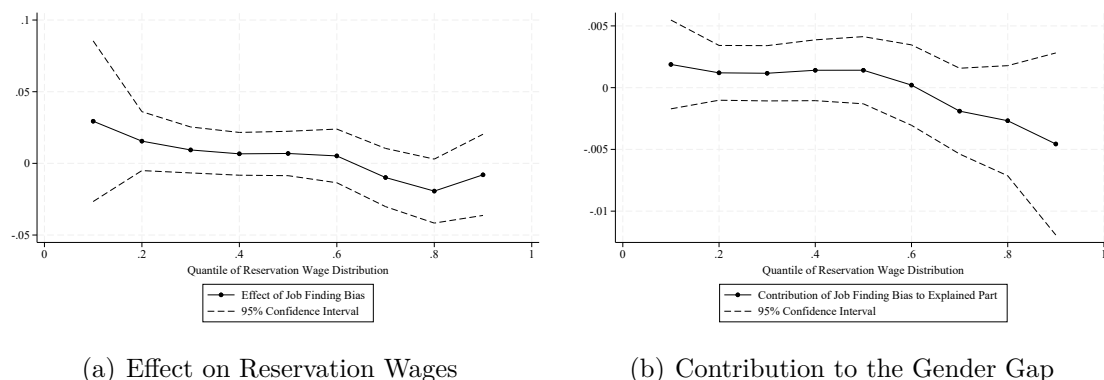
Notes: (a) shows coefficient estimates of Job Loss Bias based on unconditional quantile regressions. Job Loss Bias is standardized. (b) shows estimated contributions of Job Loss Bias to the explained part of the gender gap based on unconditional quantile decompositions. The same control variables as in Column (4) of Table 4.5 are used. The male coefficient vector is used as reference vector in (b). Job Loss Bias is standardized. Standard errors are bootstrapped (400 replications) and are clustered at the individual level. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

4.5.3 Heterogeneity Across the Wage Distribution

Lastly, I examine whether the wage effects and the contribution to the gender gap of the perceived job loss and job finding probabilities vary across the wage and reservation wage distribution, respectively. To estimate the effect of the bias measures on wages and reservation wages at different percentiles of the distribution, I run unconditional quantile regressions proposed by Firpo et al. (2009) that are based on recentered influence functions. Since I pool observations across survey years, however, it is necessary to first adjust the outcome variables for changes in the wage and reservation wage distribution over time, respectively (DiNardo et al., 1996). Hence, I regress log wages and log reservation wages on the set of survey year dummies and use the residuals of the two regressions as the outcome variables in the unconditional quantile regressions. Subsequently, I decompose the gender gaps at unconditional percentiles of the wage distributions using the method of Firpo et al. (2018), which requires to run unconditional quantile regressions separately for females and males. This allows me to quantify the contributions of the job loss and job finding bias to the gender gap at different points of the respective distribution.

The effect of the perceived job loss probability on wages across the hourly wage distri-

Figure 4.4: Effect of Job Finding Bias Across Reservation Wage Distribution



Notes: (a) shows coefficient estimates of Job Finding Bias based on unconditional quantile regressions. Job Loss Bias is standardized. (b) shows estimated contributions of Job Finding Bias to the explained part of the gender gap based on unconditional quantile decompositions. The same control variables as in Column (4) of Table 4.6 are used. The male coefficient vector is used as reference vector in (b). Job Finding Bias is standardized. Standard errors are bootstrapped (400 replications) and are clustered at the individual level. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

bution is displayed in Subfigure (a) of Figure 4.3.²⁵ The results are based on unconditional quantile regressions including the same control variables as in Column (4) of Table 4.5. The effect of the job loss bias is fairly stable between the second and ninth decile of the wage distribution. It is always close to the effect at the mean. In fact, the size of the estimated effect at the mean and at the median are almost identical. However, the results suggest a considerable deviation at the lower and end of the distribution. At the first decile the direction of the effect is reversed. That is, a higher perceived job insecurity is associated with higher wages at the lower end of the distribution. A similar but slightly different picture emerges when inspecting the contribution of the job loss bias to the gender gap across the wage distribution, as is shown in Subfigure (b). The contribution is again relatively stable across the wage distribution and is at most deciles close the value at the mean. However, it is a bit larger at the second decile of the distribution.

Figure 4.4 shows the estimated reservation wage effect of the job finding bias and its contribution to the gender gap across the reservation wage distribution, respectively.²⁶

²⁵Figure 4.A.11 in Appendix A displays results when only including full-time employed workers.

²⁶Figure 4.A.12 in Appendix A displays results when only including job-seekers who intend to work full-time.

For this analysis, the same control variables as in Column (4) of Table 4.6 are used. From the first to the sixth decile of the distribution the estimated effect is always relatively close to the effect found at the mean. From the seventh decile upwards, though, the effect turns negative. Thus, at the upper part of the reservation wage distribution higher perceived job finding chances are associated with lower reservation wages. Yet, as was the case at the mean, the estimated effect is never statistically significant. The estimated contributions of the job finding bias to the gender gap match quite well with the estimated reservation wage effects across the distribution. In the lower half of the distribution the contribution is always positive and close to the contribution at the mean. In the upper half, in contrast, the estimated contribution is either zero (sixth decile) or negative. Once more, however, the estimates are not statistically significant.

4.6 Conclusion

Despite growing evidence that individuals' beliefs about their labor market prospects can be important for their subsequent labor market choices and outcomes, gender differences with regard to job loss and job finding expectations have not been studied before. Using data from the German Socio-Economic Panel, this paper documents that there exist significant differences in beliefs about job insecurity and job finding chances between women and men. Women, on average, perceive a job loss as more and a job finding as less likely than comparable men. Since I consider the subjective beliefs in relation to objective probabilities based on a rich set of observed individual characteristics, these gender differences cannot be explained by differences in actual job risk and job finding chances.

Furthermore, my results confirm previous findings detecting a negative relationship between job loss expectations and wages and a negative relationship between job finding expectations and reservation wages. Beyond that, I demonstrate that the negative effect of perceived job insecurity on wages is mainly driven by employees without a collectively

agreed wage. This finding reinforces claims that job loss expectations affect wages due to its impact on workers' bargaining position in wage negotiations.

In terms of the gender wage gap, my results suggest that a small but significant portion of the gap can be attributed to gender differences in beliefs about job insecurity. I also find a positive contribution of job finding expectations on the gender differential in reservation wages of job-seekers, that is larger in magnitude but never statistically significant. Moreover, my findings indicate that the effects of labor market expectations on wages and reservation wages and on the corresponding gender differentials are particularly pronounced among college graduates.

All together, my findings show that lower confidence of women in keeping or finding a job can be an explanatory factor for gender disparities in labor market outcomes. While the estimated effect on the gender wage gap in the entire working population is rather small, these differences in beliefs still matter, particularly for high skilled individuals and workers who directly negotiate with employers about their pay. Further, beliefs about job loss and job finding probabilities are presumably not the only labor market expectations relevant for wage differences. Thus, future research should study in which other aspect of the labor market gender differences in beliefs emerge and how they affect observed gender inequalities in the labor market. What is more, there is a need to better understand how and why gender differences in labor market expectations arise.

References

- Adamecz-Völgyi, A. and Shure, N. (2022). The gender gap in top jobs – The role of overconfidence. *Labour Economics*, 79:102283.
- Ahrens, A., Hansen, C. B., and Schaffer, M. E. (2023). pystacked: Stacking generalization and machine learning in Stata. *The Stata Journal*, 23(4):909–931.
- Balleer, A., Duernecker, G., Forstner, S., and Goensch, J. (2023a). DP18005 Biased expectations and labor market outcomes: Evidence from German survey data and implications for the East-West wage gap. CEPR Press Discussion Paper 18005.

- Balleer, A., Duernecker, G., Forstner, S., and Goensch, J. (2023b). DP18019 Wage Bargaining and Labor Market Policy with Biased Expectations. CEPR Press Discussion Paper 18019.
- Blau, F. D. and Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3):789–865.
- Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources*, 8(4):436–455.
- Breiman, L. (1996). Stacked regressions. *Machine Learning*, 24(1):49–64.
- Briel, S., Osikominu, A., Pfeifer, G., Reutter, M., and Satlukal, S. (2022). Gender differences in wage expectations: the role of biased beliefs. *Empirical Economics*, 62(1):187–212.
- Campbell, D., Carruth, A., Dickerson, A., and Green, F. (2007). Job Insecurity and Wages. *The Economic Journal*, 117(518):544–566.
- Cortés, P., Pan, J., Pilossoph, L., Reuben, E., and Zafar, B. (2023). Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab*. *The Quarterly Journal of Economics*. qjad017.
- Dawson, C. (2017). The upside of pessimism. Biased beliefs and the paradox of the contented female worker. *Journal of Economic Behavior and Organization*, 135:215–228.
- Dickerson, A. and Green, F. (2012). Fears and realisations of employment insecurity. *Labour Economics*, 19(2):198–210.
- DiNardo, J., Fortin, N. M., and Lemieux, T. (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica*, 64(5):1001–1044.
- Eurostat (2021). Gender Pay Gap Statistics. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Gender_pay_gap_statistics. Accessed: 2024-01-12.
- Firpo, S. P., Fortin, N. M., and Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, 77(3):953–973.
- Firpo, S. P., Fortin, N. M., and Lemieux, T. (2018). Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics*, 6(2).

- Goldin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4):1091–1119.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*. Springer New York.
- Hendren, N. (2017). Knowledge of Future Job Loss and Implications for Unemployment Insurance. *American Economic Review*, 107(7):1778–1823.
- Jäger, S., Roth, C., Roussille, N., and Schoefer, B. (2024). Worker Beliefs About Outside Options*. *The Quarterly Journal of Economics*, 139(3):1505–1556.
- Mueller, A. I. and Spinnewijn, J. (2023a). Chapter 22 - Expectations data, labor market, and job search. In Bachmann, R., Topa, G., and van der Klaauw, W., editors, *Handbook of Economic Expectations*, pages 677–713. Academic Press.
- Mueller, A. I. and Spinnewijn, J. (2023b). The Nature of Long-Term Unemployment: Predictability, Heterogeneity and Selection. Working Paper 30979, National Bureau of Economic Research.
- Mueller, A. I., Spinnewijn, J., and Topa, G. (2021). Job Seekers’ Perceptions and Employment Prospects: Heterogeneity, Duration Dependence, and Bias. *American Economic Review*, 111(1):324–63.
- Niederle, M. and Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics*, 122(3):1067–1101.
- Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14(3):693–709.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- Reuben, E., Wiswall, M., and Zafar, B. (2017). Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender. *The Economic Journal*, 127(604):2153–2186.
- Spinnewijn, J. (2015). Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs. *Journal of the European Economic Association*, 13(1):130–167.

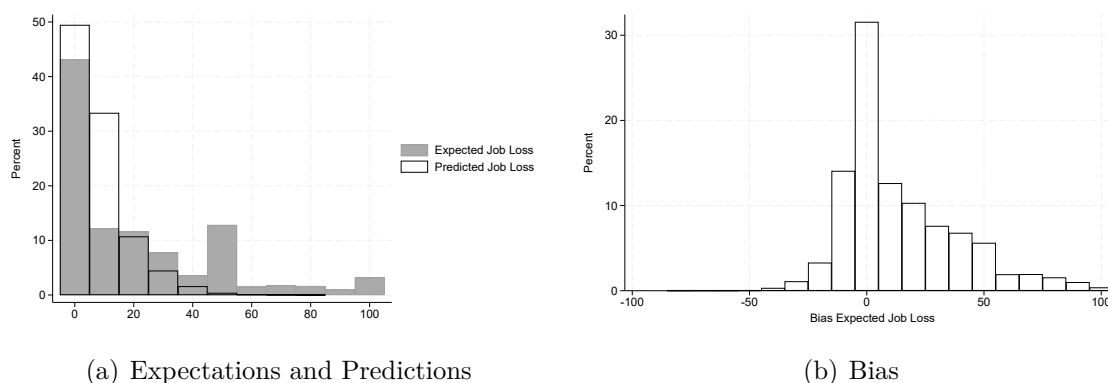
van der Laan, M. J., Polley, E. C., and Hubbard, A. E. (2007). Super Learner. *Statistical Applications in Genetics and Molecular Biology*, 6(1).

Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, 5(2):241–259.

Appendix

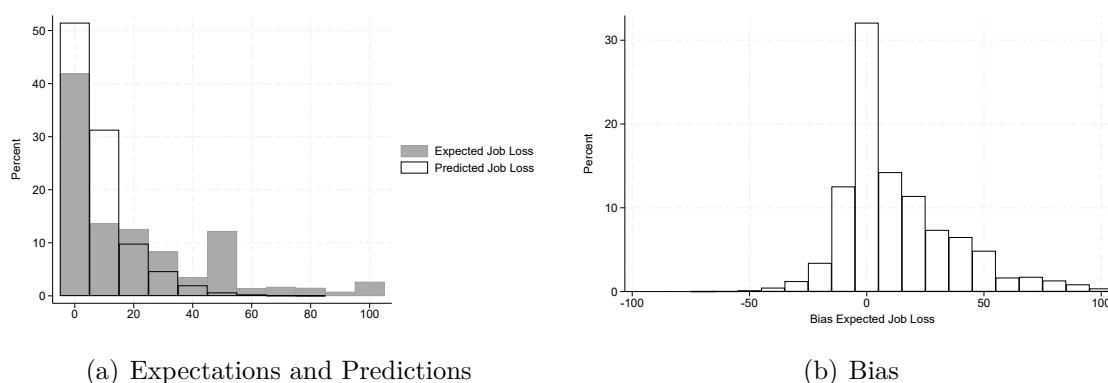
A Additional Results

Figure 4.A.1: Distribution of Expected and Predicted Job Loss Probabilities, Females



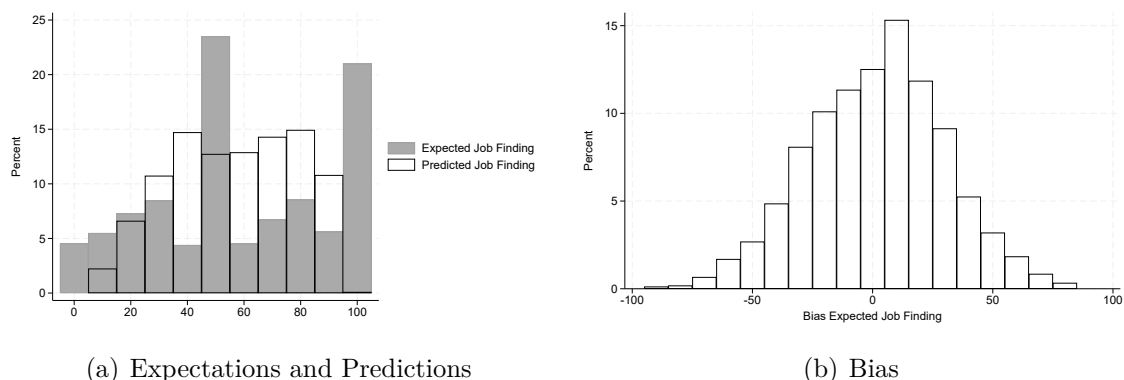
Notes: (a) displays histogram of expected and predicted job loss probabilities for females. (b) displays histogram of bias in job loss expectations (difference between expectations and predictions) for females. Predictions are rounded to the closest decile on the probability scale. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 4.A.2: Distribution of Expected and Predicted Job Loss Probabilities, Males



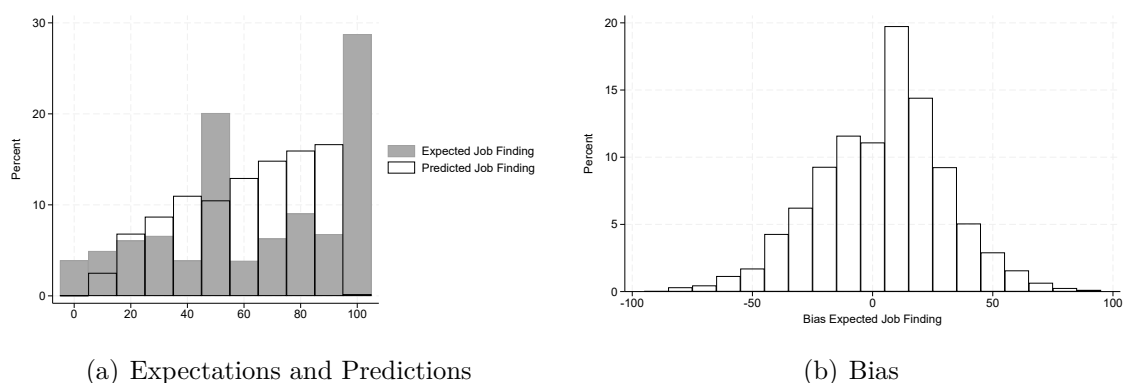
Notes: (a) displays histogram of expected and predicted job loss probabilities for males. (b) displays histogram of bias in job loss expectations (difference between expectations and predictions) for males. Predictions are rounded to the closest decile on the probability scale. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 4.A.3: Distribution of Expected and Predicted Job Finding Probabilities, Females



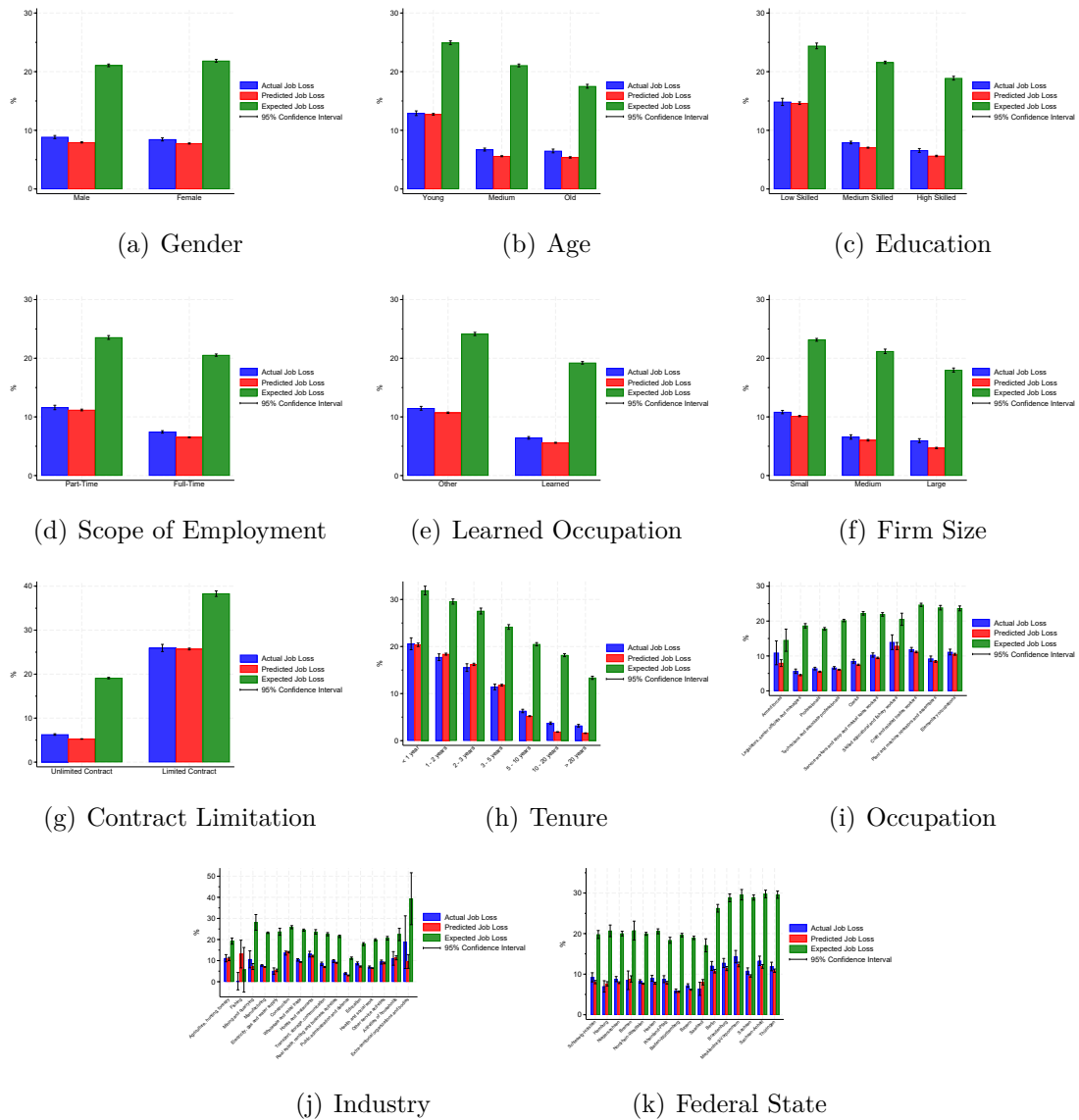
Notes: (a) displays histogram of expected and predicted job finding probabilities for females. (b) displays histogram of bias in job finding expectations (difference between expectations and predictions) for females. Predictions are rounded to the closest decile on the probability scale. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 4.A.4: Distribution of Expected and Predicted Job Finding Probabilities, Males



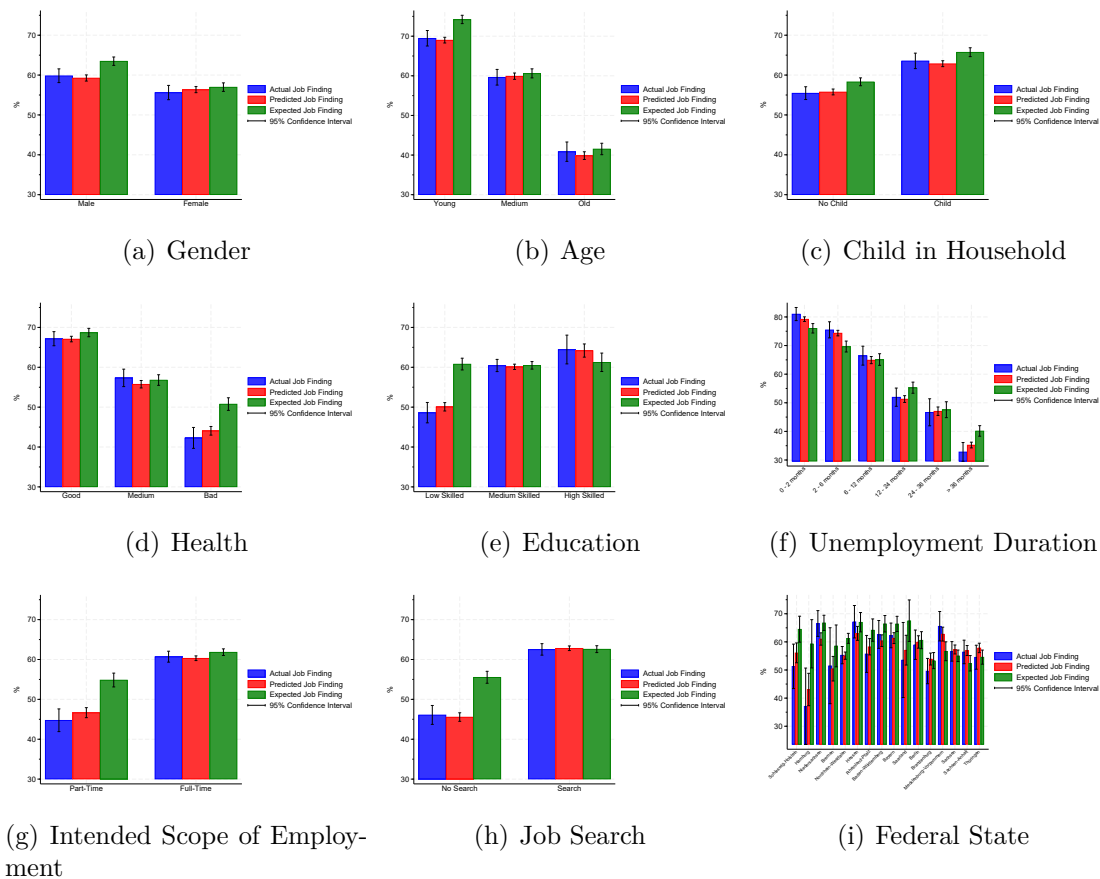
Notes: (a) displays histogram of expected and predicted job finding probabilities for males. (b) displays histogram of bias in job finding expectations (difference between expectations and predictions) for males. Predictions are rounded to the closest decile on the probability scale. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 4.A.5: Job Loss Probabilities Within Subgroups



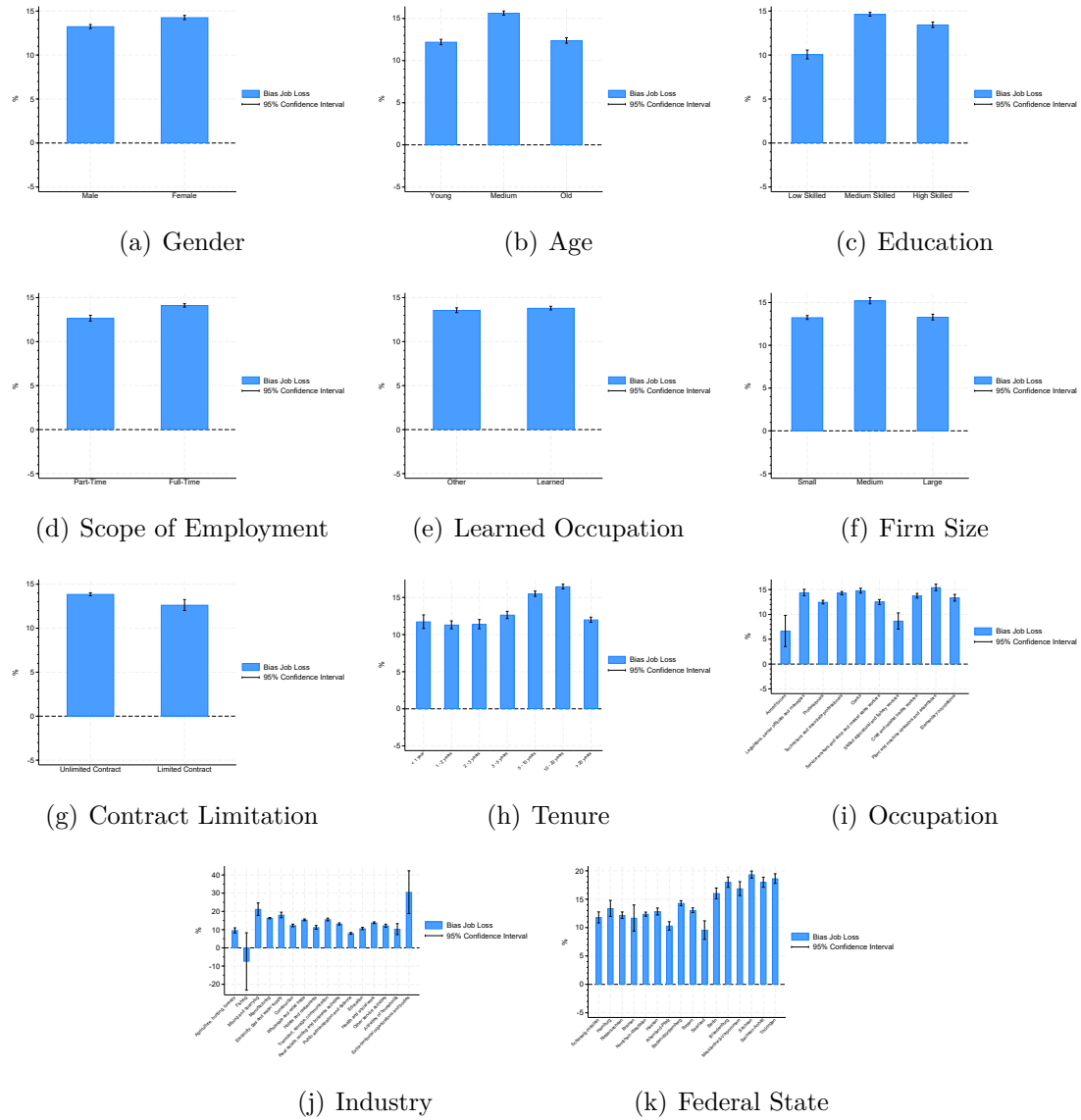
Notes: Figure displays the share of realized job losses and mean values of predicted and expected job loss probabilities within various subgroups. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 4.A.6: Job Finding Probabilities Within Subgroups



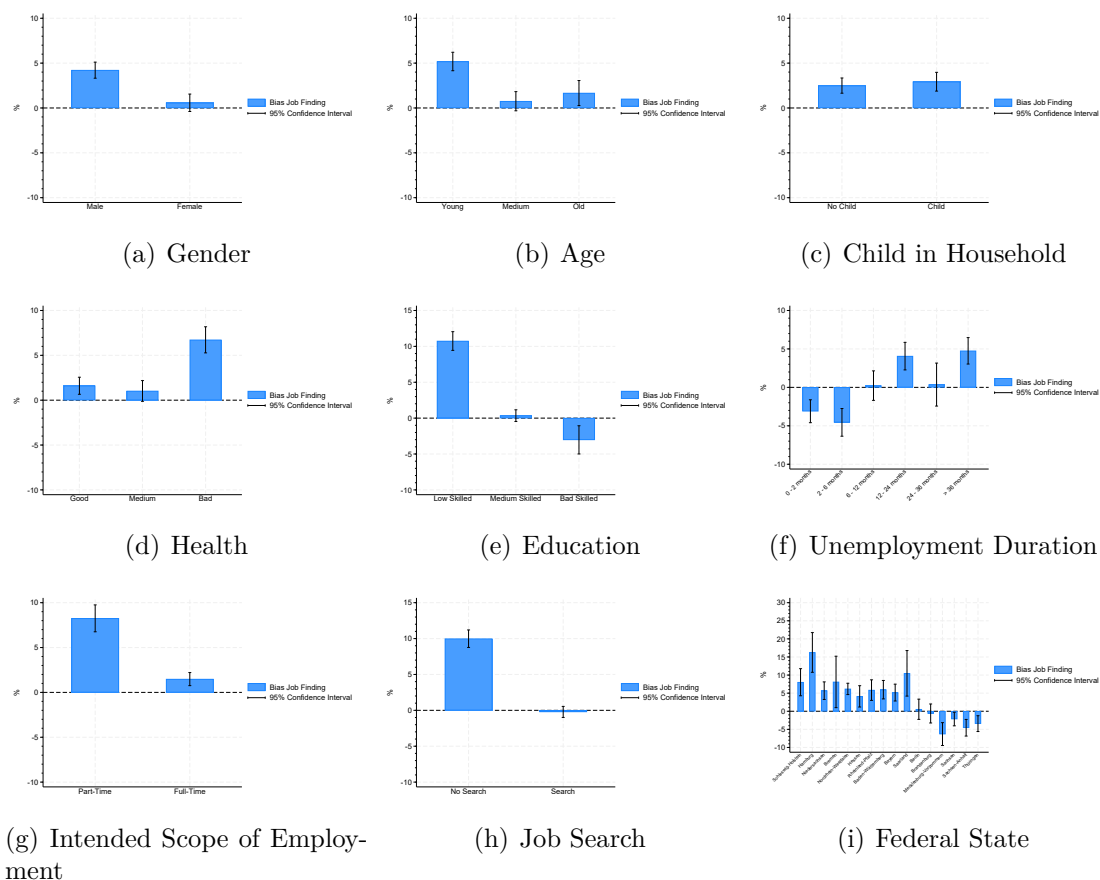
Notes: Figure displays the share of realized job findings and mean values of predicted and expected job finding probabilities within various subgroups. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 4.A.7: Job Loss Bias Within Subgroups



Notes: Figure displays mean values of bias in job loss expectations (difference between expectations and predictions) within various subgroups. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 4.A.8: Job Finding Bias Within Subgroups



Notes: Figure displays mean values of bias in job finding expectations (difference between expectations and predictions) within various subgroups. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

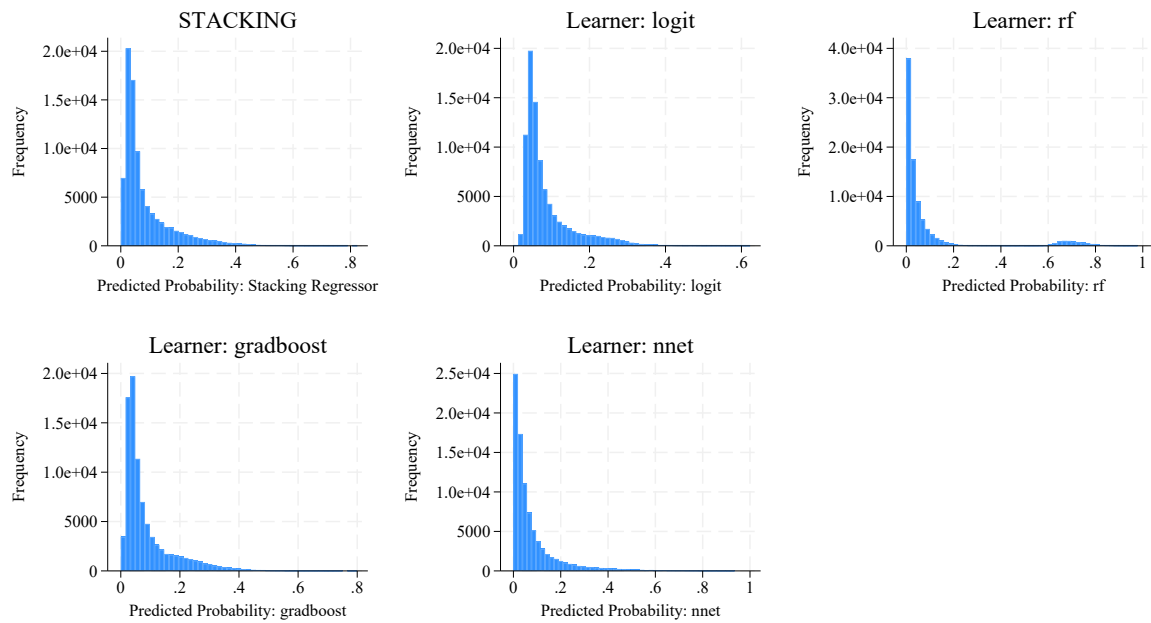
Table 4.A.1: Stacking Weights used for Predictions

	(1)	(2)
Base Learner	Job Loss	Job Finding
Logit	0.000	0.147
Random Forest	0.121	0.011
Gradient Boosting	0.716	0.413
Neural Net	0.162	0.428

Notes: Table displays the weights of different base learners chosen by the stacking algorithm for the predictions of a job loss and job finding.

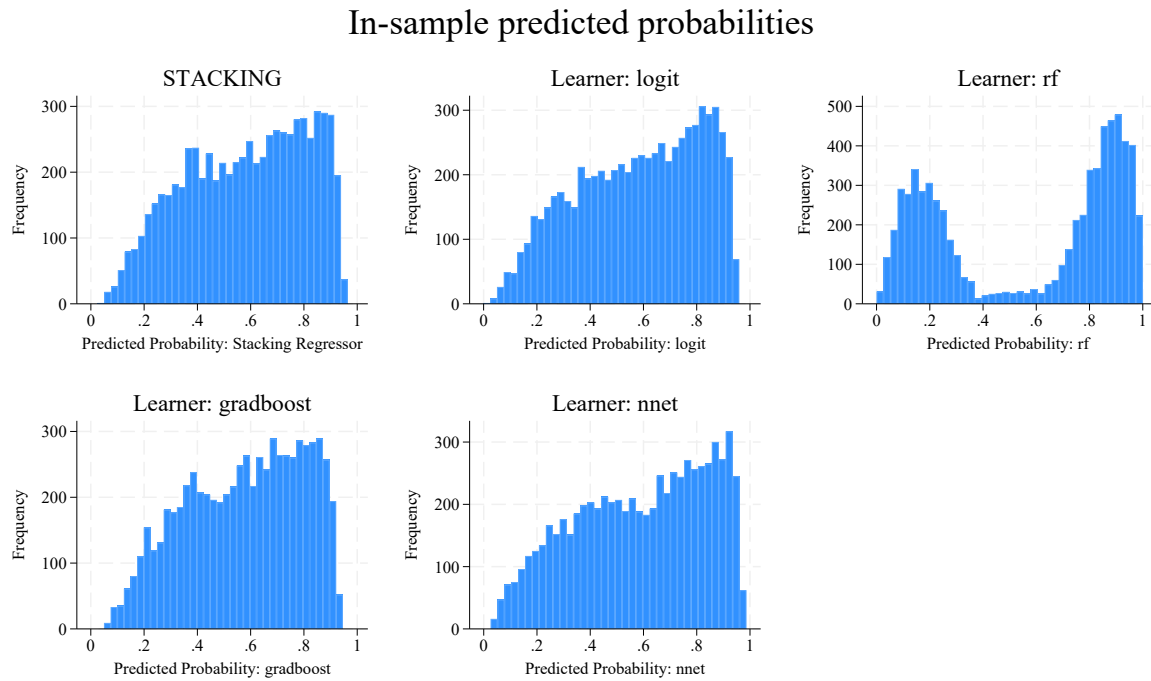
Figure 4.A.9: Predicted Job Loss Probabilities by Base Learner

In-sample predicted probabilities



Notes: Figure displays predicted job loss probabilities based on the different base learners and the stacking algorithm. Probabilities are not rounded. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 4.A.10: Predicted Job Finding Probabilities by Base Learner



Notes: Figure displays predicted job finding probabilities based on the different base learners and the stacking algorithm. Probabilities are not rounded. *Source:* German Socio-Economic Panel v37, own calculations.

Table 4.A.2: Wage Effect by Gender

Dependent Variable: Log Hourly Wage						
	Full Sample			Full-Time Only		
	All	Female	Male	All	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
Job Loss Bias	-0.0098*** (0.0024)	-0.0042 (0.0033)	-0.0186*** (0.0034)	-0.0195*** (0.0026)	-0.0142*** (0.0040)	-0.0241*** (0.0032)
Predicted Prob.	-0.0070*** (0.0004)	-0.0080*** (0.0006)	-0.0054*** (0.0005)	-0.0034*** (0.0004)	-0.0036*** (0.0007)	-0.0033*** (0.0005)
Female	-0.1027*** (0.0068)			-0.1423*** (0.0068)		
Observations	77,981	37,352	40,629	54,301	18,095	36,206
R^2	0.506	0.420	0.560	0.485	0.429	0.487

Notes: All columns include the same control variables as in Column (4) of Table 4.5. Job Loss Bias is standardized. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 4.A.3: Reservation Wage Effect by Gender

Dependent Variable: Log Monthly Reservation Wage						
	Full Sample			Looking for Full-Time		
	All	Female	Male	All	Female	Male
	(1)	(2)	(3)	(4)	(5)	
Job Finding Bias	0.0098 (0.0094)	0.0083 (0.0138)	0.0150 (0.0123)	0.0135 (0.0097)	0.0128 (0.0153)	0.0162 (0.0124)
Predicted Prob.	-0.0008 (0.0010)	-0.0029* (0.0015)	0.0001 (0.0013)	0.0003 (0.0011)	-0.0020 (0.0018)	0.0013 (0.0013)
Female	-0.1399*** (0.0197)			-0.1392*** (0.0196)		
Observations	4269	1972	2297	3,531	1,348	2,183
R^2	0.283	0.298	0.210	0.209	0.213	0.191

Notes: All columns include the same control variables as in Column (4) of Table 4.6. Job Finding Bias is standardized. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 4.A.4: Collective vs. Individual Wage Agreement (With Personality Traits)

Dependent Variable: Log Hourly Wage						
	Full Sample			Full-Time Only		
	All	Collective Wage		All	Collective Wage	
	(1)	Yes	No	(4)	Yes	No
	(1)	(2)	(3)	(4)	(5)	(6)
Job Loss Bias	-0.0079 (0.0071)	-0.0028 (0.0077)	-0.0100 (0.0098)	-0.0227*** (0.0066)	-0.0121 (0.0077)	-0.0312*** (0.0097)
Predicted Prob.	-0.0087*** (0.0014)	-0.0104*** (0.0017)	-0.0081*** (0.0019)	-0.0035** (0.0014)	-0.0055*** (0.0018)	-0.0018 (0.0020)
Female	-0.0298* (0.0160)	-0.0254 (0.0200)	-0.0354 (0.0232)	-0.0951*** (0.0147)	-0.0793*** (0.0178)	-0.0931*** (0.0218)
Observations	9,847	5,219	4,588	6,216	3,379	2,811
R^2	0.536	0.533	0.565	0.580	0.564	0.627

Notes: Results are only based on the survey year 2015. All columns include the same control variables as in Column (5) of Table 4.5, except the year fixed effects. Job Loss Bias is standardized. Standard errors are bootstrapped (400 replications) and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 4.A.5: Wage Effect by Age and Education (With Personality Traits)

Dependent Variable: Log Hourly Wage						
	Full Sample			Full-Time Only		
	(1)	(2)	(3)	(4)	(5)	(6)
(a) Age						
	Young	Medium	Old	Young	Medium	Old
Job Loss Bias	0.0002 (0.0059)	-0.0077 (0.0051)	-0.0190*** (0.0064)	-0.0168*** (0.0058)	-0.0088* (0.0049)	-0.0157*** (0.0059)
Observations	10,214	18,353	11,777	6,646	12,607	8,409
R^2	0.511	0.489	0.482	0.456	0.510	0.510
(b) Education						
	Low	Medium	High	Low	Medium	High
Job Loss Bias	0.0066 (0.0100)	-0.0093** (0.0039)	-0.0210*** (0.0068)	0.0004 (0.0111)	-0.0116*** (0.0038)	-0.0300*** (0.0074)
Observations	4,367	25,117	10,860	1,872	17,442	8,348
R^2	0.494	0.423	0.484	0.441	0.462	0.487

Notes: All columns include the same control variables as in Column (5) of Table 4.5. Job Loss Bias is standardized. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 4.A.6: Wage Decomposition by Age and Education (With Personality Traits)

	Full Sample			Full-Time Only		
	(1)	(2)	(3)	(4)	(5)	(6)
(a) Age						
	Young	Medium	Old	Young	Medium	Old
Gender Gap	0.0475** (0.0204)	0.2696*** (0.0140)	0.2537*** (0.0179)	0.0958*** (0.0166)	0.2236*** (0.0153)	0.1976*** (0.0197)
Explained Part	0.0844*** (0.0216)	0.2021*** (0.0296)	0.2729*** (0.0480)	0.0483*** (0.0172)	0.0985*** (0.0156)	0.1099*** (0.0298)
Unexplained Part	-0.0369* (0.0220)	0.0675** (0.0309)	-0.0192 (0.0510)	0.0475*** (0.0174)	0.1252*** (0.0158)	0.0878*** (0.0290)
Contribution of						
Job Loss Bias	0.0013 (0.0012)	-0.0004 (0.0005)	0.0009 (0.0007)	0.0014 (0.0010)	-0.0005 (0.0006)	-0.0000 (0.0006)
to Explained Part						
Observations	10,214	18,353	11,777	6,646	12,607	8,409
(b) Education						
	Low	Medium	High	Low	Medium	High
Gender Gap	0.0106 (0.0366)	0.2123*** (0.0112)	0.2431*** (0.0203)	0.2009*** (0.0325)	0.1879*** (0.0122)	0.2613*** (0.0210)
Explained Part	0.0741 (0.0696)	0.2354*** (0.0534)	0.1769*** (0.0218)	0.0551 (0.0504)	0.1070*** (0.0168)	0.1618*** (0.0189)
Unexplained Part	-0.0636 (0.0699)	-0.0232 (0.0541)	0.0662*** (0.0218)	0.1458*** (0.0504)	0.0809*** (0.0168)	0.0994*** (0.0189)
Contribution of						
Job Loss Bias	0.0006 (0.0021)	0.0004 (0.0005)	0.0005 (0.0010)	-0.0003 (0.0015)	0.0002 (0.0005)	-0.0003 (0.0013)
to Explained Part						
Observations	4,367	25,117	10,860	1,872	17,442	8,348

Notes: The dependent variable is the log hourly wage. The male coefficient vector is used as reference vector. All columns include the same explanatory variables as in Table 4.7 as well as personality traits. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 4.A.7: Reservation Wage Effect by Age and Education (With Personality Traits)

Dependent Variable: Log Monthly Reservation Wage						
	Age			Education		
	Young	Medium	Old	Low	Medium	High
	(1)	(2)	(3)	(4)	(5)	(6)
Job Finding Bias	-0.0192	-0.0114	0.0352	0.0020	0.0033	0.0377
	(0.0186)	(0.0186)	(0.0341)	(0.0334)	(0.0140)	(0.0844)
Observations	701	795	559	438	1,377	240
R^2	0.427	0.464	0.352	0.429	0.371	0.460

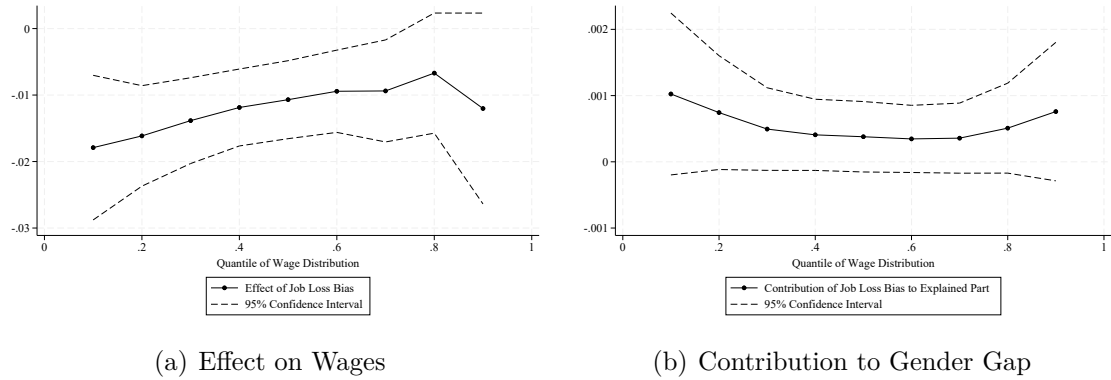
Notes: All columns include the same control variables as in Column (5) of Table 4.6. Job Finding Bias is standardized. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 4.A.8: Reservation Wage Decomposition by Age and Education (With Personality Traits)

	Age			Education		
	Young	Medium	Old	Low	Medium	High
	(1)	(2)	(3)	(4)	(5)	(6)
Gender Gap	0.2035***	0.2676***	0.3057***	0.2370***	0.2814***	0.2071*
	(0.0442)	(0.0405)	(0.0644)	(0.0613)	(0.0342)	(0.1096)
Explained Part	-0.0414	0.0487	0.3873	0.0325	0.1745**	0.2771
	(0.0657)	(0.0604)	(0.3264)	(0.1064)	(0.0701)	(0.4903)
Unexplained Part	0.2449***	0.2189***	-0.0816	0.2045*	0.1069	-0.0699
	(0.0698)	(0.0665)	(0.3332)	(0.1117)	(0.0762)	(0.4969)
Contribution of						
Job Finding Bias	0.0001	-0.0045	0.0062	0.0003	-0.0004	0.0168
to Explained Part	(0.0039)	(0.0062)	(0.0152)	(0.0107)	(0.0037)	(0.0645)
Observations	701	795	559	438	1,377	240

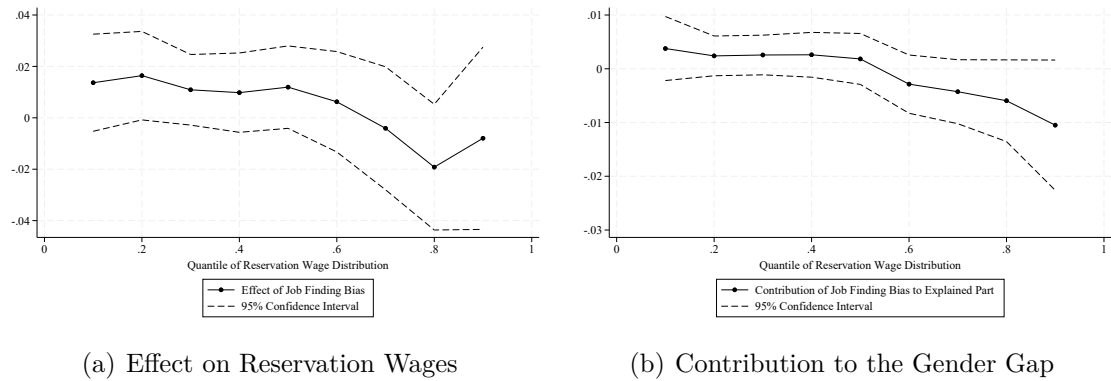
Notes: The dependent variable is the log monthly reservation wage. The male coefficient vector is used as reference vector. All columns include the same explanatory variables as in Table 4.8 as well as personality traits. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 4.A.11: Effect of Job Loss Bias Across Wage Distribution (Full-Time Only)



Notes: (a) shows coefficient estimates of Job Loss Bias based on unconditional quantile regressions. Job Loss Bias is standardized. (b) shows estimated contributions of Job Loss Bias to the explained part of the gender gap based on unconditional quantile decompositions. The same control variables as in Column (4) of Table 4.5 are used (except indicator for full-time employment). The male coefficient vector is used as reference vector in (b). Job Loss Bias is standardized. Standard errors are bootstrapped (400 replications) and are clustered at the individual level. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Figure 4.A.12: Effect of Job Finding Bias Across Reservation Wage Distribution (Looking for Full-time Only)



Notes: (a) shows coefficient estimates of Job Finding Bias based on unconditional quantile regressions. Job Loss Bias is standardized. (b) shows estimated contributions of Job Finding Bias to the explained part of the gender gap based on unconditional quantile decompositions. The same control variables as in Column (4) of Table 4.6 are used (except indicator for intention to be employed full-time). The male coefficient vector is used as reference vector in (b). Job Finding Bias is standardized. Standard errors are bootstrapped (400 replications) and are clustered at the individual level. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

B Robustness Checks

Table 4.B.1: Gender Difference in Job Loss Expectations (With Exact Predicted Prob.)

Dependent Variable: Job Loss Bias		
	(1)	(2)
Female	1.5394*** (0.3832)	1.9508*** (0.5396)
Predicted Prob.	80.5067*** (3.3590)	83.0995*** (5.1539)
Personality		✓
Observations	82,539	43,025
R^2	0.213	0.229

Notes: Both columns include additionally the same set of control variables as in Column (5) of Table 4.3. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 4.B.2: Gender Difference in Job Finding Expectations (With Exact Predicted Prob.)

Dependent Variable: Job Finding Bias		
	(1)	(2)
Female	-3.2590*** (1.1340)	-3.5817** (1.6044)
Predicted Prob.	63.6326*** (6.3985)	59.8040*** (9.1682)
Personality		✓
Observations	6,484	3,091
R^2	0.373	0.411

Notes: Both columns include additionally the same set of control variables as in Column (5) of Table 4.4. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 4.B.3: Bias in Job Loss Expectation and Wages (With Exact Predicted Prob.)

Dependent Variable: Log Hourly Wage		
	(1)	(2)
Job Loss Bias	-0.0082*** (0.0026)	-0.0051 (0.0037)
Predicted Prob.	-0.7974*** (0.0478)	-0.9175*** (0.0766)
Female	-0.1039*** (0.0068)	-0.0760*** (0.0101)
Personality		✓
Observations	77,981	40,344
R^2	0.507	0.509

Notes: Job Loss Bias is standardized. Both columns include additionally the same set of control variables as in Column (4) of Table 4.5. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 4.B.4: Bias in Job Finding Expectation and Reservation Wages (With Exact Predicted Prob.)

Dependent Variable: Log Monthly Reservation Wage		
	(1)	(2)
Job Finding Bias	0.0118 (0.0104)	0.0076 (0.0164)
Predicted Prob.	-0.1641 (0.1108)	-0.2900 (0.1776)
Female	-0.1397*** (0.0197)	-0.1294*** (0.0275)
Personality		✓
Observations	4,269	2,055
R^2	0.284	0.305

Notes: Job Finding Bias is standardized. Both columns include additionally the same set of control variables as in Column (4) of Table 4.6. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 4.B.5: Wage Decomposition (With Exact Predicted Prob.)

	Full Sample		Full-Time Only	
	(1)	(2)	(3)	(4)
Aggregate Decomposition				
Gender Gap	0.2147*** (0.0078)	0.2011*** (0.0106)	0.2106*** (0.0080)	0.1963*** (0.0107)
Explained Part	0.1678*** (0.0149)	0.1915*** (0.0202)	0.0814*** (0.0078)	0.0977*** (0.0117)
Unexplained Part	0.0468*** (0.0154)	0.0096 (0.0207)	0.1292*** (0.0075)	0.0986*** (0.0111)
Contribution To Explained Part				
Job Loss Bias	0.0005* (0.0003)	0.0003 (0.0003)	0.0005 (0.0005)	-0.0000 (0.0004)
Personality		✓		✓
Observations	77,981	40,344	54,301	27,662

Notes: The dependent variable is the log hourly wage. All columns additionally include the same set of control variables as in Table 4.7. The male coefficient vector is used as reference vector. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

Table 4.B.6: Reservation Wage Decomposition (With Exact Predicted Prob.)

	Full Sample		Looking for Full-Time	
	(1)	(2)	(3)	(4)
Aggregate Decomposition				
Gender Gap	0.2621*** (0.0193)	0.2563*** (0.0296)	0.1804*** (0.0194)	0.1745*** (0.0289)
Explained Part	0.1009*** (0.0230)	0.1055*** (0.0342)	0.0318** (0.0150)	0.0144 (0.0304)
Unexplained Part	0.1612*** (0.0293)	0.1509*** (0.0409)	0.1486*** (0.0212)	0.1601*** (0.0373)
Contribution To Explained Part				
Job Finding Bias	0.0035 (0.0030)	0.0028 (0.0054)	0.0052 (0.0041)	0.0050 (0.0080)
Personality		✓		✓
Observations	4,269	2,055	3,531	1,668

Notes: The dependent variable is the log monthly reservation wage. All columns additionally include the same set of control variables as in Table 4.8. The male coefficient vector is used as reference vector. Standard errors are bootstrapped (400 replications), are clustered at the individual level, and are shown in parentheses. *, **, and *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively. SOEP sample weights used. *Source:* German Socio-Economic Panel v37, own calculations.

C Predictors and Control Variables

Table 4.C.1: Predictors and Control Variables

Predictors of Job Loss	
Demographics:	gender, age, age squared ^b , married, child under 16 in household, migration background
Education:	school leaving degree, vocational degree, college degree
Experience:	full-time experience, full-time experience squared ^b , part-time experience, part-time experience squared ^b , unemployment experience, unemployment experience squared ^b
Job Characteristics:	tenure ^a , dummy variables for value of tenure ^b , size of firm, full-time employment, limited contract, learned occupation, public sector, civil servant, self-employed, occupation (ISCO88, 1-digit), industry (NACE, level 1)
Region:	federal state
Time:	year ^a , industry $\times$ year ^b
Predictors of Job Finding	
Demographics:	gender, age, age squared ^b , married, child under 16 in household, migration background, subjective health
Education:	school leaving degree, vocational degree, college degree
Experience:	full-time experience, full-time experience squared ^b , part-time experience, part-time experience squared ^b , unemployment experience, unemployment experience squared ^b
Job Search:	active job search in last four weeks, intended scope of employment, employment intended immediately
Unemployment:	Duration of current unemployment spell, receiving <i>Arbeitslosengeld 1</i> , receiving <i>Arbeitslosengeld 2</i>
Region:	federal state
Time:	survey year
Control Variables in Job Loss Bias Equation	
Demographics:	age, age squared, married, child under 16 in household, migration background
Education:	school leaving degree, vocational degree, college degree
Experience:	full-time experience, full-time experience squared, part-time experience, part-time experience squared, unemployment experience, unemployment experience squared
Job Characteristics:	tenure, tenure squared, size of firm, full-time employment, limited contract, learned occupation, public sector, civil servant, work satisfaction, occupation (ISCO88, 1-digit), industry (NACE, level 1)
Region:	federal state
Time:	year
Personality:	big five, positive reciprocity, negative reciprocity, external locus of control, internal locus of control, risk tolerance
Control Variables in Job Finding Bias Equation	
Demographics:	age, age squared, married, child under 16 in household, migration background, log household income
Education:	school leaving degree, vocational degree, college degree
Experience:	full-time experience, full-time experience squared, part-time experience, part-time experience squared, unemployment experience, unemployment experience squared

Continued on next page

Job Search:	active job search in last four weeks, intended scope of employment, employment intended immediately
Unemployment:	Duration of current unemployment spell, receiving <i>Arbeitslosengeld 1</i> , receiving <i>Arbeitslosengeld 2</i>
Region:	federal state
Time:	year
Personality:	big five, positive reciprocity, negative reciprocity, external locus of control, internal locus of control, risk tolerance

Control Variables in Wage Equation and Decomposition

Demographics:	age, age squared, married, child under 16 in household, migration background
Education:	school leaving degree, vocational degree, college degree
Experience:	full-time experience, full-time experience squared, part-time experience, part-time experience squared, unemployment experience, unemployment experience squared
Job Characteristics:	tenure, tenure squared, size of firm, full-time employment, public sector, civil servant, occupation (ISCO88, 1-digit), industry (NACE, level 1)
Region:	federal state
Time:	industry $\times$ year
Personality:	big five, positive reciprocity, negative reciprocity, external locus of control, internal locus of control, risk tolerance

Control Variables in Reservation Wage Equation and Decomposition

Demographics:	age, age squared, married, child under 16 in household, migration background, log household income
Education:	school leaving degree, vocational degree, college degree
Experience:	full-time experience, full-time experience squared, part-time experience, part-time experience squared, unemployment experience, unemployment experience squared
Job Search:	active job search in last four weeks, intended scope of employment, employment intended immediately
Unemployment:	Duration of current unemployment spell, receiving <i>Arbeitslosengeld 1</i> , receiving <i>Arbeitslosengeld 2</i>
Region:	federal state
Time:	year
Personality:	big five, positive reciprocity, negative reciprocity, external locus of control, internal locus of control, risk tolerance

^a Only used for random forest, gradient boosting, and neural net.

^b Only used for logistic regression.

Chapter 5

General Conclusion

Wage inequality between women and men continues to be one of the most intensively studied phenomenon in the economic literature. Numerous previous studies have emphasized the role of different explanatory factors that are helpful for the understanding of the persistent gender wage gap. Nevertheless, as the considered factors fail to explain the entire wage differential most studies find a significant unexplained gender gap. That is, women earn on average less than men with the same observed characteristics. Hence, the sources of the gender wage gap are still not fully understood and there remain open research questions. This dissertation focuses on the role of expectations and aspirations of women and men in this context. In particular, it analyzes gender gaps in wage expectations of students and reservation wages of job seekers. Understanding gender differences in expectations and aspirations about wages is important as they may transmit to gender differences in actual pay. These differences may partly be due to the fact that women tend to be less confident in their abilities and less optimistic about future outcomes. Hence, this dissertation additionally identifies gender differences in beliefs about labor market prospects such as the earnings potential relative to others, average wage levels, job insecurity, and job finding chances. Moreover, it highlights their effect on gender gaps in wage expectations, reservation wages, and wages.

Chapter 2 evaluates wage expectations of prospective university students in Germany. We estimate gender gaps across the distribution of students' expectations about their own starting salary as well as about the average starting salary of graduates in their study field.

Our findings indicate a significant gender gap in expected own salaries, which is similar in magnitude to the gender gap in realized starting salaries of college graduates. The gender gap we find in the expected average salary is a bit smaller but is still substantial, which suggests that females are less optimistic about the average wage level in their field of study. In terms of the structure across the wage expectation distribution, our results point to a larger gender gap at the lower end of the distribution. Furthermore, our study contributes to the literature by showing that biased beliefs about the earnings potential relative to others and about average wage levels play an important role in explaining the gender gap in students' expectation of their own starting salary. We demonstrate that females are less likely than males to overestimate their earnings potential relative to others and to overestimate the average starting salary. Conditional on these beliefs and other observed characteristics we do not find a significant gender gap in wage expectations.

Chapter 3 provides evidence on the gender difference in reservation wages of German job seekers. The study compares hourly reservation wages between women and men and estimates the unexplained gender gap using a variety of estimation approaches. We see that female job seekers in our sample set lower reservation wages than their male counterparts on average. The gender gap widens even further once we control for a large set of observed characteristics of individuals. All of our estimates of the unexplained gender gap are significantly in favor of men and are larger than the raw gender gap. Hence, our results indicate that the reservation wages of women are lower than those of comparable men. In addition, it is shown that some characteristics such as receiving unemployment benefits affect reservation wages of men and women differently. What is more, our results reveal that the gender gap is larger at the upper end of the reservation wage distribution. Lastly, we emphasize that the gender gap in reservation wages is more pronounced among parents. This finding suggests that childbearing plays a key role with regard to the gender reservation wage differential.

Chapter 4 analyzes gender differences in subjective beliefs about job insecurity and job finding chances. I compare respondents' perceived probabilities of losing or finding a

job to predicted probabilities that I estimate based on a large set of predictors. Whereas workers are considerably too pessimistic with regard to their job security, job seekers are a bit too optimistic about their chances of finding a job. However, women are significantly more pessimistic on average than men with regard to both expectations. These differences in beliefs between women and men cannot be explained by observed characteristics such as personality traits and risk preferences. Moreover, I show that these expectations are relevant since they are related to wages and reservation wages of respondents. As is predicted by theoretical models, I find a negative effect of job loss expectations and wages of workers and a positive effect of job finding expectations on reservation wages of job seekers. Additionally, I highlight that the effect of job loss expectations on wages is mostly driven by workers without a collectively agreed wage. Further, my results suggest that the gender differences in expectations do positively contribute to the gender wage gap. The contribution is particularly pronounced among college graduates.

All in all, the empirical results of this dissertation emphasize that gender gaps do not only appear in labor market outcomes but already arise in individuals' beliefs and aspirations with regard to their future professional career. It is shown that female university students have considerably lower expectations about their starting salary after graduation compared to males. Likewise, women who are searching for a job set lower reservation wages than comparable men. These findings have implications for the gender wage gap, as wage expectations and reservation wages can become self-fulfilling. Hence, when tackling the persistent wage inequality between women and men the focus of policy makers should also lie on wage expectations and reservation wages. To some extent the gender difference in wage expectations and aspirations may be caused by factors that also underlie the gender wage gap, such as differences in educational and occupational choices. But, as is demonstrated in this dissertation, it is also related to women's lower confidence and optimism. Female students, for instance, exhibit a lower confidence than males about their earnings potential relative to other students in their study field. In the same way, female workers and job seekers are less confident regarding their job security and job find-

ing chances compared to their male counterparts. Beyond the lower confidence in their own labor market prospects, women are also less optimistic about average outcomes, as is underlined by the observed gender gap in expected average wages of other students. Therefore, providing information to students, job seekers, and workers about their actual labor market prospects based on their characteristics could be an effective policy tool to mitigate the gender difference in these beliefs. Such policies could correct biased beliefs and as a consequence help individuals to make more beneficial career choices. For example, students and job seekers could be supported in their career choices and during their job search by informing them in more detail about relevant wage distributions. Moreover, the provided evidence suggests that policies should be particularly aimed at the highly educated, as the role of beliefs and the gender gaps are most significant in this group. Furthermore, as an implication of the pronounced gender gap in reservation wages among parents, policy makers should also focus on the (re-)integration of mothers in the labor market.

While the results of this dissertation highlight the role of beliefs for the gender wage gap, the causes of gender differences in beliefs are still not well understood. Thus, future research should try to find more insights into why and how gender differences in beliefs about labor market prospects emerge. Likewise, more research on how differences in various beliefs affect gender gaps in labor market outcomes would be valuable, since the evidence is still scarce on that front.