DISCUSSION PAPER SERIES

DP12719
(v. 2)

PERCEIVED WAGES AND THE GENDER GAP IN STEM FIELDS

Aderonke Osikominu and Gregor Pfeifer

LABOUR ECONOMICS

ISSN 0265-8003

PERCEIVED WAGES AND THE GENDER GAP IN STEM FIELDS

Aderonke Osikominu and Gregor Pfeifer

Discussion Paper DP12719
First Published 13 February 2018
This Revision 11 September 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

PERCEIVED WAGES AND THE GENDER GAP IN STEM FIELDS

Abstract

We estimate gender differences in elicited wage expectations among German University students applying for STEM and non-STEM fields. Descriptively, women expect to earn less than men and also have lower expectations about wages of average graduates across different fields. Using a two-step estimation procedure accounting for self-selection, we find that the gender gap in own expected wages can be explained to the extent of 54-69% by wage expectations for average graduates across different fields. However, gender differences in the wage expectations for average graduates across different fields do not contribute to explaining the gender gap in the choice of STEM majors.

JEL Classification: I21, J16, J31

Keywords: Gender Gap, wage expectations, college major choice, STEM

Aderonke Osikominu - a.osikominu@uni-hohenheim.de
University of Hohenheim and CEPR

Gregor Pfeifer - g.pfeifer@uni-hohenheim.de
University of Hohenheim and University College London

Acknowledgements

We estimate gender differences in elicited wage expectations among German University students applying for STEM and non-STEM fields. Descriptively, women expect to earn less than men and also have lower expectations about wages of average graduates across different fields. Using a two-step estimation procedure accounting for self-selection, we find that the gender gap in own expected wages can be explained to the extent of 54-69% by wage expectations for average graduates across different fields. However, gender differences in the wage expectations for average graduates across different fields do not contribute to explaining the gender gap in the choice of STEM majors.

1 Introduction

Occupations related to Science, Technology, Engineering, and Mathematics (STEM) feature higher wages on average as compared with non-STEM fields; moreover, STEM employment opportunities are predicted to grow in the future (U.S. Department of Commerce, 2011; U.S. Bureau of Labor Statistics, 2014). Despite these advantages, females remain under-represented in STEM occupations as well as in college degree completion. According to Blau and Kahn (2017), one of the main causes of the gender wage gap is precisely such male-female sorting into different majors and corresponding occupations, which is why encouraging women into STEM fields has become an important policy concern.

In this paper, we examine the gender gap using elicited expectations of future salaries, focusing on differences between STEM and non-STEM majors. We use a two-step estimation procedure, which accounts for self-selection in the spirit of Heckman (1979). In a first stage, we estimate the reduced form probability of choosing a STEM major using a full set of controls, which includes wage expectations for average other students across all available fields of study. In a second stage, we estimate separate earnings regressions for prospective STEM and non-STEM students, where the expected own future salary of a student serves as dependent variable and the respective Mills ratio from the first stage controls for self-selection. Moreover, wage expectations for average other students across STEM fields serve as instruments for people choosing STEM, while wage expectations for average other students across non-STEM fields serve as instruments for people choosing non-STEM. Lastly, we decompose our estimation results and elaborate on explained and unexplained parts of the gender gap we observe.

With this research, we relate to different contributions on students' choices using elicited expectations. For instance, Wiswall and Zafar (2015) as well as Arcidiacono et al. (2012) study specific determinants of college major choice, showing that beliefs about earnings and ability, but also personal tastes, play a significant role (see also Beffy et al., 2012). Arcidiacono et al. (2017) focus on the choice of occupation using elicited beliefs from a sample of male undergraduates, and show evidence of strong sorting across occupations on expected earnings.[1] Zafar (2011) analyzes how students form expectations regarding major-specific outcomes, showing that learning plays a key role within that process. Stinebrickner and Stinebrickner (2012) as well as Stinebrickner and Stinebrickner (2014) also incorporate subjective expectations into models of choice behavior, analyzing the updating of beliefs and its influence on later outcomes, while the latter contribution also slightly discusses STEM fields. Arcidiacono et al. (2016) focus on differences in graduation rates of minorities within STEM fields, estimating a model of students' college major choice, where net returns of a science major differ across campuses and student preparation. Zafar (2013) focuses on the gender pay gap and reveals that females value non-pecuniary outcomes when deciding on a specific college major, while males are incentivized by pecuniary outcomes. Gemici and Wiswall (2014) also look at gender differences, estimating a dynamic overlapping generations model of human capital investments and labor supply, and find that changes in skill prices, higher schooling costs, and gender-specific changes in home value were important to long-term trends.[2]

To the best of our knowledge, our paper is first to aim at combining these research streams by looking at gender gaps using elicited expectations at the time of college major choice with a particular focus on STEM fields. Moreover, while almost all related research exclusively focuses on the U.S.,[3] we look at students in Germany, a country that basically offers tuition-free education utilized by roughly 2.8 million university students of which about $500,000$ are freshmen. Finally, using data on slightly more than $2,000$ students, we draw on a considerably larger sample than most of the relevant literature, which usually works with some hundreds of observations.

Based on the first step of our estimation procedure, the reduced form probit of major

[1] Wiswall and Zafar (2018) estimate preferences for workplace attributes and relate them to college major choices and to actual job choices.

[2] Wiswall and Zafar (2018) state that gender differences in job preferences explain at least a quarter of the early career gender wage gap.

[3] An exception is Beffy et al. (2012), who concentrate on France.

choice shows that being female as compared with being male reduces the probability of studying a STEM major by more than 50% ceteris paribus. Moreover, an increase in expected wages for average graduates from STEM-fields by ten percent raises the probability to choose a STEM major by about four percent in relative terms. Higher expected wages for average graduates from non-STEM fields either decrease the probability of studying STEM or do not influence the decision at all.

Secondly, our wage regressions show that a ten percent increase in the expected wage of an average graduate from Mathematics and Computer Science (Natural Sciences) increases the own expected salary of a prospective STEM student by 6.1% (3.4%). Similarly, a ten percent increase in the expected wage of an average graduate from Law or Business Administration raises the own expected salary of a prospective non-STEM student by 2.1 to 2.5%. Expectations about salaries of average graduates in different fields, that capture how an average graduate in the respective field is rewarded on the labor market, serve as our exclusion restrictions: the two sets of instruments are each statistically significant at a 1% level. Results on self-selection imply that prospective students who choose STEM expect to earn more than average in both STEM and non-STEM fields, but have a comparative advantage in STEM. In contrast, those who choose non-STEM have expected future earnings below average in both STEM and non-STEM fields, but still higher in non-STEM compared with STEM.

Finally, our decomposition results suggest that gender differences in *wage expectations* are important. The total gender gap in own expected wages among prospective STEM students is with 36% more than twice as large than the gender gap of 14% among prospective non-STEM students. Of these total gaps, a substantial fraction can be explained by gender differences in the means of covariates: 54% and 69% for STEM and non-STEM, respectively. We can show that gender differences in the expectations about salaries of average graduates in different fields account almost entirely for this result. This suggests that women are not simply much less confident than men concerning the expectations about their own future remuneration. Rather, women's different perceptions about how average graduates across different fields are rewarded on the labor market lead them to expect lower future earnings for themselves than men. Contrarily, gender differences in wage expectations do not seem to explain the different *preferences for college majors* of men and women. This picture is consistent with a situation in which the major choice of women—unlike that of men—is more guided by non-monetary incentives rather than pecuniary rewards.

The remainder of the paper proceeds as follows: Section 2 discusses the data and shows first descriptive evidence. Section 3 explains our estimation strategy, and Section 4 shows corresponding results. Section 5 concludes.

2 Data

During the application processes in 2011 and 2012 at Saarland University, Germany, we surveyed students on their beliefs about future starting salaries. The survey's URL was e-mailed to all prospective students applying for admission, while only students submitting a complete application were considered. In 2011, 500 students completed the survey; in 2012, 1,561 students responded. Part of that increase is due to the fact that we were able to add two more majors (Education and Medicine).[4] More detailed information regarding the specific sample can be found in Klößner and Pfeifer (2018). There, the authors discuss, i.a., how the University and its students, their parents, the region/state, the curriculum, etc., compare to national and international figures and standards, making sure that we can indeed infer externally valid results from our data.

The survey started with questions regarding the prospective students' field of study. It was asked for which degree (Bachelor, Master, State Examination) and for which field of study the student had currently applied for. Students had to state whether they would aspire to obtain an additional degree afterwards (Master, Second State Examination, or a Doctoral Degree), and with which of those degrees they would aim to earn their first

[4] In 2011, this was not possible due to administrative reasons.

salary. About a quarter of the respondents seek to study either a business related field or Medicine, respectively, followed by Law Studies, Natural Sciences, Humanities, Education, and Mathematics/Computer Sciences. Under STEM, we subsume the fields Mathematics and Computer Science as well as Natural Sciences. The remainder comprises the group of non-STEM.

In the second part of the survey, students had to answer two different types of questions about monthly gross salaries. The first one asked students about expectations of their own monthly salary and about their estimates for average others who studied within the same field. The second type of questions asked students about their estimates of average monthly gross salaries for other students in different fields of study.

In the third part of the survey, students had to provide information on their personal and family background. The following characteristics were considered: gender; age; work experience; final grade in secondary school; whether the student's mother or father has a tertiary education degree; whether the student intends to live at her parents' home while studying; whether the student expects to receive funding from the public student loan scheme "BAfoeG";[5] the school system in secondary school; the federal state in which the student obtained her higher education entrance qualification; the importance of an above-average salary; the influence of income expectations on the student's major choice; the student's favorite branch of business and her work experience in this branch.

Table 1 shows relevant summary statistics. While we see 20.4% and 11.2% of survey respondents applying for STEM fields being male and female, respectively, we also observe distinct differences regarding wage expectations by gender: men expect to receive higher starting salaries than women, with the difference being 30.6 log points among prospective STEM students and 13.1 log points among non-STEM. Also, women expect field related starting salaries of average graduates to be lower than men, but the gender differences are less

[5]The German *Bundesausbildungsfoerderungsgesetz*, short BAfoeG, is a federal law on support in education, providing students from a weaker financial background with funding—specifically, with affordable student loans.

pronounced than for own starting salaries. In contrast to the patterns for the levels, women and men rank the different fields in the same way. Both groups expect an average graduate from Medicine to receive the highest salary and an average graduate from Humanities the lowest. All in all, these patterns raise the question whether and to what extent gender differences in the propensity to study a STEM major as well as in expected own future salaries could be driven by the fact that young men and women have different expectations about the average remunerations of different fields of study.

Table 1: Descriptive Statistics

	Males (1)	Females (2)	t-Statistic (3)
Expectations about own starting salary			
Applicants to STEM fields	8.189 (0.523)	7.883 (0.579)	10.611
Applicants to non-STEM fields	8.034 (0.515)	7.903 (0.519)	4.788
Share applying to STEM fields	0.204 (0.403)	0.112 (0.316)	4.722
Expectations about salary of average graduate by field of study			
Natural Sciences	7.999 (0.488)	7.934 (0.523)	2.435
Math/Computer Sciences	8.065 (0.486)	8.002 (0.527)	2.377
Humanities	7.749 (0.489)	7.616 (0.535)	4.925
Business Administration	7.998 (0.473)	7.885 (0.519)	4.327
Law	8.103 (0.486)	8.036 (0.527)	2.550
Education	7.901 (0.453)	7.819 (0.497)	3.330
Medicine	8.171 (0.525)	8.086 (0.533)	3.052
Observations	632	836	

Source: Applicant Survey, own calculations. *Note:* The first two columns show the means and standard deviations (in parentheses) of the variables, the last column the t-statistics of a test of equality of means. Wages are in logs.

3 Methodology

Let $\mathbf{x}_{i1}$ be a row vector of variables that influence the return to STEM fields, whereas $\mathbf{x}_{i0}$ denotes a row vector of variables that influence the return to non-STEM fields. We refer to the combined regressor vector as $\mathbf{x}_i$, where $\mathbf{x}_{i1} \subset \mathbf{x}_i$ and $\mathbf{x}_{i0} \subset \mathbf{x}_i$. Besides controls on the student's background, the importance of salary as a job value, and the highest intended degree, we consider expected salaries for average graduates in *all* other fields, i.e., Business, Medicine, Law Studies, Natural Sciences, Humanities, Education, and Mathematics/Computer Science.

As a first stage, we estimate a reduced form model for major choice

$$\Pr\{s_i = 1 \,|\, \mathbf{x}_i\} = \Pr\{\varepsilon_i \geq -\mathbf{x}_i\boldsymbol{\pi} \,|\, \mathbf{x}_i\} = \Phi\left(\mathbf{x}_i\boldsymbol{\pi}\right), \tag{1}$$

where s_i is an indicator function that takes on value one when utility from choosing a STEM field is at least as high as utility when choosing a non-STEM field. $\boldsymbol{\pi}$ is a column vector of coefficients. ε_i denotes the error term. For practical purposes, we assume the error term to be standard normally distributed, i.e. $\varepsilon_i \,|\, \mathbf{x}_i \sim \mathcal{N}(0,1)$, although this assumption as well as linearity of the index are not needed for identification since we rely on exclusion restrictions (French and Taber, 2011). The function Φ thus denotes the standard normal c.d.f.; its p.d.f. is denoted by φ. More precisely, we estimate this first stage pooling across gender and including a female dummy.

As a second stage, we estimate wage regressions

$$\ln y_{i1} = \beta_{01} + \mathbf{x}_{i1}\boldsymbol{\beta}_1 + \gamma_1\lambda_{i1} + \eta_{i1}\,, \tag{2}$$

$$\ln y_{i0} = \beta_{00} + \mathbf{x}_{i0}\boldsymbol{\beta}_0 + \gamma_0\lambda_{i0} + \eta_{i0}\,, \tag{3}$$

where $\ln y_{ij}$ denotes the log wage rate a prospective student in field j, $j \in \{0,1\}$, expects to earn after graduation and η_{ij} an error term. The terms

$$\lambda_{i1} \equiv \frac{\varphi(\mathbf{x}_i\boldsymbol{\pi})}{\Phi(\mathbf{x}_i\boldsymbol{\pi})}\,, \qquad \lambda_{i0} \equiv -\frac{\varphi(\mathbf{x}_i\boldsymbol{\pi})}{1-\Phi(\mathbf{x}_i\boldsymbol{\pi})}\,, \tag{4}$$

denote inverse Mills ratios that account for self-selection into STEM and non-STEM fields

(Heckman, 1979). If the estimate of γ_1 is positive and γ_0 is negative, this is evidence for positive selection bias in the sense that mean earnings associated with a major are higher in the subgroup of people who have chosen that major than in the total population. In line with the classical Roy (1951) model, we impose exclusion restrictions on the second stage earnings models. Expectations about the remuneration of an average graduate from the STEM fields Mathematics/Computer Science and Natural Sciences should only affect the own expected wage of prospective students in STEM fields. The reverse reasoning applies to non-STEM fields. Therefore, in the earnings equation for STEM, we consider wage expectations for average graduates in Mathematics/Computer Sciences and Natural Sciences. Analogously, for non-STEM, we use wage expectations for average students in Business, Medicine, Law Studies, Humanities, and Education. The exclusion restrictions ensure that our estimation results are not driven by distributional assumptions on the error term of the reduced form choice model because the index function of the inverse Mills ratio includes regressors that do not directly enter the earnings equation (French and Taber, 2011). Both earnings equations include the additional control variables used also in the first stage choice model and the correction term for self-selection based on the respective inverse Mills ratio, which we calculate from the first stage estimates. We again pool observations across gender.

Finally, we use our estimation results to decompose the gender difference in the means of the dependent variables in Equations (1) to (3) into a part that can be explained by gender differences in the covariates included in the models and an unexplained part. We further evaluate the individual contributions of different groups of covariates to the explained gender gap. For decomposing the log of the own expected salary, we use the familiar method developed by Blinder (1973) and Oaxaca (1973). For decomposing the probability to choose a STEM field based on the probit regression, we use the method proposed by Fairlie (2005), who adapts the Blinder-Oaxaca approach to the case of non-linear binary choice models.[6] For a related decomposition analysis of a model for college major choice, see Zafar (2013).

[6] The decompositions are implemented with the Stata modules Jann (2006) and Jann (2008).

4 Results

Columns (1) and (2) of Table 2 provide results from the reduced form probit: the first column shows the coefficients for the probability of studying STEM, and the second column gives corresponding average partial effects. While we include the full set of available control variables, some coefficients are noteworthy. Being female as compared with being male reduces the probability of studying a STEM major by eight percentage points, a relative decline by 53%, which is statistically significant at a 1% level. Moreover, and as one could have expected, an increase in expected wages for average graduates from the STEM-fields (i.e., Mathematics/Computer Science and Natural Sciences) by ten log points (approx. 10%) raises the probability to choose a STEM major by around 0.6 percentage points, about four percent in relative terms. Higher expected wages for average graduates from non-STEM fields either decrease the probability of studying STEM or do not influence the decision at all. Specifically, an increase in the expected wage for an average graduate from Business Administration by 10 log points decreases the probability to study STEM by 0.9 percentage points, a relative effect of six percent.

Columns (3) and (4) of Table 2 display the estimates of the selection-corrected wage equations for both STEM and non-STEM, with the log of the expected own future salary as the dependent variable. These columns answer the question how certain characteristics influence the expected own STEM/non-STEM salary—controlling for self-selection into field of study. Expectations about salaries of average graduates in different fields, that capture how an average graduate in the respective field is rewarded on the labor market, serve as our exclusion restrictions. Expectations about the wages of average graduates of STEM-fields affect the own expected wage of prospective STEM students, whereas expected wages of average graduates of non-STEM fields influence only the own expected wage of non-STEM applicants. All of these coefficients, which can be interpreted as elasticities, show the expected positive sign and are (highly) statistically significant. In particular, a ten percent increase in the expected wage of an average graduate from Mathematics and Computer

Table 2: Estimation Results

	Major Choice		Log Earnings	
	Coefficient	APE	STEM	Non-STEM
	(1)	(2)	(3)	(4)
Expectations about salary of average graduate in different fields (in logs)				
Natural Sciences	0.343	0.064	0.339	
	(0.194)*	(0.035)*	(0.122)***	
Math/Computer Sciences	0.296	0.055	0.611	
	(0.176)*	(0.032)*	(0.100)***	
Humanities	-0.279	-0.052		0.222
	(0.154)*	(0.028)*		(0.047)***
Business Administration	-0.470	-0.088		0.247
	(0.185)**	(0.034)***		(0.050)***
Law	0.245	0.046		0.209
	(0.170)	(0.031)		(0.057)***
Education	-0.036	-0.007		0.066
	(0.153)	(0.028)		(0.039)*
Medicine	0.071	0.013		0.138
	(0.153)	(0.028)		(0.039)***
Selected other variables				
Female dummy	-0.421	-0.081	-0.272	-0.069
	(0.088)***	(0.017)***	(0.078)***	(0.026)***
Correction term for sample selection			0.415	0.251
			(0.204)**	(0.149)*

Source: Applicant Survey, own calculations. *Note:* APE denotes average partial effect. All models include additional controls for the importance of salary as a job value (4 dummies), the degree with which one intends to leave university (3 dummies), background characteristics (5 variables) and a dummy for survey cohort as well as an intercept. Bootstrapped standard errors are shown in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

Science (Natural Sciences) increases the own expected salary of a prospective STEM student by 6.1% (3.4%). Similarly, a ten percent increase in the expected wage of an average graduate from Law or Business Administration raises the own expected salary of a prospective non-STEM student by 2.1 to 2.5%. Moreover, Table 3 shows that the sets of instruments are also jointly statistically significant, yielding Wald test p-values of virtually zero. It also provides an analogous joint test for the remaining covariates (excluding the correction term for self-selection). Going back to Table 2, the coefficient on the inverse Mills ratio at the very end of Column (3), being 0.415, suggests significantly positive self-selection of STEM applicants compared with the total population. In contrast, prospective students of non-STEM fields tend to be negatively selected (coefficient of 0.251 at the end of Column (4)), but the selection correction term is not significant at the 5% level.[7] Taken together, the results on self-selection imply that prospective students who choose STEM expect to earn more than average in both STEM and non-STEM fields, but have a comparative advantage in STEM. In contrast, those who choose non-STEM have expected future earnings below average in both STEM and non-STEM fields, but still higher in non-STEM compared with STEM. Interestingly, the patterns of self-selection differ somewhat when we estimate separate earnings regressions by gender (full results available on request). While, for men, we observe the same selectivity pattern as in the pooled estimation, prospective female students in both STEM and non-STEM fields tend to be positively self-selected compared with the total female population. In the estimations by gender, the selection correction terms are not all statistically significant, though.

Table 4 shows the decomposition results. Column (1) decomposes the estimation results of the reduced form regression of major choice displayed in Column (1) of Table 2 based on the method by Fairlie (2005). Columns (2) and (3) of Table 4 show the analogous decomposition results based on the earnings regressions in Columns (2) and (3) of Table 2. These decompositions are based on the method by Blinder (1973) and Oaxaca (1973). Focusing

[7] Recall from Equation (4), that a positive γ_1 and a negative γ_0 corresponds to positive self-selection into STEM and non-STEM fields, respectively.

Table 3: Wald Tests of Joint Significance

	Major Choice (1)	Log STEM Earnings (2)	Log Non-STEM Earnings (3)
Expectations about average salaries			
χ^2-statistic	23.785	278.355	1315.126
degr. of freedom	7	2	5
p-value	0.001	0.000	0.000
Remaining variables, except selection correction term			
χ^2-statistic	179.817	37.673	113.133
degr. of freedom	16	16	16
p-value	0.000	0.002	0.000

Source: Applicant Survey, own calculations. *Note:* Variance matrices for test statistics are bootstrapped.

on the global decomposition results for expected own earnings (Columns (2) to (3) in the top panel of Table 4), we see that the total gender gap among prospective STEM students is with 31 log points (36%) more than twice as large than the gender gap of 13 log points (14%) among prospective non-STEM students. Of these total gaps, a substantial fraction can be explained by gender differences in the means of covariates: 54% for STEM and 69% for non-STEM. The detailed decomposition results in the lower panel of Table 4 reveal that gender differences in the expectations about salaries of average graduates in their field account almost entirely for this result. For STEM, gender differences in expected salaries of average graduates contribute 16 log points to explaining the total gender gap in own expected salaries of 31 log points. For non-STEM, the corresponding figures are eight log points out of 13. These patterns suggest that women are not simply much less confident than men concerning their own future remuneration on the labor market. Rather, they tend to have different perceptions than men about how different fields of study are rewarded on the labor market, which in turn explains their different expectations about the own future salary.

The decomposition results for own expected salaries stand in sharp contrast with those

Table 4: Decomposition Results

	Major Choice	Log STEM Earnings	Log Non-STEM Earnings
	(1)	(2)	(3)
Global results			
Mean of males	0.204	8.189	8.034
	(0.016)***	(0.047)***	(0.022)***
Mean of females	0.112	7.883	7.903
	(0.010)***	(0.057)***	(0.019)***
Gender gap	0.092	0.306	0.130
	(0.018)***	(0.075)***	(0.030)***
Explained gap	0.007	0.164	0.089
	(0.011)	(0.064)**	(0.024)***
Unexplained gap	0.084	0.143	0.041
	(0.017)***	(0.046)***	(0.019)**
Contributions of covariates to the explained gap			
Expectations about average salaries	-0.009	0.158	0.076
	(0.005)*	(0.064)**	(0.023)***
Importance of salary	-0.003	-0.001	0.004
	(0.003)	(0.012)	(0.004)
Intended degree	0.008	0.012	0.005
	(0.005)	(0.026)	(0.005)
Background characteristics	-0.000	0.038	-0.004
	(0.007)	(0.026)	(0.006)
Survey cohort	0.012	-0.043	0.008
	(0.004)***	(0.034)	(0.006)

Source: Applicant Survey, own calculations. *Note:* The decomposition results in Column (1) are based on the method by Fairlie (2005), the decomposition results in Columns (2) and (3) are based on the method developed in Blinder (1973) and Oaxaca (1973). All decompositions use as basis the coefficients of a pooled model that includes a female dummy (see Table 2). Bootstrapped standard errors are shown in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

for major choice (Column (1) in the top panel of Table 4), which suggest that the total gender gap of nine percentage points cannot be explained by gender differences in the means of the covariates considered. This is also reflected in the detailed results, which show that gender differences in wage expectations cannot account for the gender gap in major choice, with a coefficient close to zero (lower panel of Table 4). Overall, these results indicate that, while gender differences in wage expectations are important, they do not seem to explain the different preferences for college majors of men and women. Separate estimations of the choice model for men and women suggest that wage expectations do not affect the major choice of women in a statistically significant way, but only of men (full results available on request). Women seem to be more guided by non-monetary incentives (personal tastes, psychic costs) than by their assessment of how majors are rewarded on the labor market. Thus, their differential wage expectations compared with men cannot explain why women are less inclined to study a STEM field than men. The fact that pecuniary rewards of major choice seem to be more important to men than to women has also been documented in Zafar (2013).

5 Conclusion

We examined gender differences among prospective students with respect to their wage expectations and the probability to choose a STEM major. Descriptively, we see that women in fact have different expectations about wages than men. For their own future salary, they expect less than men, which might be explained by looking at their expectations about wages of average graduates across different fields of study: here, women also have consistently lower expectations than men.

We estimated a binary college major choice model using the wage expectations for an average graduate across different fields and additional student characteristics as controls. Corresponding probit estimates suggest that women are ceteris paribus 50% less likely to choose STEM than men. Moreover, an increase in expected wages for average graduates from STEM fields by ten percent raises the probability to choose a STEM major by about four percent. Higher expected wages for average graduates from non-STEM fields either decrease the probability of choosing STEM or do not influence the decision at all.

In a second step, we regressed the own expected salary on the expected wages for an average graduate from the relevant fields, a correction term for self-selection into field of study, and additional student characteristics. The regression of own expected earnings among STEM applicants shows that an increase in the expected wage of an average graduate from Mathematics/Computer Science or Natural Sciences by ten percent increases the own expected salary by 3% to 6%. Analogously, an increase in the expected wages of average graduates from other fields raises the own expected salary of a prospective non-STEM student. Moreover, our findings suggest significantly positive self-selection of STEM applicants compared with the total population.

Finally, our decomposition results indicate that gender differences in wage expectations for average graduates are important for explaining gender differences in own expected salaries of both STEM and non-STEM students. This suggests that women are not simply much less confident than men about their own remuneration but that they have different perceptions about how different fields of study are rewarded on the labor market, which, in turn, affects what they expect to earn for themselves. In contrast, gender differences in wage expectations do not seem to explain the different propensities of men and women to study STEM. The latter finding is in line with the literature on college major choice, showing that for men future earnings are more important than for women, for whom non-monetary incentives play a bigger role (Zafar, 2013). With our data, we cannot directly account for differences in preferences related to non-monetary aspects, which makes it plausible that the gender gap in major choice remains largely unexplained.

References

Arcidiacono, P., Aucejo, E. M., and Hotz, V. J. (2016). University differences in the graduation of minorities in stem fields: Evidence from california. *American Economic Review*, 106(3):525–562.

Arcidiacono, P., Hotz, V. J., and Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, 166(1):3–16.

Arcidiacono, P., Hotz, V. J., Maurel, A., and Romano, T. (2017). Recovering ex ante returns and preferences for occupations using subjective expectations data. Working Paper 20626, National Bureau of Economic Research.

Beffy, M., Fougère, D., and Maurel, A. (2012). Choosing the field of study in postsecondary education: Do expected earnings matter? *The Review of Economics and Statistics*, 94(1):334–347.

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.

Fairlie, R. W. (2005). An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models. *Journal of Economic and Social Measurement*, 30(4):305–316.

French, E. and Taber, C. (2011). Identification of models of the labor market. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics, Vol. 4A, ch. 6*, pages 537–617. Elsevier Science, Amsterdam.

Gemici, A. and Wiswall, M. (2014). Evolution of gender differences in post-secondary human capital investments: College majors. *International Economic Review*, 55(1):23–56.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Jann, B. (2006). fairlie: Stata module to generate nonlinear decomposition of binary outcome differentials. Technical report, mimeo, available from http://ideas.repec.org/c/boc/bocode/s456727.html.

Jann, B. (2008). The blinder-oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4):453–479.

Klößner, S. and Pfeifer, G. (2018). The importance of tax adjustments when evaluating wage expectations. *The Scandinavian Journal of Economics*, forthcoming.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3):693–709.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146.

Stinebrickner, R. and Stinebrickner, T. R. (2014). A major in science? initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies*, 81(1):426–472.

Stinebrickner, T. and Stinebrickner, R. (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics*, 30(4):707–748.

U.S. Bureau of Labor Statistics (2014). STEM 101: Intro to tomorrow's jobs. Occupational Statistics and Employment Projections.

U.S. Department of Commerce (2011). STEM: Good Jobs Now and For the Future. Economics & Statistics Administration.

Wiswall, M. and Zafar, B. (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2):791–824.

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. (2011). How Do College Students Form Expectations? *Journal of Labor Economics*, 29(2):301–348.

Zafar, B. (2013). College major choice and the gender gap. *Journal of Human Resources*, 48(3):545–595.