GENDER GAP IN EXPECTED EARNINGS AND THE ROLE OF INFORMATION: EVIDENCE FROM GEORGIA

By

Ekaterine Khvitia

Submitted to

Central European University

Department of Economics and Business

In partial fulfillment of the requirements for the degree of Master of Arts in Economics

Supervisor: Andrea Weber

Budapest, Hungary and Vienna, Austria

2021

Abstract

This is not a subject of dispute that there exists gender pay gap. One of the determinants of the realized wages is considered to be investments to the education. This paper analyses gender gap in expected earnings of students in Georgia. The unique data set makes it possible to examine whether students had biased beliefs and how the information provision may contribute to the gender gap. Using Oaxaca-Blinder decomposition method thesis finds that highest expected earnings differentials are among those students who chose specializations of exact and natural sciences, art and humanities, and law. The largest part of the explained part of the gender gap is due to the school fixed effects. And the most important contribution of the paper is that it finds the significance and the positive effect of the information provision on the gender gap.

Acknowledgments

Hereby, I would like to thank Professor Andrea Weber for being my supervisor. Her valuable comments and constant feedbacks strongly contributed to my paper. I also would like to express gratitude to Gega Todua and Nikoloz Kudashvili who allowed me to use their data for my thesis, and for giving me thoughtful comments.

Table of Contents

1 Introduction

In the recent decades females have started participating more actively in the labor force and continue to fight for the equal payments to males. Globally, in 2020 women tend to earn 68.6% of the male earnings (World Economic Forum, 2021). Number of studies have searched for the explanations of existing gender wage gaps, finding that gender-specific differences in occupational aspirations and job-related skills (Marini & Fan, 1997), race (Ifatunji & Harnois, 2015), qualifications (Blau & Kahn, 2012), working experience (Dias, Joyce, & Parodi, 2020) are considered the most important factors in explaining the wage differentials. Furthermore, realized wages can be partly considered as returns to the investments in education (Filippin & Ichino, 2005). Therefore, one of the most crucial decisions for students is when they consider whether invest in their further education or not. This mostly holds as long as the expected returns from the investment exceeds the costs (Cohn & Hughes Jr., 1994). But how does the expected earnings among students differ by gender? Cohn and Hughes claim that both female and male students on average expect higher benefits from education in comparison with costs, however, males expected by $16,000 higher returns (Cohn & Hughes Jr., 1994). Students' decision are influenced by subjective beliefs (Wiswall & Zafar, 2015) and the awareness about the actual average earnings in the developing countries can be even more insufficient (Kudashvili & Todua, 2021). For instance, Georgia, where the women to men ratio of salaries has persistently been at around 64% throughout the last years. Moreover, the labor force participation among men is 61.8%, while for women

43.1% as of 2019[1] . Such low participation rate of females is not really surprising due to widely dominated patriarchal norms that diminishes women's economic role (Japaridze, 2012).

This thesis investigates the gender gap in the expected earnings among students in Georgia. For this purpose, I am using the data collected from the experimental survey conducted in two rounds by (Kudashvili & Todua, 2021) in Georgia. The population consists of tenth and eleventh grade[2] students from schools in the capital city of Georgia – Tbilisi. According to the survey setting, part of the schools (treated) were provided with information regarding field-specific actual average earnings and unemployment rates. Authors find that the students hold inaccurate beliefs regarding actual average earnings (Kudashvili & Todua, 2021). The unique data set gives me an opportunity to study whether the provision of information had an additional effect in the explanation of the gender gap in expected earnings. The expected earnings differentials can be observed at different stages of a lifetime: after university graduation, at the age of 30, and at the age of 45.

This paper contributes to the existing literature in the way of studying the gender gap in expected earnings through another channel – lack of information.

The aim of this paper is to answer following research questions: (i) Does the provided information contribute to shaping students' expected earnings? (ii) Does the information provision affect the gender gap in the reported expected earnings?

[1] Statistics taken from the National Statistics Office of Georgia.

[2] General education consists of twelve grades in Georgia.

For studying the gender gaps I use the Oaxaca-Blinder decomposition method ((Blinder, 1973), (Oaxaca, 1973)), which reports the gender gaps and helps to determine to what extent the independent variables contribute to the explanation of the existing gender gaps. This empirical strategy has been widely and successfully used for studying the gender pay gaps, differentials in earnings by races ((Blau & Beller, 1988), (Kim, 2010), (Fortin, Bell, & Böhm, 2017)). This method compares the means of the outcome variables. Due to such limitation, similar to (Fortin, Bell, & Böhm, 2017), I decided to divide population into centiles based on the reported expected earnings and observe gender gaps separately on those centiles as well. Moreover, for testing the significance of the treatment effect on the expected earnings I use the linear probability model, similar to (Kudashvili & Todua, 2021).

I find that there exist gender gaps in favor for males in both rounds at all stages of the lifetime. These gaps are different based on the chosen specializations. Largest gaps are reported for the exact and natural sciences, art and humanities, and law. Moreover, the earning differentials are larger in the upper centiles in comparison with the lower bound. Furthermore, the provision of information is an important factor in the change of the preferred specializations and expected earnings in both rounds. There also exists heterogeneity in treatment effect by gender for the case of expected earnings at the age of 30, reported in the second round. This fact can be observed from the O-B decomposition estimations results as well. Such result has been mainly driven by the upper centile.

The structure of the thesis following. Chapter 2 presents the literature review. In Chapter 3 the used data set is introduced and described. In Chapter 4 the used methodologies for answering the research questions are explained. Chapter 5 reports the estimation results. The thesis ends with the concluding part, where additionally I remark the limitations and further investigation possibilities.

2 Literature Review

There is a rich literature studying gender gaps in various economies. (Wellington, 1993) using questionnaires of the Panel Study of Income Dynamics conducted in 1976 and 1985, observes wage gap between white men and women. Author claims that the difference in wages has narrowed by approximately 5% over these 9 years and half of this change was due to the changes in the characteristics, more specifically, women acquired some additional trainings. These conclusions complement with (Blau & Beller, 1988) findings regarding reduction of 3% in the wage gap due to the mean levels of experience. It is notable that authors were able include potential experience in their paper, by which they might have been underestimating the actual reduction. Moreover, (Wellington, 1993) finds that men tend to receive higher earnings due to the fact that they were choosing fields where they could gain higher rewards.

One of the exceptions that used other countries than the United States as population is (Fortin, Bell, & Böhm, 2017) studying European countries. Unlike the classical approach, in this paper authors studied the gender gaps in the top jobs as well by dividing the selection into four top centiles. They have found decreasing female shares as moving up across the centiles. (Blau & Kahn, 2003) compare wage differentials between 22 countries and find one of the largest gaps of 36.4% in the Eastern European countries, including post-soviet Russia. This corresponds to the National Statistics Office of Georgia statistics stating that there exists the gender pay equal to 37.2% in 2017, despite the fact that females possessed better professional characteristics, such as, education and work experience.[3]

[3] Monthly salary statistics are taken from the National Statistics Office of Georgia.

The abovementioned papers were studying gender gaps in actual earnings, however, expectations in youth might be different across males and females as well. In fact, the beliefs regarding earnings for students might be one of the main factors in making decision regarding their major (Beffy, Fougere, & Maurel , 2012). Moreover, (Attanasio & Kaufmann, 2009) state in their paper that expected earnings are important determinants for the college attendance decision. (Filippin & Ichino, 2005) find that females expect higher wage differentials than males and that females think that this gap can be explained by the differences in characteristics across genders and due to the employer's discriminatory tastes as well. Authors claim that this kind of beliefs lead to less investments from the side of females in their education that in turn leads to such gaps. (Blau & Ferber, 1991)claim that according to the students' suggestion there is no gender gap in the starting salaries, however, female students expect lower earnings in the subsequent years in comparison with males. (Wiswall & Zafar, 2015) are constructing the perceived gender gap in earnings based on the reported population earnings for both males and females, separately for college majors. Authors find that males expect wage differential in their favor varying from -3.23% to -7.41% depending on the college major, while females report even larger gaps also in favor of men. Moreover, females expect the highest paying jobs in the field of economics or business, engineering and computer science, and the lowest earnings in - art and humanities.

Such gaps in expected earnings could be the result of the lack of information among students. Several economists have studied this issue with experimental approach. (Zafar, 2013), and (Wiswall & Zafar, 2015) have provided students with information regarding actual salaries, finding that students update their expected earnings after information prevision. Similarly, (Kudashvili & Todua, 2021) implemented experiment for testing the biasedness of beliefs. The abovementioned papers found that the expected earnings play a significant role in the

determination of major field. The effect of the information provision has been estimated with linear probability and probit models.

The strategies used for the estimation of the gender pay gap are very similar in the literature, mostly using the Oaxaca-Blinder decomposition method for explaining the wage differentials across genders. (Blinder, 1973) and (Oaxaca, 1973) have developed this methodology for studying wage discrimination. (Wellington, 1993) uses O-B decomposition, a slightly modified version of Oaxaca-Blinder decomposition on changes implemented by (Smith & Welch, 1989). Specifically, as already mentioned, (Wellington, 1993) observes the changes in the wage gaps across genders between 1976 and 1985 years by answering the question "If the returns to the independent variables were constant at their 1985 levels, what portion of the wage gap can be accounted for by changes in the means?" Classic O-B decomposition has been employed by (Khitarishvili, 2019) for explaining the gender wage gap between 2000 and 2004 in Georgia, finding that the largest part of the wage differential remains unexplained. (Fortin, Bell, & Böhm, 2017) used classic O-B decomposition, however, in the paper, the selection was divided into centiles. With such strategy authors were able to study the gender wage gap in top jobs across time periods. Moreover, (Blau & Beller, 1988) estimated wage equations separately for each race-gender groups and year. In particular, authors reported the O-B decomposition by gender results separately for each race.

This paper serves as a combination of the explanation of the gender gap in expected earnings among students and the estimation of effectiveness of the information provision in the explanation of the earnings differential.

3 Data

In this chapter I introduce the data used for the analysis and relevant descriptive statistics.

3.1 Data

For my analysis I used the experimental survey conducted by Kudashvili and Todua in Georgia. The data was collected from the students, aged between 15 and 17 (10^{th} and 11^{th} graders[4]), of 22 schools of Tbilisi – the capital city of Georgia, in 2017. The experiment was organized in two rounds. The first round was held in April, 2017 during the regular school hours and students were asked to provide information regarding their characteristics such as gender, number of siblings, parental occupations, whether they have a laptop/personal computer or not, if they are planning to hire private tutor for the national exams[5], their evaluation of their own performance on school and national level, also, for testing the biasedness, beliefs regarding unemployment rates and incomes in different fields, moreover, students indicated their preferred specializations[6] and expected future incomes at different stages of lifetime – right after graduation of the university, at the age of 30 and at the age of 45. In order to incentivize honesty from the students' side, they were promised to receive specific information, such as, university application procedures, deadlines, admission

[4] In order to enter university in Georgia, it is compulsory to have finished 12 years (grades) of school education.

[5] This is a national phenomenon - for the national exams, which is mandatory for students intending to enter university, almost every student decides to hire private tutors for specialization specific subjects. Tutoring is considered to be increasing the chances to be admitted at the top universities and receiving government scholarships based on the scores gained in the national exams. 78.3% of the surveyed students indicated such intention (Kudashvili & Todua, 2021)

[6] In the first place students were asked whether they were going to enter university and at around 95% had positive answer. Such result is mainly due to the fact that the survey was conducted only in the capital city and, what is more important, earnings higher education diploma is a cultural phenomenon (common for the Post-Soviet countries). Thus, the difference in the expected earnings will not be due to the fact whether students choose to continue to higher education or not.

requirements and competition by email based on their occupation preferences (Kudashvili & Todua, 2021). The process was monitored to eliminate the communication between students.

Schools were randomly divided into control (C) and treated (T) schools. Right after handling in the filled surveys at the end of the first round, each student in the T group was given statistics[7] about the unemployment rates and average earnings, while control schools did not receive any kind of information and explanations of difference in wages and employment chances for each major. These statistics are provided in the column 1 and 4 of the Table 1 and the grouped majors in the Appendix Table 1.

Second round of the survey was conducted one month later in May 2017. At this stage students from both C and T groups were asked to revisit their rankings of preferred specialization and expected earnings right after graduation of university, at the age of 30 and at the age of 45, also, they were asked whether they have discussed the survey with their parents.

This treatment effect will be used to examine the role direct information provision in gender gap in expected earnings.

I use students' indicated expected earnings as dependent variable to study the gender wage gap at all three stages of the lifetime - right after graduation of university, at the age of 30 and at the age of 45. As explanatory variables I am using the reported preferred specializations, whether a student has a PC/ Laptop, students' evaluation of their performance on the school level, which is a proxy for the ability, additionally, beliefs of average earnings and unemployment rates separately for

[7] Specialization specific and age specific earning statistics, unemployment rebates taken from the household survey conducted by the National Statistics Office of Georgia in 2015

each specialization[8] and without degree, I observe their parental occupations and number of siblings, furthermore, I include the school fixed effects, such as, number of parallel classes, number of students, number of teachers, whether a school is specialized or not.

Overall I have 1313 observations, 44% from control schools and 56% from treatment schools. Even though, as already mentioned, students were incentivized to provide correct information regarding their preferred choices, this might have not worked for the expected earnings and some students could play around and misreport.

3.2 Descriptive Statistics

I distinguish 3 types of earnings in the analysis:

- **Perceived** earnings that are beliefs about average earnings for each specialization, which are general beliefs for the whole population and are used as explanatory variables. Moreover, along with the beliefs regarding unemployment rates within each field, they are used to test the biasedness, so that the treatment effect is applicable. (see Table 1)
- **Expected** earnings at 3 stages of lifetime – after graduation of university, at the age of 30 and at the age of 45, which are student's personal expected earnings and are used as dependent variables in the analysis. These are different from the abovementioned beliefs since students take into consideration their evaluation of their own abilities (where do they consider themselves in the upper centiles or in the lower?), also, inflation rates,
- **Actual** earnings for each specialization, which are the statistics provided to the treatment schools.

[8] Specializations: Exact and Natural Sciences, Medical Sciences, Economics and Business, Social Sciences, Art and Humanities, Law.

From Figures 1-3 in Appendix, it is observable that males on average have higher expectations regarding their future incomes at the ages of 30 and 45 as expected, and both for males and females those are higher than the average actual earnings across all specializations. Gender gaps are largest in the cases of exact and natural sciences, art and humanities, and law. The smallest gaps, almost negligible, across those students who have indicated medical sciences, and economics and business as their first choice field. In the case of expected earnings after university graduation, females, who have chosen as their preferred specializations medical sciences, and economics and business, on average reported higher expectations than males in the same cluster. It is also notable that among those students who have listed exact and natural sciences, social sciences, and law as their preferred field, on average females reported expected earnings very close to the mean of actual earnings. Again, the gender gap in the expected earnings in the case of after graduation of university is among those students who have reported exact and natural sciences, art and humanities, and law as their first choice specialization.

Table 1: Actual and perceived earnings and unemployment rates

Educational Attainment	**Mean Earnings**			**Unemployment Rate**			**Predicted Earnings**		
	Actual	Belief	Bias	Actual	Belief	Bias	Actual	Belief	Bias
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
No University Education	504	381	-25%	8.3%	46%	450%	462	206	-55%
Tertiary Education	802	1,280	60%	13.1%	29%	123%	697	921	31%
Exact and Natural Sciences	771	940	22%	12.6%	30%	137%	673	660	-2%
Medical Sciences	673	1,432	113%	10%	25%	149%	606	1,078	78%
Economics and Business	890	1,696	91%	19.2%	28%	43%	719	1,229	71%
Social Sciences	872	1,213	39%	13.3%	30%	125%	756	849	12%
Art and Humanities	654	843	29%	8.5%	34%	303%	599	554	-7%

Law	953	1,555	63%	15.1%	29%	92%	809	1,104	36%

Notes. Columns (1)-(2) report the actual and perceived mean monthly earnings in Georgia. Columns (4)-(5) report the actual and perceived unemployment rates. Columns (7)-(8) report the predicted monthly earnings calculated as the product of mean perceived monthly earnings and employment rates. Employment rates are calculated as one minus the unemployment rate. Both actual and perceived earnings are given in GEL. GEL stands for the Georgian Lari, the average exchange rate in 2017 was approximately $1=2.4 GEL. Mean monthly earnings and unemployment rates for the individuals with tertiary education are calculated as the weighted average earnings and unemployment rates of individuals having a degree in one of the majors: exact and natural sciences, medical sciences, economics and business administration, social sciences, arts and humanities, and law. Columns (3), (6) and (9) calculate the difference between the perceived and actual figures in percentage terms. The bias is calculated as follows: $Bias = \frac{Belief - Actual}{Actual} * 100$. Actual earnings and unemployment rates are calculated using the 2015 Household Survey conducted by the National Statistics Office of Georgia. For the calculation of earnings, we considered only hired full-time employees. Unemployment rates are defined in line with the International Labor Organization (ILO) strict criteria (see page 6). https://ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/publication/wcms_675155.pdf
Source: Kudashvili and Todua, 2021 (Kudashvili & Todua, 2021)

The figures indicating bias in the columns 3, 6 and 9 of the tables help to determine whether students' perceptions regarding earnings and unemployment rates are accurate or not. Students underestimate mean earnings for individuals without a degree, while they overestimate constantly mean earnings and unemployment rates in other cases. The difference in the biases in mean earnings between individuals with and without tertiary education is 85%. For separate specializations the bias in mean earnings is largest for individuals in medical sciences (113%), followed by economics and business (91%) and law (63%). In the case of unemployment rate, the differences between actual and perceived rates are extremely large that could be explained by the fact that in Georgian public schools subject economics is not taught, therefore, students do not have a proper knowledge about economic concepts such as labor force or unemployment rate. The smallest bias has been reported in the case of economics and business.

In the table there are also reported predicted earnings, which is a product of perceived earnings and employment rates that in turn is 1 minus unemployment rates. The predicted earnings is the most accurate for individuals with degrees in exact and natural sciences (-2%) and art and humanities (-7%). Again, the largest positive bias has been reported in the field of medical sciences

(78%) and economics and business (71%), and the largest negative bias – without university degree (-55%).

Further, investigating the treatment effect on different outcomes, I observe the ratios of students who changed and did not change their preferred specialization in the second round after information provision, separately in treated and controlled schools (Table 2).

Table 2: Ratios of students who changed and did not change their first choice specializations

	Treatment	Control
Changed	25.5% (69.1%)	11.7% (30.9%)
Did not change	74.5% (53.5%)	88.3% (46.5%)

Notes: In this table is displayed the ratios of students who changed and did not change their preferred specializations between round 1 and round 2, separately for treatment and control groups.
In the normal format are indicated ratios across treatment and control groups separately, while in the parenthesis the percentages were calculated across the students who changed and did not change their first choice specializations.

In the treatment and control schools students tend to not change their first choice specializations. On the other hand, the majority of students who changed their preferred field in the second round are from the treated schools (69.1%).

Table 3: How did those students adjust their expected earnings who have changed their preferred specialization?

	Treatment						Control					
	Afteruni		30y		45y		Afteruni		30y		45y	
	M	F	M	F	M	F	M	F	M	F	M	F
Adjusted-	25%	16%	23%	14%	21%	16%	13%	16%	16%	16%	21%	5%
0	7%	11%	9%	9%	8%	12%	5%	3%	5%	5%	5%	11%
Adjusted+	20%	21%	20%	25%	23%	20%	42%	21%	39%	18%	34%	24%
Total	52%	48%	52%	48%	52%	48%	61%	39%	61%	39%	61%	39%

Notes: In this table it is shown how did those students adjust their expected earnings who have changed their preferred specializations, in percentages, across Treatment and Control groups, for the expected earnings right after graduation of university (Afteruni), at the age of 30 (30y) and at the age of 45 (45y), across male and female students. Adjustment is displayed among those students who have lowered their expectations (Adjust-), who did not change (0), and who increased (Adjust+). Males are represented more in this cluster both in control and treatment groups, in comparison with males.

Table 4: How students adjusted their expected earnings in the whole population?

	Treatment						Control					
	Afteruni		30y		45y		Afteruni		30y		45y	
	M	F	M	F	M	F	M	F	M	F	M	F
Adjusted-	20%	20%	18%	20%	19%	19%	11%	18%	9%	18%	11%	16%
0	8%	12%	11%	10%	10%	12%	6%	11%	7%	8%	6%	10%
Adjusted+	20%	20%	19%	22%	19%	21%	29%	26%	30%	28%	29%	28%
Total	48%	52%	48%	52%	48%	52%	46%	54%	46%	54%	46%	54%

Notes: In this table it is shown how students adjusted their expected earnings in the whole population, in percentages, across Treatment and Control groups, for the expected earnings right after graduation of university (Afteruni), at the age of 30 (30y) and at the age of 45 (45y), across male and female students. Adjustment is displayed among those students who have lowered their expectations (Adjust-), who did not change (0), and who increased (Adjust+). Females are represented more in treatment and control groups in comparison with males.

According to the Table 3, in the whole population, the ratio of students who didn't change their expected earnings at all three stages: after graduation, at the age of 30 and at the age of 45, exceeds in the treatment group in comparison with the control group. Mostly, both males and females were adjusting in the positive direction, this trend is more visible in the control group. Among those female students who changed their preferred schools this tendency seems to hold, however, males tend to lower their expected earnings in treatment schools, while in the control schools the majority of mails were increasing their expectations (Table 4). This negative adjustment in the treated schools among males is not really surprising since from the Figures 1-3 of Appendix is noticeable that they had higher expectations in comparison with females and the average actual earnings as well. This tendency makes me believe that the information provision had an impact.

However, along with the treated schools, large proportion of students from the controlled schools also have adjusted their expected earnings in the second round. Of course this could be just noise[9],

[9] In the experimental survey, for the specializations students' choices were incentivized, as already mentioned, surveyors promised to send the information regarding the universities, admission process and some other academic indicators based on the first choice, therefore, students could be more attentive to the specialization choices but not towards the expected earnings.

but on the other hand, even in the control schools students could have started thinking about it after the first round and asked their acquaintances, parents regarding the earnings, meaning they took the survey seriously and came for the second round more prepared. If this is the case, the results could be tracked by the fact whether a student discussed with parents or not. Overall, at around 78% of the students has discussed the survey with their parents.

Table 5: What ratios of students discussed with their parents and adjusted their expectations?

	Control						Treatment					
	Afteruni		30y		45y		Afteruni		30y		45y	
	A	NA	A	NA	A	NA	A	NA	A	NA	A	NA
D	76%	84%	76%	84%	76%	84%	72%	83%	73%	80%	74%	78%
ND	24%	16%	24%	16%	24%	16%	28%	17%	27%	20%	26%	22%

Notes: This table displays the ratios of students who discussed (D) and did not discuss (ND) the survey after the first round among those students who adjusted (A) and did not adjust (NA) their expected earnings, observing separately for the controlled and treated schools.

Based on the Table 5, after discussion with parents in both treatment and control groups more students did not adjust. The discussion rate with their parents is higher in the controlled groups in comparison with treatment groups. In the control group 76% adjusted, so this part can be counted as the explanation and not the noise. And those 24% who adjusted but did not discuss with their parents, they could have discussed with their friends, or made research on their own, however, still, part of it can be the noise.

Table 6: What ratios of students discussed with their parents and adjusted their expectations across those observations who adjusted their preferred specializations?

	Control						Treatment					
	Afteruni		30y				Afteruni		30y			
	A	NA		A	NA		A	NA		A	NA	
D	6%	3%	6%	3%	6%	4%	16%	5%	16%	4%	16%	5%
ND	4%	1%	4%	1%	4%	2%	13%	2%	12%	3%	10%	5%

Notes: This table displays the ratios of students who discussed (D) and did not discuss (ND) the survey after the first round among those students who adjusted (A) and did not adjust (NA) their expected earnings, observing separately for the controlled and treated schools, clustered for those students who have changed their preferred specializations

Table 6 shows that the discussion rate is higher in the treated schools. Those 6% who discussed with parents and adjusted their expected earnings in the control group can be counted as the explained part and not the noise, and for the remaining 4% the same logic can be applied as above.

Moreover, in both groups more students adjusted their expected earnings after discussion with their parents.

How are the expected earnings distributed across genders? As can be seen from the Figures 4-9 in Appendix, females consider their expected earnings more in the lower bounds, especially in the cases of the expected earnings at the age of 30 and 45, while males indicate more in the upper bounds. The thresholds are at around the medians for each case, exact numbers are presented in the Table 2 in Appendix. It is evidential from this table that students are increasing their expected earnings on average in the second round in comparison with the first round.

Table 7: How are genders distributed among treatment and control groups?

	Male	Female
Treatment	57.7%	55.3%
Control	42.3%	44.7%

Notes: The table indicates how male and female students were assigned in the treatment and control groups. Meanings that 57.7% of male students were from treated schools and 42.3% from controlled schools, same reasoning for female students.

The table 7 indicates what fraction of males is in the treatment group and in control group and the same for females. Having 57.7% of males were assigned to the treated schools and from females - 55.3%.

4 Methodology

In this chapter I present empirical strategies first to show the significance of the imposed treatment effect, and in the second part I describe the methodology used to explain the gender gap in expected earnings.

4.1 Treatment effect significance

To measure the impact and the significant of provision of information, I used Ordinary Least Squares (OLS) and Probit models, similarly to Kudashvili and Todua (2021).

First I test the significance on the change of preferred specialization. For this purpose I use the simple linear probability model on treatment variable (1), then I include covariates (2), and in the final phase I run the Probit on the model with covariates. In order to test the significance of treatment effect along with the gender, I include the interaction between treatment and gender in the analysis. Again, first I use the simple linear probability model on treatment, gender and interaction (3), afterwards I include other covariates as well (4), and in the final part I run the Probit on the model with covariates.

$$change_i = \alpha + \beta_1 treatment_i + \varepsilon_i, \text{ for each individual } i \quad (1)$$

$$change_i = \alpha + \beta_1 treatment_i + \sum_{j=2}^{n} \beta_j X_{ji} + \varepsilon_i, \text{ for each individual } i \quad (2)$$

$$change_i = \alpha + \beta_1 treatment_i + \beta_2 female_i + \beta_3 treatment * female_i + \varepsilon_i, \quad (3)$$

for each individual *i*

$$change_i = \alpha + \beta_1 treatment_i + \beta_2 female_i + \beta_3 treatment * female_i + \sum_{j=4}^{n} \beta_j X_{ji} + \varepsilon_i, \quad (4)$$

for each individual *i*

In all (1)-(4) regressions the dependent variable is the dummy indicating whether a student changed its preferred specialization or not, and X are the rest of covariates excluding treatment variable in (2) and excluding treatment, gender and interaction parts in the regression (4).

In order to estimate the significance of treatment effect and the interaction part on the expected earnings, I first run the abovementioned 4 regressions with the dependent variables of expected earnings in round 1 and round 2 for all three stages of lifetime: after university graduation, at the age of 30 and at the age of 45.

4.2 Gender gap

For studying the gender gap in expected earnings I follow the Oaxaca-Blinder decomposition method.

The population has been divided into two groups, male (M) and female (F) students. In the analysis are used three different dependent variables:

$$Y_i = \begin{cases} expected\ earnings\ after\ graduation\ of\ university_i \\ expected\ earnings\ at\ the\ age\ of\ 30_i \\ expected\ earnings\ at\ the\ age\ of\ 45_i \end{cases}$$

First, the Oaxaca-Blinder estimates $Y_i = X_i\beta + \varepsilon_i$ for each gender separately. While being able to observe the gender-specific means of dependent and explanatory variables, the O-B method fits model separately in the two subsamples:

$\bar{Y}^M = \bar{X}^M \hat{\beta}^M$, for males

$\bar{Y}^F = \bar{X}^F \hat{\beta}^F$, for females

The construction of outcome difference between the means of the two groups is based on the counterfactuals. The desired counterfactual should answer the question: "What would be the average expected earning of females if they had the same mean characteristics as males in this population?" or vice-versa, namely, $\bar{Y}^{F}_{counterfactual} = \bar{X}^{M}\,\hat{\beta}^{F}$. Consequently, the O-B decomposition helps to determine whether the gender gap in the means of the expected earnings is due to the difference in characteristics, by constructing explained and unexplained parts:

$$\bar{Y}^{F} - \bar{Y}^{M} = \bar{X}^{F}\,\hat{\beta}^{F} - \bar{X}^{M}\,\hat{\beta}^{M} = \underbrace{(\bar{X}^{F} - \bar{X}^{M})\,\hat{\beta}^{M}}_{\substack{explained\ by\ the\ difference \\ in\ characteristics}} + \underbrace{\bar{X}^{F}\,(\hat{\beta}^{F} - \hat{\beta}^{M})}_{unexplained\ part}$$

In particular, first I run the O-B decomposition separately on the expected earnings after university graduation, at the age of 30 and at the age of 45 reported in the first round and in the second round, with and without treatment variable included as an explanatory one, along with other covariates. Having the treatment effect in the first round estimation is a demonstration of the randomization check, while for the second round I would like to observe the significance of the treatment variable. Among the other covariates I have created pooled subsets:

Table 8: Grouped covariates

Groups	Variables	Notes
Specializations	Dummy variables of each specialization: exact and natural sciences, medical sciences, economics and business, social sciences, art and humanities, and law.	For avoiding multicollinearity problem, exact and natural sciences dummy has been omitted. Specializations' effect has been reported separately for each field and for pooled as well.
Personal	Whether a student has a PC/ Laptop, number of sisters and brothers, their own evaluation of their performance on the school level.	

Beliefs	Perceived mean earnings and unemployment rates separately for each specialization and for individuals without a degree.	
Parents	Dummy variables for mother and father professions of: exact and natural sciences, medical sciences, economics and business, social sciences, art and humanities, and law, or without a degree	
School	Number of parallel classes, number of students, number of teachers, whether a school is specializing in a specific subject or not.	

In the Figures 4-9 in Appendix it is observable that the gaps are different before a certain threshold for each expected earnings, as already mentioned in descriptive statistics section these thresholds are very close to the median. Therefore, similar to (Fortin, Bell, & Böhm, 2017), for further analysis I divide the subsets into centiles, more specifically, into two parts - students who indicated their expected earnings below median (0-50%) and above median (50-100%). And apply O-B decomposition by gender separately for each centile.

Since the treatment effect is a dummy variables, in the O-B decomposition the part of explained by treatment variable would be interpreted as "What would be the gap in expected earnings between males and females if among females the distribution in the treated schools would be the same as among males?" Therefore, for the treatment effect to be significant it is essential to have different ratios of males and females being in the treated groups. Based on the Table 7, this is not the case, therefore, I expected treatment effect to be insignificant in the above-mentioned analysis. But the unexplained part of treatment effect should be indicating its importance.

5 Estimations Results

In this chapter are presented the results of the estimation strategies described in the previous section.

5.1 Treatment effect significance

Table 9: Treatment significance on specialization change

	Linear		Probit	Linear		Probit
	1	2	3	4	5	6
treatment	0.11****	0.11****	0.53***	0.13****	0.12****	0.70***
	(0.02)	(0.02)	(0.11)	(0.03)	(0.03)	(0.16)
female	-	0.05**	0.24**	0.08**	0.07**	0.47***
		(0.02)	(0.10)	(0.03)	(0.03)	(0.18)
treatment*female	-	-	-	-0.04	-0.04	-0.34
				(0.04)	(0.05)	(0.22)
Covariates[a]	No	Yes	Yes	No	Yes	Yes
_cons	.09****	-0.06	1.01	0.06***	-0.08	-0.84
	(0.02)	(0.37)	(83.79)	(0.02)	(0.37)	(95.92)
Observations	1259	1083	1083	1259	1083	1083
R^2	0.02	0.06		0.03	0.06	

Notes: Significance levels: **** p<0.001, *** p<0.01, ** p<0.05, * p<0.1, standard errors are presented in the parentheses. Dependent variable is a dummy variable illustrating whether a student changed preferred specialization between round 1 and round 2 or not.

[a] *Covariates:* first choice specialization, if a student has a PC/Laptop, number of siblings, their own evaluation of their performance on the school level, beliefs regarding mean expected earnings and unemployment rates separately for each specialization and for individuals without a degree, parental occupations.

Columns 1 and 2 correspond to the equations (1) and (2), while columns 4 and 5 – (3) and (4), in the columns 3 and 6 probit model is run on the corresponding models with covariates.

From Table 9 it is evidential that treatment effect has a significant positive impact, in particular, in the treated schools it is by 11% more probable that students have changed their preferred specialization in comparison with the student from controlled schools. The gender variable is significant separately as well, but the interaction is not (Columns 4-5) meaning that there is no heterogeneity in the treatment effect on the specialization changes by gender.

Table 10a: Treatment significance on expected earnings, round 1

	Afteruni				30y				45y			
	1	2	3	4	5	6	7	8	9	10	11	12
treatment	-99.53	0.76	-130.91	-36.71	-164.54	3.51	-253.86	-154.34	-477.89	-185.86	-549.65	-438.40
	(67.81)	(67.84)	(92.40)	(91.05)	(151.54)	(150.08)	(205.71)	(201.32)	(295.05)	(290.18)	(400.76)	(389.33)
female	-	174.95**	202.08**	127.12	-	483.01 ***	610.04 ***	281.57	-	935.27 ***	1268.21 ***	613.00
		(70.80)	(101.92)	(104.96)		(156.64)	(226.90)	(232.08)		(302.87)	(442.04)	(448.83)
treatment* female	-	-	55.15	83.56	-	-	155.18	351.93	-	-	84.83	563.02
			(135.47)	(135.32)			(301.61)	(299.23)			(587.58)	(578.67)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
_cons	1173.61 ****	468.97	1081.76 ****	506.96	2829.02 ****	-759.96	2551.73 ****	-599.97	5111.52 ****	-807.08	4535.06 ****	-551.14
	(50.94)	(1162.05)	(58.14)	(1164.01)	(113.84)	(2570.67)	(152.98)	(2573.81)	(221.65)	(4970.43)	(298.02)	(4977.51)
Observations	1313	1132	1313	1132	1313	1132	1313	1132	1313	1132	1313	1132
R^2	0.0016	0.0616	0.0109	0.0620	0.0009	0.1081	0.0175	0.1093	0.0020	0.1037	0.0173	0.1045

Notes: Significance levels: **** p<0.001, *** p<0.01, ** p<0.05, * p<0.1, standard errors are presented in the parentheses. Dependent variable is the reported expected earnings after university graduation (afteruni) in round 1(Columns 1-4), at the age of 30 (30y) between in round 1 (Columns 5-8), at the age of 45 (45y) in round 1 (Columns 9-12). In the columns 1, 2; 5, 6; 9, 10 the analysis is provided for the treatment effect, while in the columns 3, 4; 7, 8; 11, 12 for the treatment effect, gender and their interaction. In turn columns 1 and 3, 5 and 7, 9 and 11 are the simplest OLS regressions on the core covariates, in the columns 2 and 4, 6 and 8, 10 and 12 other explanatory variable are also included.

Covariates: first choice specialization, gender (female), if a student has a PC/Laptop, number of siblings, their own evaluation of their performance on the school level, beliefs regarding mean expected earnings and unemployment rates separately for each specialization and for individuals without a degree, parental occupations.

Expected earnings are measured in Georgian Lari (GEL); the average exchange rate in 2017 was approximately $1=2.4 GEL

Table 10b: Treatment significance on expected earnings, round 2

	Afteruni				30y				45y			
	1	2	3	4	5	6	7	8	9	10	11	12
treatment	-226.14 ****	-121.98 **	-190.02 **	-94.54	-416.28 ***	-278.59*	-181.70	-42.46	-497.13	-206.93	-259.42	15.59
	(61.12)	(61.85)	(82.51)	(83.01)	(149.87)	(152.13)	(200.48)	(203.92)	(309.56)	(316.06)	(415.74)	(424.11)
female	-	303.83 ****	416.62 ****	338.84 ****	-	892.86 ****	1429.51 ****	1194.18 ****	-	1660.78 ****	2486.54 ****	1944.75 ****
		(64.55)	(91.00)	(95.69)		(158.78)	(221.12)	(235.08)		(329.87)	(458.57)	(448.92)
treatment* female	-	-	-96.74	-61.17	-	-	-562.86*	-526.43*	-	-	-623.52	-496.10
			(120.97)	(123.38)			(293.93)	(303.09)			(609.55)	(630.36)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
_cons	1208.41 ****	332.29	1019.04 ****	304.49	3067.89 ****	-298.14	2418.11 ****	-537.45	5452.36 ****	-1754.34	4322.12 ****	-1979.86
	(45.92)	(1059.39)	(61.35)	(1061.24)	(112.59)	(2605.78)	(149.08)	(2607.02)	(232.56)	(5413.59)	(309.17)	(5422.11)
Observations	1313	1132	1313	1132	1313	1132	1313	1132	1313	1132	1313	1132
R^2	0.0103	0.0851	0.0376	0.0853	0.0059	0.1401	0.0507	0.1425	0.1085	0.1228	0.0393	0.1233

Notes: Significance levels: **** p<0.001, *** p<0.01, ** p<0.05, * p<0.1, standard errors are presented in the parentheses. Dependent variable is the reported expected earnings after university graduation (Afteruni) in round 2(Columns 1-4), at the age of 30 (30y) between in round 2 (Columns 5-8), at the age of 45 (45y) in round 2 (Columns 9-12). In the columns 1, 2; 5, 6; 9, 10 the analysis is provided for the treatment effect, while in the columns 3, 4; 7, 8; 11, 12 for the treatment effect, gender and their interaction. In turn columns 1 and 3, 5 and 7, 9 and 11 are the simplest OLS regressions on the core covariates, in the columns 2 and 4, 6 and 8, 10 and 12 other explanatory variable are also included.

Covariates: first choice specialization, gender (female), if a student has a PC/Laptop, number of siblings, their own evaluation of their performance on the school level, beliefs regarding mean expected earnings and unemployment rates separately for each specialization and for individuals without a degree, parental occupations.

Expected earnings are measured in Georgian Lari (GEL); the average exchange rate in 2017 was approximately $1=2.4 GEL

Table 10a and 10b report the estimates of treatment effect significance but on the expected earnings after graduation of university, at the age of 30 and at the age of 45 in round 1 and round 2. Treatment effect is almost everywhere negative for the expected earnings at each stage of the lifetime, but significant ones are in the second round rather than the first round (Table 10b). More precisely, students from treated school indicate by 226.14 GEL lower expected earnings after university graduation than in the controlled schools in the second round. These gaps between treated and controlled school increase for the expected earnings at the age of 30 and at the age of 45 as well. Overall, in the treated schools students on average expect earnings after university graduation in the second round equal to 982.27 GEL, while in controlled schools – 1208.41 GEL.

There clearly exists strong heterogeneity in the treatment effect by gender for the expected earnings at the age of 30in the second round (Column 7, 8 of Table 10b).

5.2 Gender gap

The results of Oaxaca-Blinder decompositions by gender of expected earnings is presented in this subsection.

Table 11: O-B Decomposition

	Round 1			Round 2					
	Afteruni	30y	45y	Afteruni		30y		45y	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Difference (*F-M*)	-201.29 ***	-548.51 ****	-1079.35 ****	-325.73 ****	-325.73 ****	-1011.72 ****	-1011.72 ****	-1958.74 ****	-1958.74 ****
Explained	-18.56	-7.45	-147.54	-50.65*	-45.32	-208.54 **	-198.26 **	-509.84 ***	-504.16 ***
% of gap	*9%*	*1%*	*14%*	*16%*	*14%*	*21%*	*20%*	*26%*	*26%*
Unexplained	-182.73 ***	-541.05 ****	-931.81 ***	-275.07 ****	-280.41 ****	-803.17 ****	-813.46 ****	-1448.89 ****	-1454.57 ****
% of gap	*91%*	*99%*	*86%*	*84%*	*86%*	*79%*	*80%*	*74%*	*74%*
Mean of control group	1173.61	2829.02	5111.52	1208.40		3067.88		5452.36	
				explained					
Treatment	-	-	-	-	6.13	-	11.83	-	6.53
Specializations	-12.73	-37.58	-47.66	-1.85	-2.19	-63.93	-64.59	-117.02	-117.38
Med	*7.77*	*16.79*	*42.73*	*7.95*	*7.53*	*-1.97*	*-2.78*	*-3.79*	*-4.24*
EconBus	*-20.14*	*-69.81 ** *	*-129.71 ** *	*-20.25** *	*-19.98** *	*-34.27*	*-33.77*	*-77.80*	*-77.52*
Soc	*-5.81*	*-7.26*	*-24.02*	*-3.05*	*-3.27*	*-13.36*	*-13.80*	*-29.51*	*-29.75*
ArtHum	*2.21*	*11.63*	*37.16*	*11.26*	*11.66*	*-9.92*	*-9.15*	*9.04*	*9.46*
Law	*3.23*	*11.05*	*26.17*	*2.22*	*1.87*	*-4.39*	*-5.07*	*-14.95*	*-15.32*
Personal	0.96	27.21	27.57	5.43	5.60	44.74*	45.07*	54.58	54.76
Beliefs	-25.21	-71.08	-126.30	-21.99	-24.27	-67.47	-71.86	-128.13	-130.55
Family	12.90	9.39	3.76	5.87	4.40	-13.03	-15.88	-58.01	-59.58

CEU eTD Collection

School	5.51	64.60	-4.91	-38.11**	-34.98*	-108.84 **	-102.81 **	-261.26 ***	-257.93 ***
				unexplained					
Treatment	-	-	-	-	86.24	-	383.42*	-	476.96
Specializations	-150.82	-555.30	-889.65	-123.86	-129.78	-1183.44 ***	-1197.40 ***	-1886.03 **	-1898.97 **
Med	*32.85*	*75.56*	*333.38**	*-10.59*	*-5.88*	*-60.29*	*-48.51*	*47.08*	*58.15*
EconBus	*145.35**	*287.68**	*736.70 ***	*20.59*	*18.01*	*110.15*	*106.81*	*266.72*	*265.66*
Soc	*26.86*	*67.62**	*203.51 ***	*-3.70*	*-2.81*	*38.77*	*41.14*	*75.10*	*77.36*
ArtHum	*2.37*	*-41.44*	*47.34*	*-41.87*	*-41.63*	*-135.79*	*-137.55*	*-215.39*	*-218.42*
Law	*-10.52*	*-96.54*	*-50.32*	*-41.41*	*-38.77*	*-171.86**	*163.06**	*-324.08*	*-314.42*
Personal	208.72	433.87	829.80	153.77	129.30	207.53	156.55	1662.46	1620.51
Beliefs	-364.62	-729.28	-2185.66 *	-365.34	-351.19	-1246.61 **	-1250.90 **	-1705.91	-1736.62
Family	-34.76	924.27	1295.14	-117.96	-94.59	1002.19	1105.43	2212.76	2339.69
School	1232.04	2016.34	3166.18	-47.73	-1192.78	3948.02	-726.78	4129.36	-1502.81
Observations	1132	1132	1132	1132	1132	1132	1132	1132	1132

Notes: Significance levels: ****p<0.001, ***p<0.01, **p<0.05, *p<0.1.

Dependent variables are: expected earnings after university graduation (Afteruni), at the age of 30 (30y), at the age of 45 (45) in the first round (Columns 1, 2 and 3, respectively) and in the second round (Columns 4, 5; 6, 7 and 8, 9).

Gender gap is calculated as $\bar{Y}^F - \bar{Y}^M$.

Covariariates: **Specializations:** medical sciences, economics and business, social sciences, art and humanities, law (exact and natural sciences has been omitted) – has been reported separately for each field and for pooled as well. **Personal**: if a student has a PC/Laptop, number of siblings, their own evaluation of their performance on the school level (self ranking on the school level); **Beliefs**: beliefs of expected earnings and unemployment rates separately for each specialization and for individuals without a degree; **Family**: Mother and Father professions; **School**: number of parallel classes, number of students, number of teachers, if a school is specialized or not

From Table 11 it is clear that female students expect less earnings at all stage of a lifetime in comparison with males and 20% and 26% of this gap can be explained by the characteristics in the second round for the age of 30 and for the age of 45, respectively. It is also notable that from specializations economics and business tend to explain the gender gap more in comparison with other fields. Moreover, the provision of information contributes poorly to the explanation of the gap, which is not completely surprising. Table 7 illustrated that from males 55.7% were studying the treatment schools and from females – 55.3%. Since these ratios are almost equal that is why in the O-B decomposition treatment effect turns out to be insignificant. This reasoning is coming mainly from the fact that in the O-B decomposition the coefficient of the treatment effect is "what part is explained by the difference in the distributions of males and females in the treatment groups". Therefore, that is why the treatment coefficient is always close to zero and insignificant in the explained part.

However, there was a clear heterogeneity in the treatment effect by gender (Table 10b) that should be displayed in the unexplained part of the information provision for the case of expected earnings at the age of 30 in the second round. Indeed, the unexplained part of the treatment effect turns out to be significant for the expected earnings at the age of 30 with a positive sign (Table 11 Column 7). Meaning that the information provision contributed to the gender shrinkage equal to 383 GEL. Furthermore, in the unexplained part specializations and beliefs seem to be important factors in the gender gap of the second round expected earnings.

As already mentioned in the methodology chapter, based on the Figures 4-9 in Appendix the gaps differ below and above medians. Thus, I implemented the O-B decomposition on centiles as well.

Table 12: O-B Decomposition on centiles

	30y			
	Round 1		Round 2	
	<50%	>50%	<50%	>50%
	1	2	3	4
Difference (*F-M*)	10.18	-598.11 **	-39.96	-1142.65 ****
Explained	18.58	232.20*	-19.08	112.92
% of gap	*183%*	*-39%*	*48%*	*-10%*
Unexplained	-8.40	-830.31 ****	-20.88	-1255.56 ****
% of gap	*-83%*	*139%*	*52%*	*110%*
	10.18	-598.11 **	-39.96	-1142.65 ****
		explained		
Treatment	0.88	68.58	-0.91	19.71
Specializations:				
Med	*-0.17*	*45.94*	*-4.02*	*22.92*
EconBus	*-5.14*	*-34.87*	*-2.20*	*-8.01*
Soc	*-0.65*	*9.38*	*-0.93*	*-24.22*
ArtHum	*-14.80**	*-0.53*	*-15.35**	*-1.76*
Law	*-0.17*	*39.76*	*-1.83*	*-5.52*
Personal	14.93 **	-13.67	7.42	28.04
Beliefs	-13.64	115.50	-0.19	90.66
Family	-2.18	-33.46	0.73	-18.93
School	39.52 ***	35.56	-1.79	10.04
		unexplained		
Treatment	-109.24 ****	50.61	-61.78*	783.64 ***
Specializations:	-69.76	581.62	3.02	-107.97
Med	*-38.83 **	*259.74 **	*-30.70**	*-44.62*
EconBus	*-6.01*	*300.65*	*30.38*	*89.03*
Soc	*16.42*	*147.89 **	*14.83*	*78.98*
ArtHum	*-20.82*	*166.03*	*-9.14*	*11.81*
Law	*-20.52*	*-292.69 **	*-2.35*	*-243.18**
Personal	59.46	11.40	189.45*	-302.00
Beliefs	72.39	-937.27	97.60	-749.33
Family	-177.12	1160.74	311.53**	637.97
School	914.24	5904.35	1496.86	-9506.79
Observations	583	543	567	555

Notes: Significance levels: $^{****}p<0.001$, $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Dependent variables are: expected earnings at the age of 30 (30y) in the first round (Columns 1, 2) and in the second round (Columns 3, 4).

Gender gap is calculated as $\bar{Y}^F - \bar{Y}^M$. The analysis has been conducted separately on the centiles. As the threshold median values have been used. For the first round the median of the expected earnings at the age of 30 is equal to 1600 GEL, while for the second round – 1800 GEL (Table 2 in Appendix). Expected earnings are measured in Georgian Lari (GEL); the average exchange rate in 2017 was approximately $1=2.4 GEL

Covariariates: **Specializations:** medical sciences, economics and business, social sciences, art and humanities, law (exact and natural sciences has been omitted) – has been reported separately for each field and for pooled as well. **Personal**: if a student has a PC/Laptop, number of siblings, their own evaluation of their performance on the school level (self ranking on the school level); **Beliefs**: beliefs of expected earnings and unemployment rates separately for each specialization and for individuals without a degree; **Family**: Mother and Father professions; **School**: number of parallel classes, number of students, number of teachers, if a school is specialized or not

Since the only significant value has been reported in the case of expected earnings at the age of 30, the O-B decomposition on centiles has been implemented on that stage of a lifetime. Concentrating on the median-specific decompositions, from Table 12, it is evidential that in the first half (below median) the gender gaps are small in both rounds, while in the upper centile in the round 1 males were indicating much higher expected earnings in comparison with females at all stages of lifetime, and these gaps are increasing even more in the second round. Furthermore, the specializations seem to have significant effect in the upper centiles of expected earnings of the first round, in the unexplained parts. Moreover, the provision of information has a significant effect on the expected earnings at the age of 30 in the lower centiles of both rounds and in the upper centile of the second round, in the unexplained parts. However, only in the upper centile of the second round is of the same sign as in the second round of Table 11. Meaning that the upper centile has driven the gender gap shrinkage after information provision.

Is such result due to the change of specialization? One possible explanation for this might be that more female students changed their specializations towards those fields where the actual earnings are higher than their initial choice, in comparison with male students. For estimation, the same approach as for the Table 9 has been implied. The estimation results are presented in the Table 13.

Table 13: Treatment effect on the direction of change of specialization

	Wagechange			RelativeWagechange		
	Linear		Probit	Linear		Probit
treatment	-0.06***	-0.06***	-0.66****	-0.03	-0.03	-0.32
	(0.02)	(0.02)	(0.21)	(0.02)	(0.02)	(0.20)
female	-0.03	-0.01	-0.20*	-0.01	-0.01	-0.12
	(0.02)	(0.02)	(0.25)	(0.02)	(0.02)	(0.24)
treatment*	0.01	0.002	0.11	-0.03	-0.02	-0.11

female						
	(0.03)	(0.03)	(0.30)	(0.03)	(0.03)	(0.29)
Covariates	No	Yes	Yes	No	Yes	Yes
_cons	0.97****	0.96****	-0.05	0.97****	0.95****	-0.89
	(0.02)	(0.27)	(206.16)	(0.01)	(0.24)	(220.94)
Observations	1259	1083	914	1206	1093	1093
R^2	0.0154	0.0807		0.0126	0.0378	

Notes: ****p<0.001, ***p<0.01, **p<0.05, *p<0.1.

The direction of the change in specialization is based on the actual wages. If the second round choice of specialization has higher average actual earnings in comparison with the first round choice then the direction is considered positive. First round specialization actual earnings is wage1, second round – wage2. The corresponding beliefs of the first round specialization is belief1, of the second round – belief2.
Dependent variables are: dummy variable for the direction of change of specialization actual wages between rounds, taking value 1 if (wage2-wage1) is positive or equal to 0, and 0 otherwise (Columns 1-3), and dummy variables for the direction of change of specialization actual wages between round, taking value 1 if ((wage2-wage1)-(belief2-belief1)) if positive or equal to 0, and 0 otherwise (Columns 4-6).

[a]*Covariates:* first choice specialization, if a student has a PC/Laptop, number of siblings, their own evaluation of their performance on the school level, beliefs regarding mean expected earnings and unemployment rates separately for each specialization and for individuals without a degree, parental occupations.

Columns 1 and 2 correspond to the equations (1) and (2), while columns 4 and 5 – (3) and (4), in the columns 3 and 6 probit model is run on the corresponding models with covariates.

Same kind of estimations have been tested for the (wage2-wage1), (wage2-wage1)/wage1, ((wage2-wage1)-(belief2-belief1)), on the total population and on the centiles as well. But no significant results.

For the direction of the specialization changes, there is no heterogeneity in treatment effect by gender. Therefore, I would assume that in the upper centiles are mostly those female students who are career oriented and the provision of information accelerates their expectations, while the lower centiles are not altered by the treatment effect.

6 Conclusion

This thesis has examined the gender gap in expected earnings among students. This paper contributes to the existing literature in terms of examining inaccurate information among students as the source of the existing gender gap in expected earnings.

Analysing the experimental survey data from Georgia, students indeed hold incorrect beliefs regarding field-specific average earnings and unemployment rates (Kudashvili & Todua, 2021). Furthermore, I find the largest gaps in the fields of exact and natural sciences, art and humanities, and law. Overall, a persistent gender gap exists in favour for male students at all stages of a lifetime, almost doubling in the second round in comparison with the first round. The second round differences can be explained on average by 20%, specifically, the largest and most important factor turns out to be the school fixed effects. On the other hand, the information provision contributes in a positive way to the gender gap in the second round reported expected earnings at the age of 30. Moreover, this has been mainly driven by those students who reported expected earnings above the median value, while the lower bound has not been affected.

Such results might be important for policymakers since the information provision decreases the gender gap in the expected earnings at the age of 30.

The explanation part of the existing gender gap could be expanded by including some core variables that could not be observed from the collected data, such as, family income, student ability and performance (grades, test scores). Moreover, authors were able to incentivise honest responses regarding the preferred specializations by promising to send relative information, however, this is not the case for the expected earnings and students could play around and misreport.

For further investigation, the survey can be conducted in other regions of the country besides the capital city. Moreover, along with the above-mentioned missing core covariates, students could be asked beliefs regarding field-specific average earnings separately by gender. This would give additional information about perceived earning differentials by each gender. Afterwards, providing the gender-specific average earnings information for each field could contribute even more to the explanation of the gender gap.

References

Attanasio, O. P., & Kaufmann, K. M. (2009). Educational Choices, Subjective Expectations, and Credit Constraints. *NBER Working Paper 15087*.

Beffy, M., Fougere, D., & Maurel , A. (2012, February). Choosing the Field of Study in Postsecondary Education: Do Expected Earnings Matter? *The Review of Economics and Statistics, 94*(1), 334-347. doi:10.3386/w21655

Blau , F. D., & Beller, A. H. (1988, July). Trends in Earnings Differentials by Gender, 1971-1981. *ILR Review, 41*(4), 513-529. doi:10.2307/2523587

Blau , F. D., & Kahn, L. M. (2003, January). Understanding International Differences in the Gender Pay Gap. *Journal of Labor Economics, 21*(1), 106-144. doi:10.1086/344125

Blau, F. D., & Kahn, L. M. (2012). "The Gender Pay Gap: Have Women Gone as Far as They Can?". In J. Brueggemann, *Inequality in the United States: A Reader* (1 ed.). New York, United States. doi:10.4324/9781003071709

Blau, F., & Ferber, M. A. (1991). Career Plans and Expectations of Young Women and Men: The Earnings Gap and Labor Force Participation. *Journal of Human Resources, 26*(4), 581-607.

Blinder, A. S. (1973, Autumn). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources, 8*(4), 436-455. doi:10.2307/144855

Cohn, E., & Hughes Jr., W. W. (1994, June). A benefit-cost analysis of investment in college education in the United States: 1969–1985. *Economics of Education Review, 13*(2), 109-123. doi:10.1016/0272-7757(94)90002-7

Dias, M. C., Joyce, R., & Parodi, F. (2020, Winter). The gender pay gap in the UK: children and experience in work. *Oxford Review of Economic Policy, 36*(4), 855–881. doi:10.1093/oxrep/graa053

Elchanan Cohn, S. P. (1995, March). The wage effects of overschooling revisited. *Labour Economics, 2*(1), 67-76. doi:10.1016/0927-5371(95)80008-L

Filippin, A., & Ichino, A. (2005). Gender wage gap in expectations and realizations. *Labour Economics, 12*(1), 125-145.

Fortin, ,. N., Bell, B., & Böhm, M. (2017, August). Top earnings inequality and the gender pay gap: Canada, Sweden, and the United Kingdom. *Labour Economics, 47*, 107-123. doi:10.1016/j.labeco.2017.05.010

Ifatunji, M. A., & Harnois, C. E. (2015, November). An Explanation for the Gender Gap in Perceptions of Discrimination among African Americans: Considering the Role of Gender Bias in Measurement. *Sociology of Race and Ethnicity, 2*(3), 263-288. doi:10.1177/2332649215613532

Japaridze, E. (2012). *Assessment of the Gender Equality Policy in Georgia by Women's Organizations.*

Khitarishvili, T. (2019, June). GENDER PAY GAPS IN THE FORMER SOVIET UNION: A REVIEW OF THE EVIDENCE. *Journal of Economic Surveys, 33*(4), 1257-1284. doi:10.1111/joes.12321

Kim, C. (2010, June). Decomposing the Change in the Wage Gap Between White and Black Men Over Time, 1980-2005: An Extension of the Blinder-Oaxaca Decomposition Method. *Sociological Methods & Research, 38*(4), 619-651. doi:10.1177/0049124110366235

Kudashvili, N., & Todua, G. (2021). Information, Perceived Returns and College Major Choices.

Marini, M. M., & Fan, P.-L. (1997, August). The Gender Gap in Earnings at Career Entry. *American Sociological Review, 62*(4), 588-604. doi:10.2307/2657428

Oaxaca, R. (1973, October). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review, 14*(3), 693-709. doi:10.2307/2525981

Smith, J. P., & Welch, F. R. (1989, June). Black Economic Progress After Myrdal. *Journal of Economic Literature, 27*(2), 519-564.

Wellington, A. J. (1993, Spring). Changes in the Male/Female Wage Gap, 1976-85. *The Journal of Human Resources, 28*(2), 383-411. doi:10.2307/146209

Wiswall, M., & Zafar, B. (2015, April). Determinants of College Major Choice: Identification using an Information Experiment. *The Review of Economic Studies, 82*(2), 791–824. doi:10.1093/restud/rdu044

World Economic Forum. (2021). *Global Gender Gap Report.*

Zafar, B. (2013). College Major Choice and the Gender Gap. *Journal of Human Resources, 48*(3).

Appendix

Table 1: College major fields (as seen by respondents)[10]

1. Exact and Natural Sciences: Mathematics, Computer Science, Physics, Chemistry, Biology, Biochemistry, Geography, Geology, Ecology, Electrical and Mechanical Engineering, Transportation, Agriculture.

2. Medical Sciences: Medicine, Pharmacy, Dentistry, Public Health.

3. Economics and Business: Economics, Business Administration, Tourism, Management, Marketing, Accounting.

4. Social Sciences: Sociology, Politics, Journalism, Media and Communication, Political Studies, International Relations.

5. Art and Humanities: Philosophy, History, Archeology, Ethnology, Cultural Studies, Art History, Language Studies, Pedagogical Studies, Sports, Drama, Choreography.

6. Law: International Law, Public Law, Criminal Law, Civil Law.

Source: Kudashvili and Todua, 2021 (Kudashvili & Todua, 2021)

[10] College major fields were grouped based on actual major groups at Georgian universities in 2015.

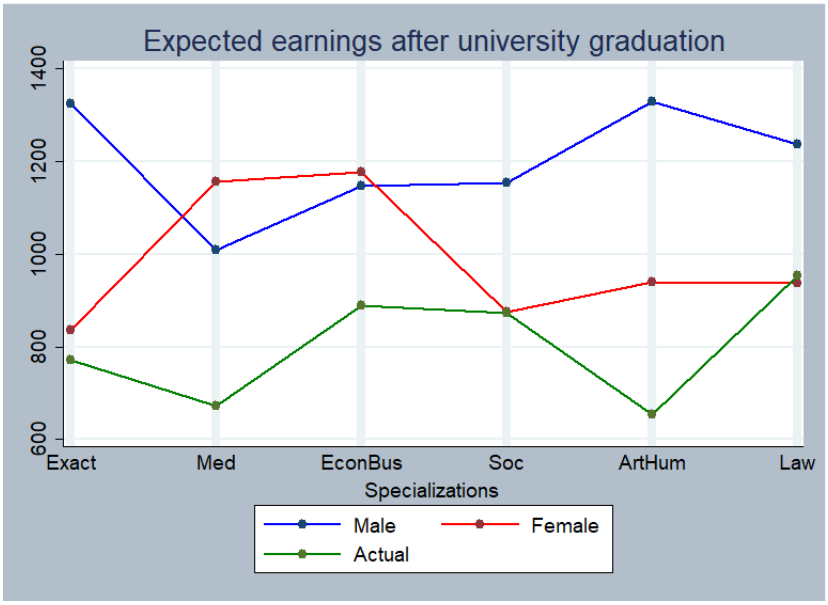


Figure 1: Expected earnings after university graduation by gender and actual earnings for each specialization

Notes: On the axis line are marked 6 specializations: Exact and Natural Sciences, Medical Sciences, Economics and Business, Social Sciences, Art and Humanities, and Law, respectively from the left to the right. Blue line represents average expected earnings after university graduation among male students for each specialization, red line corresponds to the mean of expected earnings after university graduation among female students for all 6 fields, and the green line illustrates the average actual earnings across the abovementioned 6 specializations. Both mean actual and expected earnings are given in Georgian Lari (GEL), the average exchange rate in 2017 was approximately $1=2.4 GEL

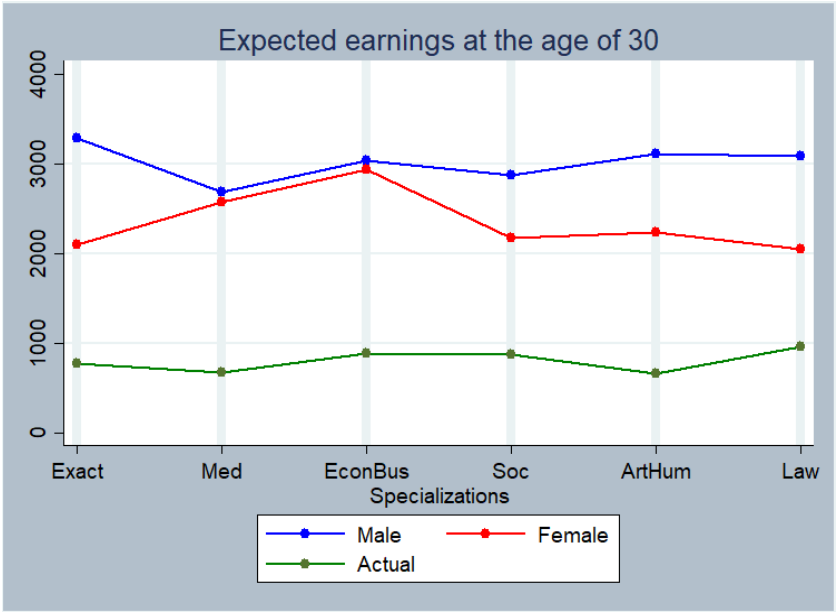


Figure 2: Expected earnings at the age of 30 by gender and actual earnings for each specialization

Notes: On the axis line are marked 6 specializations: Exact and Natural Sciences, Medical Sciences, Economics and Business, Social Sciences, Art and Humanities, and Law, respectively from the left to the right. Blue line represents average expected earnings after at the age of 30 among male students for each specialization, red line corresponds to the mean of expected earnings at the age of 30 among female students for all 6 fields, and the green line illustrates the average actual earnings across the abovementioned 6 specializations. Both mean actual and expected earnings are given in Georgian Lari (GEL), the average exchange rate in 2017 was approximately $1=2.4 GEL

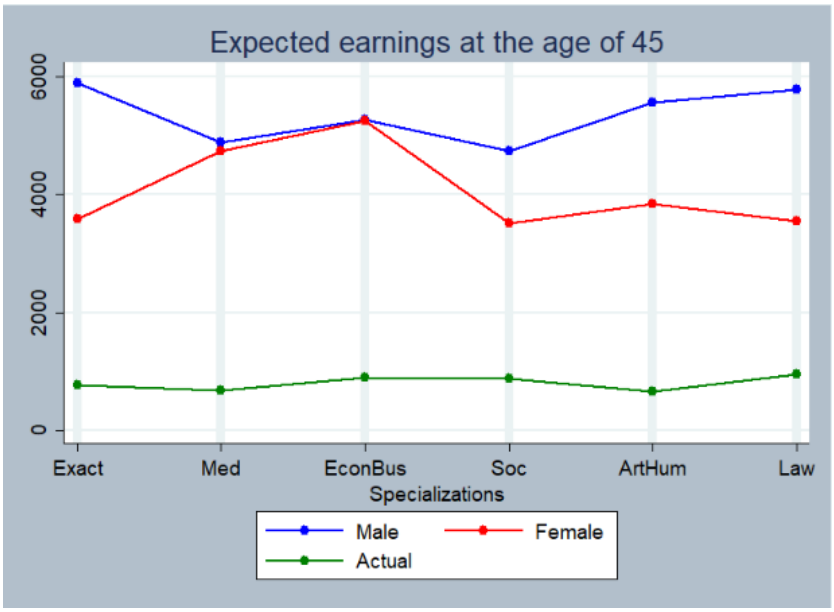


Figure 3: Expected earnings at the age of 45 by gender and actual earnings for each specialization

Notes: On the axis line are marked 6 specializations: Exact and Natural Sciences, Medical Sciences, Economics and Business, Social Sciences, Art and Humanities, and Law, respectively from the left to the right. Blue line represents average expected earnings after at the age of 45 among male students for each specialization, red line corresponds to the mean of expected earnings at the age of 45 among female students for all 6 fields, and the green line illustrates the average actual earnings across the abovementioned 6 specializations. Both mean actual and expected earnings are given in Georgian Lari (GEL), the average exchange rate in 2017 was approximately $1=2.4 GEL

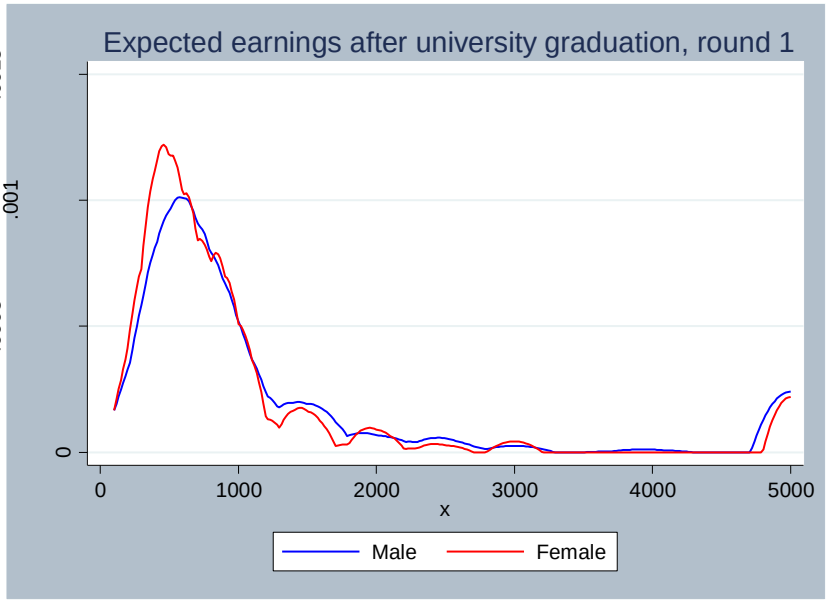
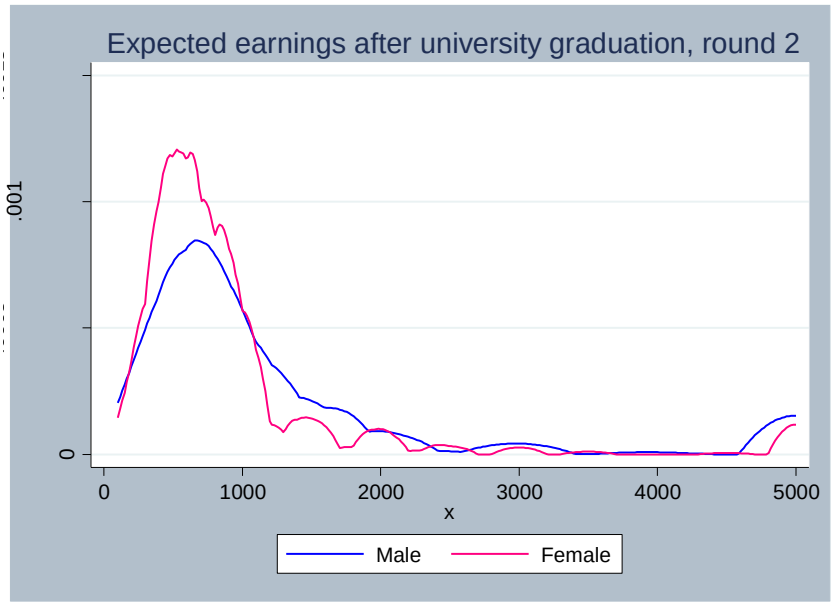


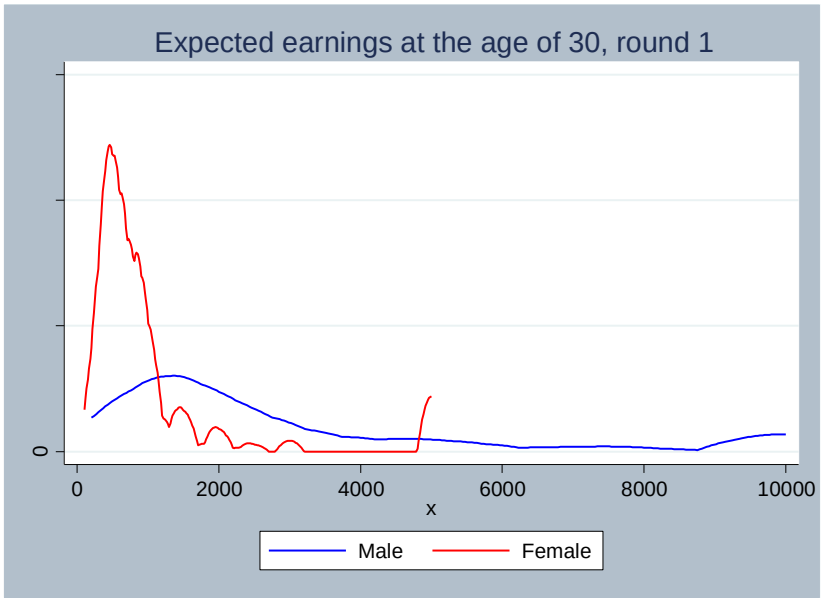
Figure 4: Density of expected earnings after university graduation, round 1

Notes: Densities of the first round reported expected earnings after university graduation for each gender – blue line is for males, red – females. Expected earnings are measured in Georgian Lari (GEL); the average exchange rate in 2017 was approximately $1=2.4 GEL



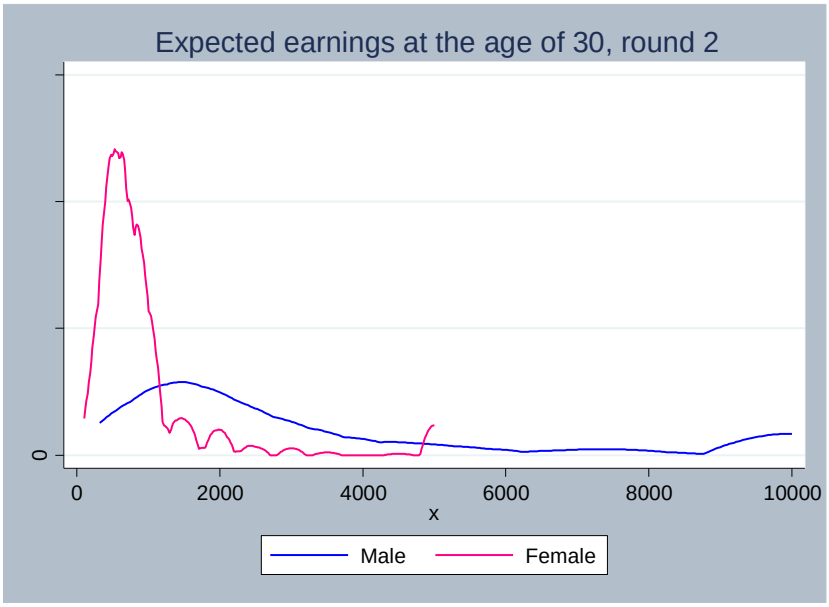
Notes: Densities of the second round reported expected earnings after university graduation for each gender – blue line is for males, red – females. Expected earnings are measured in Georgian Lari (GEL); the average exchange rate in 2017 was approximately $1=2.4 GEL

Figure 5: Density of expected earnings after university graduation, round 2



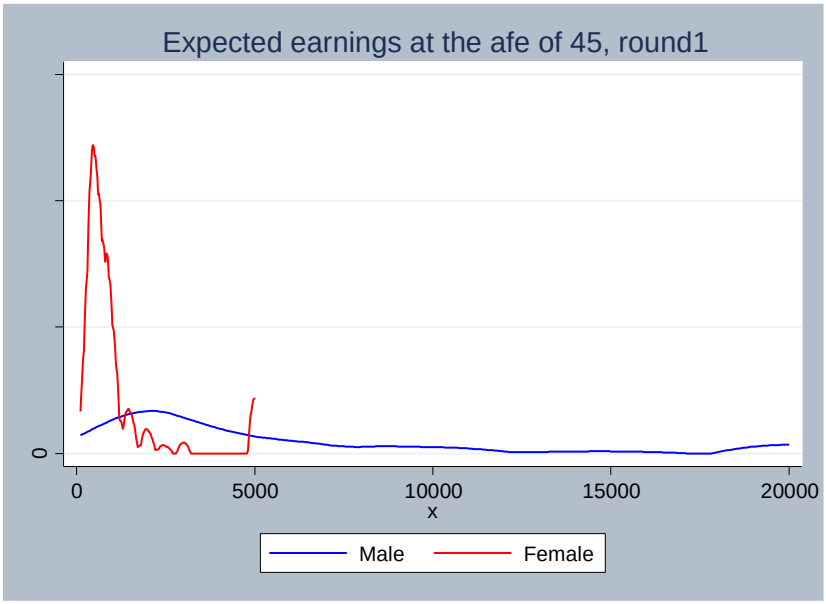
Notes: Densities of the first round reported expected earningsat the age of 30 for each gender – blue line is for males, red – females. Expected earnings are measured in Georgian Lari (GEL); the average exchange rate in 2017 was approximately $1=2.4 GEL

Figure 6: Density of expected earnings at the age of 30, round 1



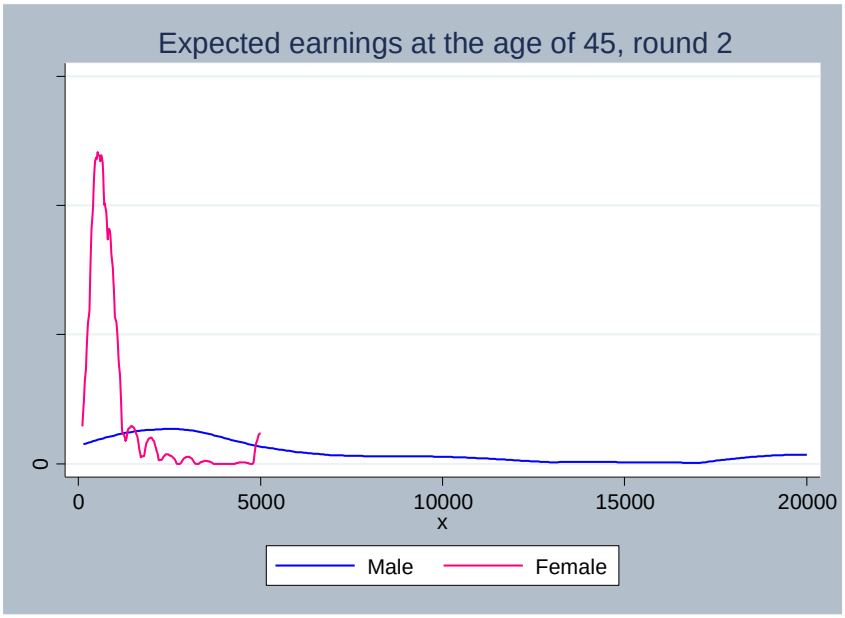
Notes: Densities of the second round reported expected earningsat the age of 30 for each gender – blue line is for males, red – females. Expected earnings are measured in Georgian Lari (GEL); the average exchange rate in 2017 was approximately $1=2.4 GEL

Figure 7: Density of expected earnings at the age of 30, round 2



Notes: Densities of the first round reported expected earningsat the age of 45 for each gender – blue line is for males, red – females. Expected earnings are measured in Georgian Lari (GEL); the average exchange rate in 2017 was approximately $1=2.4 GEL

Figure 8: Density of expected earnings at the age of 45, round 1



Notes: Densities of the second round reported expected earningsat the age of 45 for each gender – blue line is for males, red – females. Expected earnings are measured in Georgian Lari (GEL); the average exchange rate in 2017 was approximately $1=2.4 GEL

Figure 9: Density of expected earnings at the age of 45, round 2

Table 2: Median values of expected earnings

	Afteruni	30y	45y
Round 1	700	1600	2800
Round 2	750	1800	3000

Notes: In this table are indicated the overall median values for expected earnings after university graduation(Afteruni), at the age of 30 (30y), and at the age of 45 (45y), separately for round 1 and round 2. All of the values are measured in Georgia Lari (GEL); the average exchange rate in 2017 was approximately $1=2.4 GEL.