

GENDER DIFFERENCES IN
UNIVERSITY APPLICATIONS IN
UKRAINE: DO SOCIAL FACTORS
MATTER?

by

Inna Omelchenko

A thesis submitted in partial fulfillment of
the requirements for the degree of

MA in Economic Analysis

Kyiv School of Economics

2022

Thesis Supervisor: _____ Professor Olga Kupets

Approved by _____
Head of the KSE Defense Committee, Professor

Date _____

Kyiv School of Economics

Abstract

GENDER DIFFERENCES IN
UNIVERSITY APPLICATIONS IN
UKRAINE: DO SOCIAL FACTORS
MATTER?

by Inna Omelchenko

Thesis Supervisor:

Professor Olga Kupets

There is a lot of evidence that women are less presented in technological specialties than men. Researchers tried to investigate possible reasons for the unbalanced distribution in universities and the labor market. This thesis explores whether various social factors impact applicants' choices regarding specialization.

Using data from EDBO, ESS, and the State Statistic Service of Ukraine, we run the logistic model. The results showed that different indicators, like the larger share of religious people and those who follow traditions and have a negative attitude towards the LGBT community in the regions, decrease the probability of females applying for STEM degrees. In this work, we gave evidence that there are still people in society that have archaic views and are prejudiced by stereotypes about gender roles.

The thesis results could be considered in different policies that aim to raise the technological and cultural potential of the country and balance the gender roles.

TABLE OF CONTENTS

Chapter 1. INTRODUCTION.....	1
Chapter 2. LITERATURE REVIEW	4
Chapter 3. METHODOLOGY	7
Chapter 4. DATA	10
4.1 University applications data	10
4.2 ESS and State Statistics Service of Ukraine data.....	18
4.3 Final dataset	22
Chapter 5. ESTIMATION RESULTS	24
Chapter 6. CONCLUSIONS AND POLICY RECOMENDATIONS.....	30
6.1 Conclusions.....	30
6.2 Policy recommendations	31
WORKS CITED	32
APPENDIX A.	34

LIST OF FIGURES

<i>Number</i>	<i>Page</i>
Figure 1. Distribution of applications in majors by gender in 2018-2020.....	13
Figure 2. Distribution of STEM applicants by school region and gender	16
Figure 3. Distribution of Humanities applicants by school region and gender....	17

LIST OF TABLES

<i>Number</i>	<i>Page</i>
Table 1. Major categorizations.....	12
Table 2. EIE scores descriptive statistics (2018-2020).....	14
Table 3. I group. Individual enroller’s characteristics.....	15
Table 4. S group. Aggregated social factors characteristics per region and gender	19
Table 5. L group. Aggregated labor force and education level information per region and gender	20
Table 6. Descriptive statistics for the social data.....	22
Table 7. Estimation results	25
Table 8. APE results	26
Table A-9. University applications datasets description	34

ACKNOWLEDGMENTS

I would like to express my greatest appreciation to the Armed Forces of Ukraine and to all other people, organizations, and government structures that gave me the opportunity to proceed with my thesis work in this tough period for our country.

I would like to express my gratitude to my thesis advisor Olga Kupets. I am grateful for her guidance, meaningful advice, patience, and continuous support in this way.

Also, I thank all research workshop professors for their feedback, which helped to improve my thesis work.

Thanks to my family and friends who have supported and inspired me during the most challenging times this year.

LIST OF ABBREVIATIONS

APE. Average Partial Effect.

EDBO. Electronic Database on Education.

EIE. External Independent Evaluation (ZNO).

ESS. European Social Survey.

ILO. International Labor Organization.

IT. Information Technologies.

OECD. Organization for Economic Co-operation and Development.

PISA. Program for International Student Assessment.

STEM. Science, technology, engineering, and mathematics.

TVET. Technical and Vocational Education and Training.

UNDP. United Nations Development Program.

UNFPA. United Nations Population Fund.

Chapter 1

INTRODUCTION

A small share of females among graduates of technical majors is a highly observed global trend. According to Aphichoke and Das (2019), there are only 13.8% of the OECD countries (4 from 29) where the share of females with tertiary education in Science, Math, and Computing is equal to or slightly bigger than the share of males. Moreover, data from UNESCO shows that humanities and arts are more popular among women. This field of study is female-dominated in 60% of the countries (among 103), and only 8% of them are male-dominated. Following this further, the data on occupational segregation by gender (among 80 countries from the ILO database) shows that there are strictly male-dominant sectors: electrical trades, legislation, information, communications technology, etc. Teaching, healthcare, and sales occupations are more typical for women. Various researches among a number of countries specifically highlighted different social factors that can impact this unequal distribution.

Ukraine is not an exemption in this distribution. Looking closely at gender situation for graduates from STEM programs in tertiary education, there is 29% of women (UNDP 2020). This result is much lower compared to the humanities and arts field, where the value reached 79% (MES 2018). Looking at the Ukrainian labor force, the share of females working in technology-related occupational groups is only 21%¹ (Ukrstat 2021).

As for now, Information Technologies is one of Ukraine's most progressive and widely represented STEM fields. However, gender disbalance is also presented here. According to Sakhno et al. (2021), with an increasing number of students

¹ Calculated value. The aggregated sum of employed women among technology-related occupational groups (Ukrstat, 2021).

in IT majors and achieving gender balance by 2025, the economy of Ukraine could earn up to 32.7 million USD by 2030 from a smaller pay gap. Currently, only 20% of female graduates from IT majors are female. Despite the positive dynamics of female involvement in this industry: for the last ten years (2011-2020), the share of women in IT has drastically increased from 7% to 25% (DOU 2020), the gender inequality is still significant.

With this result, the unequal distribution between females and males in education and the following technical job occupation can be observed in Ukraine, even in such progressive sphere as IT. Therefore, it is important to look closer at the possible reasons for this issue.

One of the reasons may be the parents' approach to raising children. According to the PISA survey results, the parents' opinion was highlighted as one of the factors that impact enrollers' choices (OECD 2019). Interviewed girls whose parents encouraged them into engineering were more likely to choose a STEM major. Generally, all respondents recognized their parents' support as the primary key to their professional realization during their education.

As for Ukraine, gender stereotypes can influence specialty choices. According to UNFPA results, 30% believe IT and other engineering are more male-dominated (Vološevych and Podnos 2021). Alternatively, 51% of Ukrainians consider cleaning and raising children a female field. Additionally, 21% of parents persuade their children to choose a profession that suits their gender. From these results, we can conclude that outdated prejudices still prevail in society.

In the same time, Ministry of Education and Science recognized the school education system as one of the reasons for increasing gender stereotypes in Ukraine (MES 2018). The first steps to solving this issue have already been done. Since 2017, the division by gender in labor lessons in schools has been canceled. Also, the anti-discrimination examination of schoolbooks was launched. The

results of the examination showed that more than 70% of illustrations of highly skilled intellectual professions are performed by men, and only in images with humanities professions, such as teaching, women often appear. Moreover, most household illustrations depict women.

This work studies the causes of gender differences in Ukrainian university applications, particularly whether the social factors have any impact on the enrollers choices based on the data from EDBO, ESS, and the State Statistics Service of Ukraine.

The preliminary hypothesis suggests that various social factors of Ukrainian regions like higher self-identified level of religion, negative attitude towards the LGBT community, the lower level of happiness, the lower level of education in the region negatively impact the possibility of enrolling in the STEM major for women.

The previous literature review has shown several similar works for Ukraine, where authors tried to determine the factors influencing students' choices. However, most of them are restricted by one region or even a university. This work, on the contrary, covers all of Ukraine and be based on a massive sample of new data.

By processing the data of university applications for the 2018-2020 period, State Statistics Ukraine data, and the results of the European social survey, this thesis set a target to research the hypothesis. The regression is conducted on the cross-sectional data. In order to estimate whether different social factors affect enrollers' choices logit model is used.

This thesis is constructed in the following way. Chapter 2 gives the literature review. Chapter 3 represents the methodology that was used. Chapter 4 consists of a data description. Chapter 5 shows the main results obtained in the research. Conclusions and policy recommendations can be found in the last chapter 6.

Chapter 2

LITERATURE REVIEW

Researching the possible reasons for the gender differences in major choices is a popular topic investigated in different countries from various perspectives. Overall, in most researched papers, authors emphasize that the different social factors impact enrollers' choices.

The Delaney and Devereux (2020) and PISA results (OECD 2019) highlighted the influence of female insecurity on their major decision. First, researchers analyzed how the university and major choice for students with an equal level of knowledge variate by gender. In particular, the authors examined the social component – the aspiration and risk management among men and women in Ireland's college applications for 2015-2017. This study shows that females are more insecure about their achievements. Because of any possibility of not being enrolled in any major, they are more likely to decrease the requirements of the university and program they want to enroll in. On the other hand, the males with equal achievements show a higher level of confidence and set higher expectations about future institutions. Similar results were previously demonstrated in PISA results (OECD, 2019). The research revealed that females' choice is often affected by their insecurity. Every two out of three girls in OECD countries report that the possible failure makes them doubt their plans for the future. The same indicator for males is lower; fewer than one out of two boys have the same issue. Moreover, in most countries, boys were more likely to express a positive attitude toward competition. Consequently, considering all these and the existence of gender stereotypes of Ukrainian society presented in Chapter 1, in pair with the social pressure they create, the influence of female insecurity is highly likely to be present in Ukraine.

Speer (2017) presents the idea that males are more successful in mathematics, and science, thus, passing the pre-college exams at a higher level, which, in turn, explains the large part of the major gap. Based on the National Longitudinal Survey of Youth 1979 and 1997, Speer (2017) has researched the impact of pre-college academic factors by gender. The paper highlights that males are more successful in mathematics and science subjects, while females are likelier to show high results in humanities. As a result, these factors affect pre-college preparation and consequently explain a large part of the gender gap. However, this theory fails in Ukraine because of the high female performance in EIE. According to Vox Ukraine, women have a higher average grade for EIE in mathematics and non-technical subjects (Mykhailyshyna 2018). Thus, the males' propensity for math and higher achievements in pre-university exams cannot be the main factor that explains differences in major choices for Ukraine.

Black et al. (2015) show how belonging to particular ethnic and racial groups can impact enrollers' decisions. The authors used Texas public school graduates' 2008-2009 data to research the various college choices based on the behavior of students who belong to different ethnic and racial groups. The researchers found out that the application gap for these groups is not explained by varying levels of pre-university preparation or the previous school quality and rate. The research shows that students that belong to a particular minority are more likely to choose their college based on the high share of people of the same ethnicity or race. Additionally, the high school acquaintances' achievements in a particular educational institution can affect the enroller's choice in favor of this college.

On the other hand, various races and ethnic groups are not so widely represented in Ukraine. However, the same effect of choosing a major based on the large proportion of the people from one city or village could be observed for residents of the long distant regions. So, this study can be especially relevant if, instead of ethnic groups, we consider residents of specific remote locations.

Additionally, evidence of social factors' influence was found by Samokhin and Kohut (2017). Based on EIE-2016 survey results, they determined the possible factors influencing high-school students' performance in Ukraine. The authors have found a gap between urban and rural schools and showed that parents' education is crucial. Apart from that, some other factors turned out to be significant. Among them are the field of work and the parents' position, the ability to prepare for the test with a tutor, and gender. Research shows that socio-economic circumstances significantly affect learning outcomes despite the significant role of individual factors.

Generally, the literature review showed that social factors play a significant role in choosing a specialty by women. This research aims to investigate whether certain social factors could push women away from the tech industry in Ukraine.

Chapter 3

METHODOLOGY

This thesis investigates whether different social factors can impact the probability of enrolling in the technical specialty enrollers of both genders. The following model specification provides a methodology for indicating how various social prerequisites native to particular regions affect applicants' choices.

We considered several methods to determine the efficient specification. As the initial conditions, we have a categorical dependent variable and a set of categorical and continuous independent variables. In conjunction with the primary goal of the thesis – to predict the probability of the specific situation happening, we reviewed two models, multinomial logistic regression and logistic regression.

On the grounds that, initially, after major classification, we had four categories of the dependent major variable – Humanities, STEM, Medicine, and Military², a multinomial model was reviewed as a potential appropriate one for the current research. However, the last two categories represent only a minor part of the final dataset. In particular, 1.5% of all applications in the medicine category and less than 0.1% in the military major. Besides that, as can be seen from Figure 1, we have almost the same gender distribution for humanities and medicine categories, as well as for technical majors and military categories. This, in turn, is quite apparent results insofar as a military specialty is considered as more male activity in society (like STEM). Furthermore, medicine is recognized as a more female's occupation like the major humanities category (Volosevych and Podnos 2021).

Considering all the above, we proceeded further only with two major categories that take the largest share of the dataset (98.4%) – Humanities and STEM.

² The classification was made within the scope of this thesis

Since we left only two categories in the final dataset, the binary response model is considered the main one for the current study. The logit regression model is applied because of the binary dependent variable and different types of independent variables (including categorical). This model is supported by the literature review (Black et al. 2015) for estimation.

It is expected to research how the educational preferences of applicants differ depending on their origins. For this purpose, we set the logit model on the binary dependent major variable and a set of different characteristics as independent variables in regression

The logistic regression specification for the current research has the following form:

$$P(\mathit{major} = 1|x) = G(I_i\beta_1 + H_i\beta_2 + S_i\beta_3) \quad (3.1)$$

where G is the logistic function (theoretical model description see Wooldridge (2018)):

$$G(z) = \frac{e^z}{1 + e^z} \quad (3.2)$$

and I – individual characteristics (from the main enrollers’ dataset), L – aggregated labor force and education level information per region and gender, S – aggregated social factors characteristics per region and gender. Detailed information on the variables and their expected impact can be found in Table 3 - Table 5 presented in the next chapter.

The dependent variable is binary and presented by *major*, where

- 1 – if a student applied to a STEM major,
- 0 – if a student applied to a Humanity major.

Following this, the output of the logistic regression model presents the coefficients of the independent variables and their significance. In contrast to the linear model, we cannot interpret coefficients so directly. The regression output gives a clear picture of the magnitudes of the different factors, but to get a complete interpretation, we need to obtain a marginal effect.

We proceed with the average partial effect (APE) in order to compute a scale factor. It is calculated by averaging the individual partial effects in the dataset. The theoretical equation can be seen below (Wooldridge 2018):

$$APE = n^{-1} \sum_{i=1}^n \{G[\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \dots + \hat{\beta}_{k-1} x_{ik-1} + \hat{\beta}_1 (c_k + 1)] - G(\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \dots + \hat{\beta}_{k-1} x_{ik-1} + \hat{\beta}_1 c_k)\} \quad (3.3)$$

The APE allows us to evaluate how the probability will change if the predictor or independent variable increases by one unit.

Chapter 4

DATA

Cross-sectional data were used to research social factors' impact on high school graduates' choices. In particular, the thesis is based on the following data: university applications datasets for 2018-2020 and data about high schools (obtained from EDBO), the ESS dataset, and the data from the State Statistics Service of Ukraine (Ukrstat 2021).

4.1 University applications data

The individual-level data on university applications for bachelor's degrees from 2018 to 2020 was obtained by individual request from an EDBO source. It has three separate datasets for the 2018-2020 years, each with 15 variables. The description information on the data frame size and variables are presented in Appendix Table A-9.

The datasets consist of all enrollers' applications – for a unique student, this number varies from 1 to 15. For the data processing optimization, it was decided to reduce the number of applications to the top 3 per person. Considering that other applications have the same specialty choice, only in different universities, this amount is sufficient to represent enrollers' major choice.

After merging, a dataset with 1,388,849 observations for three years was obtained. This data contains two types of enrollers:

- 1) applicants only graduated from school, who live in their hometowns, and are likely to be affected by their relatives, friends, etc.

- 2) university/college graduates (enrollers that are eager to get a second degree) who have already relocated from their homes and are more likely not to be affected by certain factors of the home region for some time.

Since in this thesis, the question of social factors impacts the enroller's hometown region is researched, only the first type of applicants is considered. The second type of applicants could already be influenced by the factors of the region where they studied after school, so we did not take into account it.

However, this data contains only applicants' school names but not their origin region, which is crucial for further estimations. We used the high school dataset to map the school's name variable with the corresponding region. Since we match by string variable, many data were missing at the first iterations (this happens because school names in different datasets were included in different ways due to mistakes during manual input). We performed data normalization to avoid data loss and increase the number of matched names. In particular, we lowercased all data, removed unnecessary symbols (like prefixes to the school names), then identified and deleted all redundant whitespaces. These actions helped to increase the number of observations by about two hundred thousand. As a result, the dataset with 671,486 observations and 27 variables was created. After data preprocessing, the gender distribution among all applicants is equal to 44% for males and 56% for females.

Further, to simplify specialty identification categorization based on official Ukrainian sources (Osvita 2020) was conducted. All majors were distributed into four categories: Humanities, STEM, Medicine, and Military³. The content of each category is presented in Table 1.

³ The classification was made within the scope of this thesis

Table 1. Major categorizations

Humanities	STEM	Medicine	Military
Culture and art	Agricultural sciences and	Healthcare	Military
Education	food	Veterinary	sciences,
Philology	Architecture and	medicine	national
Journalism	construction		security, state
Law	Automation and		border security
Management and	instrumentation		
administration	engineering		
Public	Biology		
administration	Chemical and		
Service sector	bioengineering		
Social and	Electrical engineering		
behavioral sciences	Electronics and		
Social work	telecommunications		
Theology	Information Technology		
International	Mathematics and statistics		
relationship	Mechanical engineering		
Civil Security	Production and technology		
	Science		
	Transport		

Even though STEM has the largest number of specializations (major sub-category) – 56 in 12 majors compared to 42 in Humanities, the second one is more popular among enrollers. The distribution of applications among majors is following Humanities – 67.7%, STEM – 30.7%, Medicine – 1.5%, and Military – 0.1%. As expected, there are more females in Humanities as well as Medicine and a bigger number of males applying for technical majors and the military sphere (Figure 1). Due to the small number of medicine and military applications, we proceed further only with the Humanities and STEM categories.

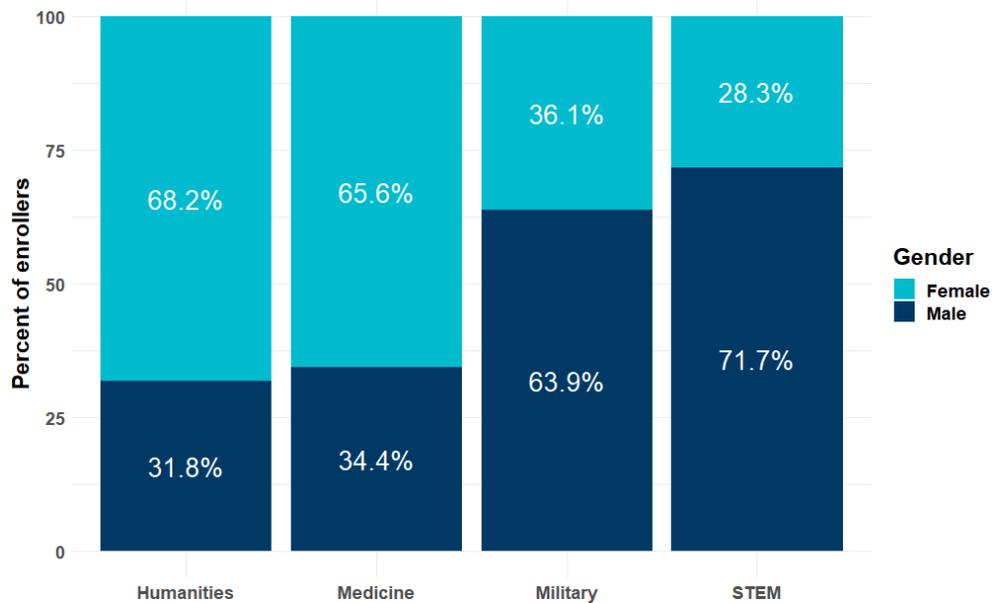


Figure 1. Distribution of applications in majors by gender in 2018-2020

Although, there are more man applications in STEM majors, women demonstrate promising results in technical subjects during EIE. Females' scores are not only better for humanitarian subjects like English, Ukrainian language, and History but also on one level or higher in technical exams like Mathematics and Physics. High results in exams are additional evidence against the stereotype that females are not able to succeed in technical subjects on one level as males. It is a good indicator that women have good pre-university preparation and are inclined toward science during their school education. The detailed descriptive statistics on EIE results are presented in Table 3 below.

Table 2. EIE scores descriptive statistics (2018-2020)

subject	gender	min	mode	median	mean	max	std dev
History of Ukraine	Female	100	135	151	150.4	200	24.1
History of Ukraine	Male	100	135	147	147.2	200	24.6
English	Female	100	124	136	140.2	194	23.2
English	Male	100	102	130	132.3	189	23.4
Biology	Female	100	158	151	150.7	200	24.6
Biology	Male	100	142	144	145.5	200	24.4
Geography	Female	100	119	149	148.5	200	24.2
Geography	Male	100	149	154	152.2	200	24.5
Mathematics	Female	100	155	153	151.3	200	25.6
Mathematics	Male	100	117	153	152.4	200	26.2
Ukrainian language and literature	Female	100	186	170	165.2	200	22.4
Ukrainian language and literature	Male	100	148	156	154.5	200	23.8
Physics	Female	100	135	154	152.3	200	26
Physics	Male	100	138	149	149.8	200	26.1
Chemistry	Female	100	154	152	152.3	200	26.8
Chemistry	Male	100	132	157	155.4	200	28.3

The University application dataset mainly provides information on enrollers' scores, universities, and specializations. However, while combining with the high school dataset (when we matched school names with corresponding regions), we obtained additional information about the type of the school: rural or not, and whether it is a boarding school. Therefore, from these two data sets, we got variables for the *I* group – individual applicant characteristics. A detailed description of the variables and expected impact can be found in Table 3 below.

Table 3. I group. Individual enroller’s characteristics

Variable	Description	Expected impact
Gender	Indicates gender: male – 1, female – 0.	Males are expected to be more likely to apply to a STEM major.
Belonging to the village (rural)	Indicates whether enroller is from village: 1 – yes, 0 – no.	Various traditions, high levels of religion, and stereotypes are present in villages. So, female applicants from the village are more likely to not go against preconceptions and not apply to the technical major.
Belonging to the boarding school	Indicates whether enroller is from the orphanage/boarding school: 1 – yes, 0 – no.	Children in boarding schools are under specific isolation and are not as influenced by social factors as those who live in full-fledged families. Accordingly, such things as the level of religion or traditions do not affect them so much. As a result, belonging to a boarding school increases the probability of entering a technical specialty.
Average score	Average grade for the enroller’s high school diploma.	Since STEM specialties have more heightened requirements of the pre-university level of knowledge, it is expected that enrollers with a higher grade for a high school diploma are more likely to apply to a technical major.
Overall score	Overall grade for the enroller’s three EIE results.	Since STEM specialties have more heightened requirements of the pre-university level of knowledge, it is expected that enrollers with a higher grade for a EIE are more likely to apply to a technical major.

Looking closer at the distribution of STEM applicants by their region and gender (Figure 2), there are considerably more enrollers from such technologically advanced, densely populated regions as Dnipro, Kharkiv, and Kyiv city. Moreover, a small share of females and an overall small number of applicants from the Ivano-Frankivsk region, that are considered more traditional (more people follow old social behavior). For the humanities, the distribution looks pretty much the same, only reversed (Figure 3). Instead of a larger share of men in the regions, we have a larger share of women.

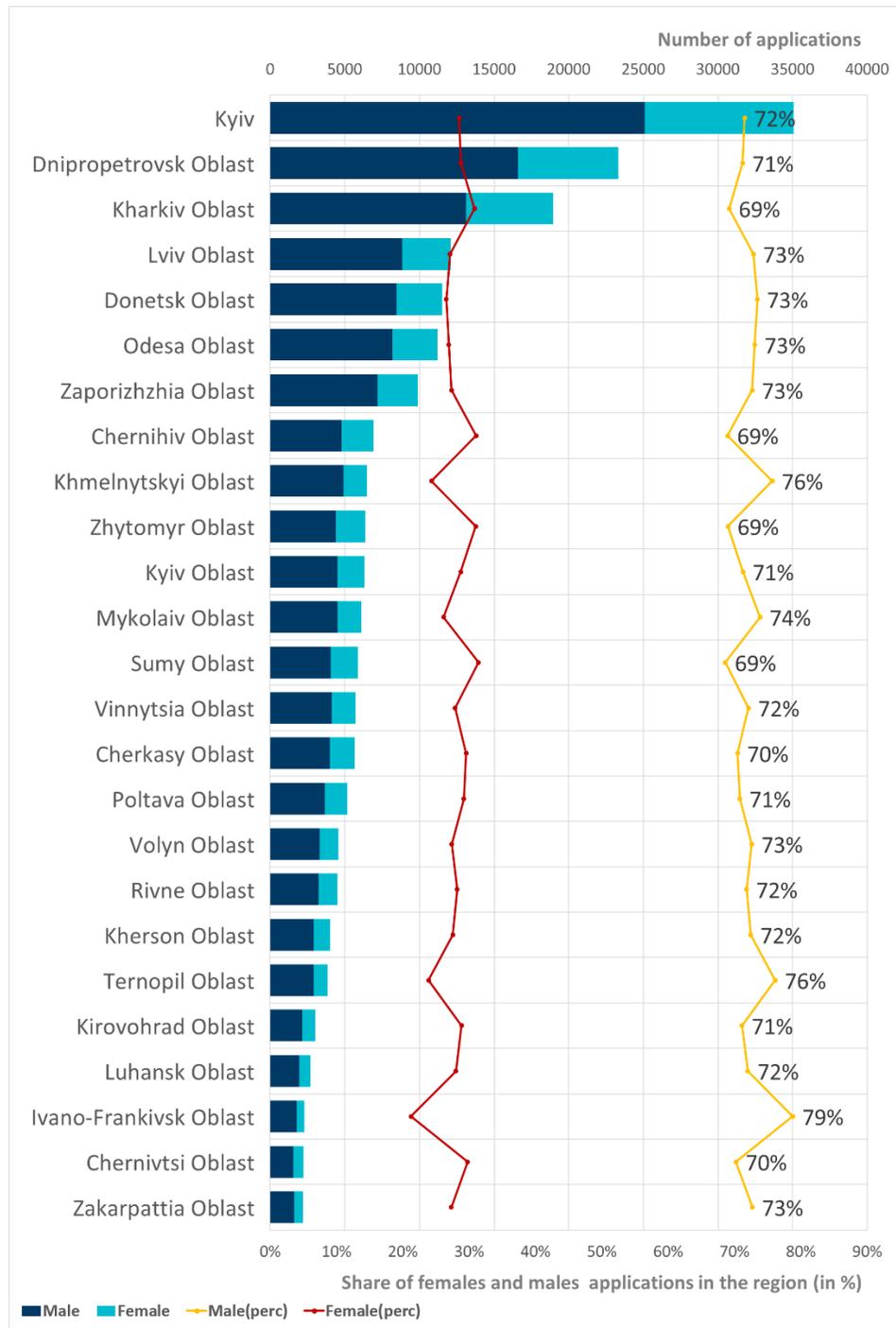


Figure 2. Distribution of STEM applicants by school region and gender

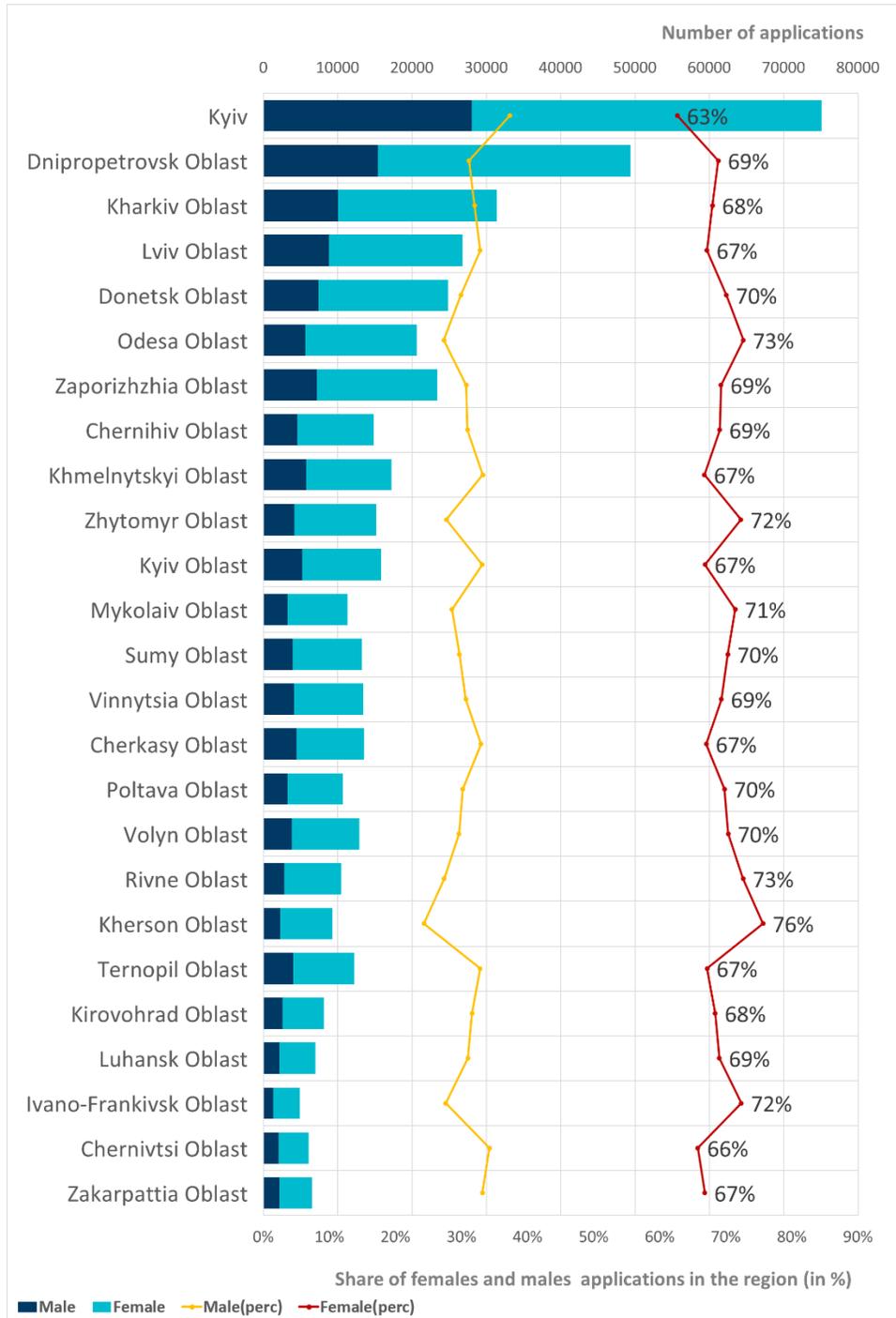


Figure 3. Distribution of Humanities applicants by school region and gender

4.2 ESS and State Statistics Service of Ukraine data

The second part of the data preparation was processing the European social survey. The ESS dataset is considered for determining the level of presence of different social factors in regions in Ukraine. It consists of 2,178 observations and 347 variables.

The most recent results of ESS for Ukraine were released in 2013. Despite a time gap between social factors data and university applications data, it was chosen for research because of the following reasons:

- These are the only publicly available dataset on social factors that are distributed by region and gender;
- Stereotypes, prejudices, and other social factors tend to change slowly in the society;
- Enrollers to the universities that graduated in 2018-2020 are highly likely to be affected by people surveyed in 2013 because these part of society represent the current situation with social factors among adults by region and gender.

So, it is convenient to use ESS data for this topic research.

ESS provides a wide variety of variables that cover such topics as attitudes in society towards politics, culture, media, education, gender equality, and social behavior. Moreover, it presents information about of religion and traditions' importance among regions. Considering the research question of the current thesis, after analysis, different social factors were selected in two groups:

- *S* – aggregated social factors characteristics per region, and per region and gender;
- *L* – aggregated labor force and education level information per region, and per region and gender.

Table 4. S group. Aggregated social factors characteristics per region and gender

Variable	Description	Expected impact
Average level of life satisfaction	The variable shows how on average, surveyed people are satisfied with their lives per gender and region.	People with a higher level of life satisfaction are “freer” and more open in their views. So, they are more likely to choose a technical degree, even if it is against social stereotypes.
LGBT share	The variable shows what share of surveyed people per region have a negative and strongly negative opinion about the LGBT community.	The higher share of people with negative opinions towards the LGBT indicates the presence of prejudice in the region. Therefore, it potentially negatively affects enrollers’ choices.
Religion share	The variable shows what share of surveyed people per region indicates themselves with a high level of religion (7-10 out of 10).	The high level of religion partially represents conservatism, following the stereotypes about women (like, females should have the role of housewife) in society. Thus, the greater value of religion share in a region is likely to have a negative impact on females’ choices and a positive impact on males’.
Traditions share	The variable shows what share of surveyed people per region considered important to follow traditions (5-6 out of 6).	It is expected that a higher share of people that follows traditions negatively affects the possibility that females will consider a technical major.

Data about education level, employment, and wages per gender and region are also present in the ESS dataset, which we proceeded above. However, social surveys are not the best source for this type of information since they represent a too small share of the population, while the state statistics (Ukrstat) contain more precise indicators for all regions. For that reason, we selected new data from the State Statistics Service of Ukraine (Ukrstat 2021) to obtain more accurate information about the wage and employment ratios between genders in regions and the education situation within society in Ukraine. These variables are presented in the L group – aggregated labor force and education level information per region and gender.

Table 5. L group. Aggregated labor force and education level information per region and gender

Variable	Description	Expected impact
Wage Ratio	The variable represents a ratio of women's wages to the man. In particular, it shows how much, on average, females' wages differ from males' per region.	The higher value could indicate the lower level of social pressure caused by stereotypes. Thus, it positively affects the females' possibility to apply for a technical major.
Employment Ratio	The variable represents a ratio of the share of women's employment to the man per region.	The higher value could indicate the absence of social pressure caused by stereotypes. Thus, it positively affects the females' possibility to apply for a technical major.
Share of people with high level of education per region	The variable shows a share of people with high education per region.	The increasing share of the educated population potentially decreases the impact of old prejudice. Therefore, this variable is expected to positively impact enrollers' choices, specifically females.
Share of people with TVET (college) level of education per region	The variable shows a share of people with TVET (college) education per region.	As with the previous one, increasing in this indicator could positively impact enrollers' choices.

Indicators from these groups allowed us to obtain an overview of the situation with social factors in the regions.

The lowest value for the level of life satisfaction and share of women who identify as highly religious are in Kirovohrad Oblast. Despite the low values for previous indicators, the share of females that found it essential to follow traditions is among the highest in this region.

The most satisfied life females are lived in Volyn Oblast. Following this further, the biggest share of highly-religious women is in Ivano-Frankivsk and Lviv Oblast.

We have a similar distribution of pick values among regions for the men. As well as previously, Kirovohrad Oblast has one of the lowest share of males who determine themselves as highly religious. Although, there is a large share of them who follows traditions. Also, a considerable share of highly religious men lives in Lviv Oblast. While in Kherson Oblast, men do not have an elevated level of life satisfaction.

Since the western regions have always been considered more religious, this result is expected. It confirms the general trends in the country.

Zakarpattia Oblast can be characterized as one with the lowest employment ratio, the share of people with high and TVET education. At the same time, Donetsk oblast has the lowest wage ratio.

The most progressive in high education and position with positive opinions towards the LGBT community are Kyiv city and Sumy Oblast, correspondingly. At the same time, the largest share of people with a negative attitude towards diverse communities is in the Volyn oblast.

In addition, Luhansk Oblast seems like a center of TVET education since there is a large share of people with college degrees. Rivne Oblast has the most reasonably balanced employment ratio, which is equal to almost one.

In order to get generalized numerical characteristics of the following factors, descriptive statistics were prepared (Table 6).

Table 6 shows essential statistics for proceeded variables. Overall, for various social indicators where we have gender distribution, females have the highest values for the max column. In particular, there are larger share of women in regions with the high level of following traditions, religion and life satisfaction.

Further, it could be seen that the TVET (college) level of education is more prevalent among regions contrary to the high level (universities).

As for labor indicators, like wage and employment ratio, we received the following results. On average, women officially earn about 20% less than men. Also, the presence of males in the labor market is 17% higher than females.

Table 6. Descriptive statistics for the social data

Variable	Gender	min	max	mean	std dev
Average level of life satisfaction	Female	3.07	7.14	5.09	0.81
	Male	3.48	6.11	5.05	0.60
LGBT share		37.000	94.100	56.264	14.523
Religion share	Female	13.33	75.00	43.19	17.58
	Male	5.41	64.44	24.20	15.83
Traditions share	Female	11.10	91.70	59.06	15.81
	Male	18.75	80.00	48.95	16.82
Wage Ratio		0.624	0.953	0.802	0.069
Employment Ratio		0.675	1.011	0.829	0.063
Share of people with high level of education per region		0.148	0.538	0.259	0.068
Share of people with TVET (level of education per region)		0.189	0.566	0.433	0.089

4.3 Final dataset

To obtain a final dataset, we performed data cleaning, normalization, and matching of two types of data:

- university applications: data with enrollers applications combined with the information about high schools;
- social characteristics data: ESS data and indicators from the State Statistics Service of Ukraine,

The final dataset has 643,694 observations and overall 40 social and individual indicator variables. The model proceeded in the current work includes only variables from three groups ***I***, ***S***, and ***L*** described above.

Chapter 5

ESTIMATION RESULTS

This chapter provides the estimations results of the current research. The logistic regression was used to examine the effect of various social factors on both genders' probability of entering a technical degree in Ukrainian universities for 2018-2020 period. For precise results interpretation, average partial effect (marginal effect) results are also presented in the current section.

Table 7 below shows the logistic model results described in Chapter 3. The first (1) column presents the regression results applied for the enrollers from the final dataset. Further, to increase simplicity in the interpretation and obtain accurate estimations for females and males, we run separate regressions on both genders. Thus, column (2) and (3) contains the results for females and males correspondingly.

Following this further, the current regression result presented in Table 7 gives us only a general overview of the significance and magnitude of the social factors' coefficients of our model.

Nevertheless, it allows us to answer the central question of this thesis work, in particular, about the importance of other components behind applicants' grades. Analyzing the significance of the considered variables, it is pretty straightforward that social factors affect all enrollers and both genders separately. Different individual characteristics of applicants, aggregated values for widespread social factors by region and gender, and highlighted labor force indicators information affect the probability of choosing degree types among high school graduates in Ukraine.

Table 7. Estimation results

	<i>Dependent variable:</i>		
	(1) All	(2) Females	(3) Males
Male (gender)	1.814*** (0.008)		
Rural	0.042*** (0.013)	-0.041** (0.020)	0.106*** (0.017)
Belonging to the boarding school	0.212*** (0.022)	0.333*** (0.035)	0.162*** (0.028)
Average score (school diploma)	0.200*** (0.003)	0.165*** (0.006)	0.223*** (0.004)
Overall score (EIE)	0.003*** (0.0002)	0.006*** (0.0003)	0.001*** (0.0003)
Average level of life satisfaction	-0.110*** (0.007)	-0.027** (0.011)	-0.255*** (0.011)
LGBT share	-0.005*** (0.0003)	-0.004*** (0.0005)	-0.002*** (0.0004)
Religion share	0.002*** (0.0003)	-0.002*** (0.001)	0.003*** (0.0004)
Traditions share	-0.005*** (0.0002)	-0.004*** (0.0004)	-0.002*** (0.0003)
Wage Ratio	-0.969*** (0.058)	-0.318*** (0.092)	-1.433*** (0.077)
Employment Ratio	0.082 (0.071)	0.198* (0.120)	0.491*** (0.094)
Share of people with high educ per region	0.186*** (0.053)	0.540*** (0.087)	-0.491*** (0.072)
Share of people with TVET educ per region	0.290*** (0.058)	0.430*** (0.094)	-0.276*** (0.079)
Constant	-2.596*** (0.089)	-3.832*** (0.156)	0.219* (0.116)
Observations	643,694	364,400	279,294
Log Likelihood	-345,129.300	-155,250.100	-189,442.600
Akaike Inf. Crit.	690,286.700	310,526.200	378,911.300

Note:

* p** p*** p<0.01

However, as mentioned before, to obtain an accurate results interpretation, we need to consider the marginal effect due to the specifics of the logit model. We used the average partial effect method to find how the probability will change if the independent variable increases by one unit. Obtained results for APE are presented in Table 8.

Table 8. Marginal effects for logit model

	<i>Dependent variable:</i>		
	major_binary		
	(1) All	(2) Females	(3) Males
Male (gender)	0.325*** (0.008)		
Rural	0.008*** (0.013)	-0.005** (0.020)	0.026*** (0.017)
Belonging to the boarding school	0.038*** (0.022)	0.043*** (0.035)	0.039*** (0.028)
Average score (school diploma)	0.036*** (0.003)	0.021*** (0.006)	0.054*** (0.004)
Overall score (EIE)	0.001*** (0.0002)	0.001*** (0.0003)	0.0003*** (0.0003)
Average level of life satisfaction	-0.020*** (0.007)	-0.004** (0.011)	-0.062*** (0.011)
LGBT share	-0.001*** (0.0003)	-0.001*** (0.0005)	-0.001*** (0.0004)
Religion share	0.0004*** (0.0003)	-0.0003*** (0.001)	0.001*** (0.0004)
Traditions share	-0.001*** (0.0002)	-0.001*** (0.0004)	-0.001*** (0.0003)
Wage Ratio	-0.173*** (0.058)	-0.041*** (0.092)	-0.348*** (0.077)
Employment Ratio	0.015 (0.071)	0.026* (0.120)	0.119*** (0.094)
Share of people with high educ per region	0.033*** (0.053)	0.070*** (0.087)	-0.119*** (0.072)
Share of people with TVET educ per region	0.052*** (0.058)	0.056*** (0.094)	-0.067*** (0.079)
Constant	-0.465*** (0.089)	-0.497*** (0.156)	0.053* (0.116)
Observations	643,694	364,400	279,294
Log Likelihood	-345,129.300	-155,250.100	-189,442.600
Akaike Inf. Crit.	690,286.700	310,526.200	378,911.300

Note:

*p**p***p<0.01

As it has been pointed out the results of the considered logit model confirmed the preliminary hypothesis suggested in the current thesis work. Various social factors impact female and male applicants' probability of choosing a technical degree in Ukraine. As demonstrated in Table 8 above, in most cases of output results, the

variables' magnitude coincides with the expected effects described in Table 3 – Table 5.

Increasing the share of people with a negative opinion against the LGBT community in society provokes a decrease in the probability of choosing a STEM major. Keeping everything constant, an increase by one unit of the indicator involves decreasing probability by 0.1 percentage points. Even though the value is small, it is statistically significant at a 0.01 level. Such presumptions among the population can potentially indicate intolerance and outdated attitudes towards certain population groups.

Similar in magnitude and value, results were received for the traditions share, which means that an increase in this indicator negatively impacts the considered probability. This variable has the same effect for both genders. One possible interpretation is that people with high adherence to traditions may generally not accept the modern technologies that students in technical specialties currently teach. Such groups could be more sympathetic to non-high-tech majors or some TVET type of education in relation to both women and men.

On the other hand, as was assumed previously, an increase in the enrollers' scores entails a growth in the probability. In particular, keeping everything constant, one unit change in average grade for the high school diploma involves a 2.1 and 5.4 percent point rise in the probability of applying to the STEM degree for females and males correspondingly. The effect of the EIE general score is lower than for high school average grade, but it is also statistically significant at 0.01 level. As was supposed earlier, we got such results because a technical degree required a more thorough knowledge of the school program. Unlike humanitarian subjects, sciences subjects (like Physics, Mathematics, or Chemistry) are challenging to catch up in a short time during, for example, the first months of study at the university. Thereby it is essential to have a sufficient amount of knowledge before. Thus,

considered scores above are good indicators of the level of pre-university preparation.

Following this further, belonging to a boarding school also increased the probability of applying for school graduates. The effect is 4.3 and 3.9 percentage points for women and men. The hypothesis that describes this is that students who live in boarding schools do not being in touch as much with carriers of stereotypes as those who live in families and constantly interact with relatives who could instill some prejudice. On the contrary, children in boarding schools are under specific isolation and are not as influenced by social factors.

Moreover, belonging to the village (rural variable) has oppositely different effects for females and males. Women not from the city are less likely to apply for a technical degree. Since villages in Ukraine are a center of maintaining traditions, a high level of religion, and sometimes various superstitions, especially regarding women, these results are expected. On the other hand, males who graduated from village school are 2.6 percentage points more likely to enroll in STEM majors. This can be explained by the fact that the standard of living in villages is lower than in cities. So, men may seek financial independence faster, so they prefer to enroll in more high-paying fields, such as IT (a part of STEM).

As could be seen, a larger share of people identifying themselves as highly religious decreases the probability of females' choice of technical degree. Religion could be accounted for conservatism in society. While the controversial situation exists for the males. Such values indicate the existence of the "old" views on the female role in the regions.

Unexpected results were received for wage ratio and the average level of life satisfaction. On the whole, these two variables have a negative magnitude. Thus, increasing any of it, keeping other factors constant, would decrease the probability of applying to a STEM degree for both genders. However, these potentially could

appear because of the fact that when the unbalance and unfavorable situation is present in the society, then the sense of justice and the desire to improve the situation begin to work more powerfully. Also, the higher wages in humanities specialties could balance the ratio and attract more applicants to following with it in the universities.

In contrast to the previous one, the employment ratio between the shares of employed females and males in the regions positively affects the probability. Since some stereotypes assume that women should not work, this variable also partially presents an adequate position (free of prejudices) in the region. This is additional evidence in favor of the expected output described in the section above.

The share of people with high and TVET levels of education have pretty similar values. However, both of these variables differ in magnitudes for women and men. As for the first one, a higher share of educated people in the region positively impacts the probability of enrolling in a technical major. While for the males, we have opposite results.

CONCLUSIONS AND POLICY RECOMMENDATIONS

6.1 Conclusions

In this thesis, we performed an estimation to discover whether different social factors impact applicants' choices. For this purpose, we use three sources of information: EDBO, ESS, and the State Statistic Service of Ukraine. The logistic model results show that various aggregated factors affect enrollers' major choices in regions, both for women and men. Further, an analysis of the output was conducted.

As it has been pointed out above, social factors have a statistically significant impact on recent school graduates applying for a university degree. Moreover, the most vulnerable category of applicants is women from rural areas. They are heavily influenced by high-level religions, customs, and potentially stereotypes. The estimation results provided more evidence about the females' role in society. People are prejudiced by stereotypes and separate male and female specialties.

Looking closer to the results of the coefficient on LGBT share, traditions share variables, we proved that the conservative view is still widely present among Ukrainians.

To better determine the impact of social factors, this study lacks personalized data on applicants. Or at least a more comprehensive sample of more recent research on opinions in the regions.

However, all things considered, the statistical significance of these variables is a signal that indicates the existence of social pressure on the applicants during their school education.

6.2 Policy recommendations

There are considerably more men in the technical degree in Ukraine (71.7% versus 28.3% females). At the same time, the percentage of students of both gender in STEM majors in Ukraine overall is relatively lower in non-technical fields. According to the UNDP Gender Profile of Ukraine, among all tertiary graduates, there are 38% of men and 14% of women with technical majors (UNDP 2020). Since the technological sphere is fundamental for the country's development, small involvement in this sector could affect the country's rate in the global index of innovations. This way, it is crucially essential to increase scientific potential in Ukraine.

The Government of Ukraine has recognized this issue, and to accelerate and stimulate the interest and participation of schoolers in a technical industry, the Decree of the President of Ukraine № 31/2020 was implemented. The document announced the 2020-2021 study year as the "Year of Mathematics." In particular, the development of mathematical expertise of students for application in social, economic, technological, scientific, and other spheres of public life was set as one of the main goals.

This is a promising step on the part of the state. It is vital to continue further and implement various policies that promote the development of the students.

However, it is essential also to put attention on the social opinions and attitudes among Ukrainian population. One of the possible options could introduction additional lectures in the school on different socially significant topics. In particular, it is crucial to promote children's education in schools and show different potential opportunities for development regardless of gender.

WORKS CITED

- Aphichoke Kotikula, Das Smita. 2019. "Gender-Based Employment Segregation: Understanding Causes and Policy Interventions". *World Bank, Jobs Working Paper; No. 26*. <https://openknowledge.worldbank.org/handle/10986/31510>
- Black Sandra E., Cortes Kalena E., Lincove Jane Arnold. 2015. "Apply Yourself: Racial and Ethnic Differences in College Application". *NBER, Working paper 21368*. <https://www.nber.org/papers/w21368>
- Delaney Judith and Devereux Paul J. 2020. "Gender Difference in College Applications: Aspiration and Risk Management". *Economics of Education Review, 80, 102077*. <https://www.iza.org/publications/dp/13973/gender-differences-in-college-applications-aspiration-and-risk-management>
- DOU. 2020. Women in IT: Portrait, Career and Salary. Analytics. <https://dou.ua/lenta/articles/it-woman-2020/>
- Ministry of Education and Science (MES). 2018. News. <https://mon.gov.ua/ua/news/divchata-stanovlyat-79-vstupnikiv-na-gumanitarni-nauki-ta-mistectvo-i-lishe-23-na-inzhenerni-specialnosti-liliya-grinevich-na-2-mu-uzhk>
- Mykhailyshyna Daria. 2018. "The fate of the graduate. What factors affect the results of external evaluation.". <https://voxukraine.org/longreads/zno/index.html>
- OECD. 2019. PISA 2018 Results (Volume II): Where All Students Can Succeed, PISA, *OECD Publishing, Paris*. <https://doi.org/10.1787/132219b1-en>
- Osvita. 2020. Areas of knowledge and specialties of higher education. <https://osvita.ua/vnz/76723/>
- Sakhno Hanna, Yuzkiv Victoria, Kobernik Oleksandra, Myrasova Olga, Robeiko Anastasiia. 2021. "How much will Ukrainian girls win? From entering the it specialty?". <https://ces.org.ua/benefit-from-the-increase-in-girls-involvement-in-it/>

- Samokhin Ihor, Kohut Iryna. 2017. "Socio-economic factors in the inequality of learning outcomes in high school".
<https://cedos.org.ua/researches/sotsialno-ekonomichni-factory-u-nerivnosti-rezultativ-navchannia-u-serednii-shkoli/>
- Speer Jamin. 2017. "The Gender Gap in College Major: Revisiting the Role of Pre-College Factors". *Labour Economics*, vol. 44, issue C, 69-88, 2017.
<https://www.sole-jole.org/assets/docs/16332.pdf>
- Ukrstat. 2021. "Labour Force of Ukraine".
http://www.ukrstat.gov.ua/druk/publicat/kat_u/2021/zb/07/zb_r_s_2020.pdf
- UNDP. 2020. Gender Profile of Ukraine.
<https://www.undp.org/ukraine/gender-profile-ukraine>
- Volosevych Inna, Podnos Varvara. 2021. "What are Ukrainian men and women guided by when choosing a profession: the results of a national survey on key factors and stereotypes".
https://ukraine.unfpa.org/uk/gender_career_choices
- Wooldridge, J.M. 2018. *Introductory Econometrics: A Modern Approach*. Michigan State University: Seventh Edition, Cengage Learning, Boston, USA.

APPENDIX A.

Table A-9. University applications datasets description

Dataset	Number of observations	Variables
bach_2018	830211	<ol style="list-style-type: none"> 1) enrll_id – enrollers id 2) inst_id – university id 3) inst_name – university name 4) region – university region 5) compet_score – overall enroller’s score
bach_2019	819544	<ol style="list-style-type: none"> 6) subj_1 – the name of EIE subject 1 7) score_1 – score of EIE subject 1 8) subj_2 – the name of EIE subject 2 9) score_2 – score of EIE subject 2 10) subj_3 – the name of EIE subject 3 11) score_3 – score of EIE subject 3
bach_2020	777793	<ol style="list-style-type: none"> 12) average_score_prev_inst – average score of high school transcripts 13) specialty – chosen specialty 14) gender 15) school_name – high school name 16) enrll_year – enrollment year