TALLINN UNIVERSITY OF TECHNOLOGY
School of Business and Governance
Department of Economics and Finance

Kevin Orion

IMPACT OF SKILLS TO GENDER WAGE GAP IN ESTONIA

Master's thesis

Programme TAAM, specialisation Economic Analysis

Supervisor: Simona Ferraro, PhD

Tallinn 2021

I hereby declare that I have compiled the thesis independently
and all works, important standpoints and data by other authors
have been properly referenced and the same paper
has not been previously presented for grading.
The document length is 12,053 words from the introduction to the end of conclusion.

Kevin Orion ……………………………
(signature, date)
Student code: 163624TAAM
Student e-mail address: kevin.orion07@gmail.com

Supervisor: Simona Ferraro, PhD
The paper conforms to requirements in force

……………………………………………
(signature, date)

Chairman of the Defence Committee:
Permitted to the defence
…………………………………
(name, signature, date)

TABLE OF CONTENTS

ABSTRACT

The aim of the thesis is to investigate the differences in skills among adult individuals and between genders in Estonia, country with the highest gender wage gap in Europe, to understand how these differences might drive wages and contribute to explaining the gender wage gap. The focus will be on skill use and cognitive skills. Men are using numeracy, planning and information technology skills more at work than women. Women are more endowed in literacy skill, men in numeracy and problem-solving. Although, the differences in skill use and cognitive skills between genders are small. The author uses ordinary least squares regression, quantile regression and the Oaxaca – Blinder decomposition to show that skills do have an impact on the gender wage gap based on the Programme for the International Assessment of Adult Competencies dataset. Women are more rewarded for their use of reading and and information technology skills; men are rewarded more for their cognitive numeracy. Controlling for skills, the explained part of the gender wage gap increases by two percentage points. Oaxaca – Blinder decomposition results show that some of the explained part of the gender wage gap is driven by the higher endowment of cognitive numeracy skill of men.

Keywords: skills, human capital theory, gender wage gap, Estonia, Oaxaca – Blinder decomposition, PIAAC

INTRODUCTION

"*Men and women are not the same. And they will not be the same. That does not mean that they can not be treated fairly.*"
Jordan B. Peterson

Gender wage gap is a thoroughly studied topic in social sciences. Yet, there is still room for further analyses to understand its origins and causes. In the recent years, many studies have been published approaching the gender wage gap from different angles: from socioeconomics (Kleven et al, 2019) to psychology (Heckman & Kautz, 2012). With the rise of available data at hand, the understanding of the gender pay gap has grown but it still begs to be scrutinised.

Since it is quite a controversial topic, it has been a pressing issue in the media as well. On the surface, there seems to be inequality looming at the expense of women in the developed countries[1]. Estonia is known to have the largest gender wage gap in the Europe Union[2]. Since gender wage gap is topical in Estonia and is subject to frequent media attention, and Statistics Estonia even routinely publishes dedicated statistics on the topic, it was author's interest to delve deeper into it and investigate it more. Although, gender wage gap has been extensively studied in Estonia by authors (Anspal & Rõõm, 2010; Halapuu, 2015), the Programme for the International Assessment of Adult Competencies (PIAAC) dataset and a different approach that takes into account cognitive skills and skill use may give additional value to the debate.

The human capital is the capital that individuals have accumulated during their lifetime. It takes forms as knowledge, skills and competencies, attributes that the individuals can take advantage of to create economic value. Studies on human capital show how and which skills influence labour productivity. Earnings are believed to be dependent on labour productivity, thus knowledge, skills,

[1] According to Eurostat's 2019 data (table sdg_05_20) women's gross hourly earnings were on average 14.1% lower than those of men in the Europe Union
[2] According to Eurostat's 2019 data (table sdg_05_20) women's gross hourly earnings were on average 21.7% lower than those of men in Estonia

competencies and attributes, in turn, drive earnings. It is important to understand what affects earnings in order to analyse the gender wage gap.

The aim of the thesis is to investigate the differences in skills among Estonian individuals and between genders to understand how these differences might drive wages and contribute to explaining the gender wage gap. The focus will be on skill use and cognitive skills.

The thesis seeks to answer the following questions:

1. How large is the difference in skills between genders?
2. How might different skills predict wage levels?
3. How might skills affect the gender wage gap?

The following hypotheses are tested:

H_1: The skills that predict higher wages are more prominent in men than women.

H_2: Controlling for skills reduces the unexplained part of the wage gap.

The first chapter provides an overview of the relevant literature that was used in writing the thesis. Firstly, the author describes the history of human capital theory and then provides the literature. Secondly, the author provides different literature on the gender wage gap. Thirdly, the author gives insight into how skills and tasks may affect productivity of the labour force and the gender wage gap. Lastly, the author provides overview of the gender wage gap studies that are conducted on Estonia.

The second chapter of the thesis firstly presents the framework of the Survey of Adult Skills. Then, the author introduces the dataset, the variables and provides descriptive statistics of the data. The third sub-chapter will explore the methods used in the thesis: the ordinary least square regression (OLS), the quantile regression and the Oaxaca – Blinder decomposition. Lastly, the results of the analyses are presented by method type.

The third chapter of the thesis provides the conclusions summing up the results of the analyses and comparing the findings with other works.

I would like to thank my supervisor, Simona Ferraro for her support and help; and Merike Kukk for her support in finding the way to write the thesis. I am responsible for the content and for any errors that the thesis may contain.

1. LITERATURE OVERVIEW

In this first chapter of the thesis, the author provides an overview of the literature on human capital, gender wage gap, separately for Estonia, and skills and tasks.

1.1. Human capital theory

“The most valuable of all capital is that invested in human beings.” (Marshall, 1920)

The definition of human capital has probably its roots in the 18th century when Adam Smith, ahead of his time, noted that individuals accumulate skills and abilities that in turn fuel the economic activity. The labour was usually primarily seen as a collective mass intended for physical work and not for individualistic capabilities. It was not widely recognised that once learned skills, as a way of capital, were carried on by the individuals or as Adam Smith put it: “a capital fixed and realised, as it were, in his person”. It was not until the 1960s when economist started to incorporate the inherent abilities into a new kind of capital- human capital. (OECD, 2007)

Gary Becker, believed to be one of the three founders of human capital theory, writes in a foreword to the Oxford Handbook of Human Capital (Burton-Jones & Spender, 2012) that he started in the 1950s with a modest goal to calculate the rates of return on investments in education. In his prior research on the discrimination against minorities in American labour markets, he had seen that earnings of individuals of different ethnicity increased with years of schooling. The theory that grew out of his work is the opposite of modest. He wrote a book in 1960 on the theory called “Human Capital”. He was awarded the Nobel Prize in Economics in 1992 for “having extended the domain of microeconomic analysis to a wide range of human behaviour and interaction, including nonmarket behaviour.” (Nobel Media AB, 2021). Human capital theory is widely known in economics, although, Becker is not the only one who has contributed. James Mincer wrote two highly influential books in 1958 (“Investment In Human Capital and the Personal Income Distribution”) and 1974 (“Schooling, Experience, and Earnings”) which were seen as breakthroughs. Theodore William Schultz cemented the phrase “human capital” in his 1960’s article “Capital Formation by Education”: “I propose to treat education as an investment in man and to treat its consequences as a form of capital. Since education becomes a part of the person receiving it, I shall refer to it as *human capital.*” (Schultz 1960, 571). Schultz was later awarded the Nobel Prize in Economics in 1979 for research into human capital theory and economic development (Nobel Media AB, 2021).

Human capital is defined by the Organisation for Economic Co-operation and Development (OECD) as the knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic well-being. (OECD, 2007) Human capital theory tries to answer a question “why do dissimilar people earn disparate levels of income?”. The human capital theory states that higher income is gained by higher productivity, which is itself led by a greater amount of human capital.

Human capital theory can be easily visualised in a three-way causal chain: Human capital → Productivity → Income. However, not all agree with the latter. A criticism of the theory sees it as too simplistic and states that there is no agreed upon definition of human capital. Also, the criticism posits that productivity differences cannot be objectively compared. Blair suggests to use social hierarchy rank as a more relevant variable to explain differences in income. (Blair, 2018)

In addition to the importance of education and work experience (Mincer, 1974), Mincer also emphasised the significance of on-the-job training. He found that in terms of costs, training was as important as formal education for the male labour force (Mincer, 1962). With age comes the gain of experience and biological growth but also the decline of the latter. Productivity will thus increase with age up to certain point when it will start to degenerate through biological decline. (Mincer, 1958) It has been noted by Mincer and Polachek that being away from the labour market depreciates earning power. The higher the amount of human capital the higher the depreciation rate. However, women without children suffer from a lower depreciation rate probably because most of the time that is spent away from working is spent on job seeking. (Mincer & Polachek, 1974)

1.2. Gender wage gap

Men and women are on average of the same intellectual capability as measured in intelligence quotients (IQ) tests. However, there tend to be differences in cognitive skills, usually at the extremes of the skills distributions. Men outperform in advanced mathematics, women gravitate toward better results in language usage, reading comprehension and writing. (Halpern & LaMay, 2000) Blau & Kahn (2017) reviewed the literature on psychological attributes, like risk aversion and the propensity to compete or negotiate, and their impact on the gender wage gap. They found that gender differences in psychological attributes can explain a small to moderate part of the gender wage gap. (Blau & Kahn, 2017)

Female labour market participation is affected by both marital status and the presence of children. Earnings are not affected by marital status but are affected, most probably, by the fact of having children. Three times out of four, having a child has a negative influence on female earnings and in 25% of cases has a positive effect. The earnings setback can be up to 25%. (Dolton & Makepeace, 1987)

Women suffer a wage penalty when they have children even when controlled for human capital. Mothers with one child suffer a four percent wage penalty while mothers of two or more children face a 12% drop in wages. (Waldfogel, 1997) Using administrative data from Denmark, it was found that the impact of having children is large on women but almost non-existent on men. The earnings penalty is estimated to be around 20%. Having children influences hours worked, occupation, sector and firm choices. The unexplained gender wage gap is mainly caused due to children, but the authors also indicate that the unexplained gap could contain an element of discrimination against having children. (Kleven et al., 2019) There is a minor gender working hours gap before having children but after the first birth, there is a substantial migration of female workforce from full-time work to part-time work or even out of the labour force. Even after the children have grown up and left home, there is a larger propensity for women to work part-time than men. (Paull, 2008)

Dias et al. (2018) argue that closing the gender gaps in the proportions of full-time and part-time work between men and women would not, fully, solve the gender wage gap. They conclude that working part-time also explains lower women’s wages. This is due to work experience in full-time employment that leads higher wages whilst in part-time employment does not. (Dias et al., 2018)

Other factors that drive the gender wage gap are firm-specific wage premiums and the bargaining power of employees. Women tend to work more than men at firms that offer lower wage premiums to both genders. At the same time, women receive 90% of their male colleagues' wages at equivalent firms. (Card et al., 2015)

One of the most egalitarian occupations in the United States (US) is pharmacy. It has a low gender wage gap where female pharmacist earn four to seven percent less than their male counterparts. This can be explained by the features of the profession: uniform training, substantial use of information technology and standardisation of products. All of the mentioned allows for an increased substitutability among the employed. Work experience was found to have a little effect on earnings. There seems to be sizeable penalty in working fewer hours and the share of pharmacists working shorter hours has risen together with the relative earnings of females to males. The earnings of pharmacists are mostly a function of their working hours. (Goldin & Katz, 2016)

Men and women work in different occupations and this segregation is usually seen as an advantage for men. However, Blackburn et al., (2001) argue that the measure of segregation must be divided into two parts: horizontal and vertical, in order to assess any inequalities. Otherwise, there will be illogical findings like egalitarian Sweden where there is a high level of segregation and a low gender pay gap. (Blackburn et al., 2001)

Horizontal segregation measures difference without inequality while vertical segregation measures occupational inequality through difference in pay and social stratification. It was found that the segregation measures have an inverse relation to them: the higher the horizontal segregation, the lower the inequalities between genders. On the other hand, the latter could also mean that there is increased discrimination against women the more they compete with men. The authors also found that men are ahead in terms of remuneration but women are more highly situated in social stratification. Women work with more highly educated people, have wider access to higher status networks and are generally working on healthier occupations. (Jarman et al., 2012)

1.3. Skills and tasks

Based on the US labour market data, it was found that between 1968 and 1990, the returns to cognitive and people skills increased whilst the return motor skills decreased. The gender wage gap shrank 20% in the 80s while women were employed more in cognitive- and people- intensive occupations compared to men. (Bacolod & Blum, 2010) Men are materially more endowed in motor skills than women but the differences in cognitive skills are small. The decline of returns to motor skills in 1980-2000 makes up for the extensive amount in the drop of the gender wage gap. During the 2000's, the quicker growth of women's cognitive skills compared to men caused the gender wage gap to shrink. (Yamaguchi, 2016)

It has been shown that computer technology can replace workers who perform routine tasks. On the other hand, the technology can complement workers performing non-routine tasks that require problem-solving skills, complex communications, creativity and flexibility. This will increase the demand for more highly educated labour force. (Autor et al., 2003) With the introduction of personal computers (PCs) from 1980s to 2000, markets that had a wider educational gap experienced a drop in the gender wage gap. This could be explained by the fact that female labour participation is higher in industries where soft skills are needed and PCs are complementary to them and substitute to hard skills. (Beaudry & Lewis, 2012)

On the other hand, it has been shown that robotisation may lead to a raise in the gender wage gap. Based on 20 European countries from 2006 to 2014, it was shown that a 10% expansion in robotisation would yield a 1.8 percent increase in the gender wage gap. This rise in the wage gap stems from countries with higher gender wage gaps and due to the expansion of men's earnings in semi- to high-skilled occupations. (Aksoy et al., 2020) Over many decades, there has been an increase in earnings for more highly educated individuals relative to high school graduates. This phenomenon can be seen as a progressively increased return to skills over time. The demand for skills is linked to the progress of technology and, in particular, to the skill bias of technical change. With the technological progress, there has been an increased relative demand for skills. On the other hand, there has also been a growth in the supply of skills because public investment in schooling and willingness to acquire education have risen. (Acemoglu & Autor, 2010)

Social skills have become more important in the labour market. Based again on the US data. The return to social skills was higher in the 2000s than in the 1980s or 1990s. Teamwork furthers labour productivity via comparative advantage as employees differ in their abilities to perform the wide array of workplace tasks. Higher social skills enable workers to trade tasks with each other more easily and thus contributing to efficiency. Cognitive skills and social skills have complemented each other. Conversely, it was found that cognitive skills have become less important over time. (Deming D. J., 2017)

Another study found that higher paying jobs require higher levels of cognitive and social skills in tandem. The authors looked at the following skills in the US job vacancies from 2010 to 2015: cognitive, social, character, writing, customer service, project management, people management, financial, computer (general), computer (specific). Their measure of job skills add explanatory power to their analysis and they conclude that it explains about 12% of variance in wages across enterprises. (Deming & Kahn, 2018) Personality traits like the Big Five (conscientiousness, openness to experience, extraversion, agreeableness, neuroticism), in addition to cognitive skills, affect the economic success of individuals. These psychological traits each affect productivity uniquely in different tasks, although cognitive capability is homogeneously more important in increasingly complex tasks. (Heckman & Kautz, 2012)

Evidence from the PIAAC dataset showed that higher cognitive skills in numeracy, literacy and problem-solving are tied to higher wages in all 23 participating countries. It was found that returns to skills effect differed greatly between countries: some countries experience twice the returns to skills compared to the lower returns to skill countries. (Hanushek et al., 2015) Based on the Austrian PIAAC dataset, it was found that when controlled for cognitive skills (literacy and numeracy), skill use (influencing, planning, writing, numeracy, reading, information and communication technology), skill matching and flexibility of working hours, the unexplained part of the wage gap decreases by six to nine percentage points across the earnings distribution. (Christl & Köppl–Turyna, 2020)

1.4. Literature for Estonia

Gender wage gap has been topical in Estonia and it has been under investigation for decades. It has also been a subject of policy analysis, (Rõõm & Kallaste, 2004). Probably the most comprehensive overview in Estonia on the literature on gender wage gap studies has been done by Anspal et al., (2009). The Estonia's labour market is the most segregated labour market in Europe, both horizontally as segregation into different occupations and, vertically such as segregation into different occupation levels. Moreover, typical to Nordic and Baltic countries, female employment rate is very high. Men prefer jobs in construction, transportation and storage, agriculture, forestry and fishing sectors while women prevail in health care and social work, education, financial and insurance activity, accommodation and food services. (Anspal & Rõõm, 2010)

Over the last decades, the trend of a larger proportion of men acquiring higher education has reversed and women have overcome men in education. If education would be the only factor influencing wages, women would earn more than men in Estonia. (Anspal & Rõõm, 2010)
Estonian women are the 3rd most educated in Europe after Finland and Iceland[3]. The gender education gap is quite wide in Estonia: 51.8% of women have a higher degree while only 31% of men. Women with children earn on average 1.2 percent less per child than their childless counterparts. The same does not apply to men. The country also has the highest gender pay gap in Europe, larger in the private sector. Also, the pay gap increases with the size of the enterprise. (Anspal & Rõõm, 2010)

In a recent study by Vahter and Masso (2018), foreign owned firms have also larger gender wage gaps than domestic firms in Estonia. This could be explained by the fact that foreign enterprises have higher work commitment requirements and, on average, men tend to be more willing to work overtime, longer hours or during inconvenient hours. Also, women with young children have a higher wage penalty working in a foreign firm than working in a domestic firm. (Vahter & Masso, 2018) Moreover, it was found that firm-level factors are able to explain 35% of the gender wage gap. Results showed that due to the sorting effect, men are working more in high-wage enterprises and women more in low-wage enterprises. In addition, women receive lower pay than men in the same firm due to the bargaining effect. (Masso et al., 2020)

In Estonia, adults scored better on average in literacy and numeracy and below average in problem solving compared to other survey participants. On the other hand, higher educated people are the least proficient in problem solving compared to other countries equivalents. Information and Communication Technology (ICT) skill use frequency is higher in Estonia than the average but fewer people use computers in everyday work (64% to 69% average). Aggregate results analysis on country level found that education drives skill levels. Problem-solving skills and employment

[3] According to Eurostat's 2019 data (table edat_lfs_9903) showing tertiary education holders among 25-64-year olds

levels are also highly correlated. On the overall, Estonia is mostly more homogenous than other countries with smaller gender and education discrepancies. Differences between genders in skills are rather minor. (Halapuu & Valk, 2013)

Halapuu, in 2015, estimated a gender wage gap of 21.3% of which 30.2% was the explained component, applying the Oaxaca-Blinder decomposition. For men, higher wages were associated with higher numeracy proficiency and work autonomy/flexibility. For women, ICT skill use and highest education level were associated with higher wages. Quantile regression results showed a small correlation between numeracy and literacy with wage levels across the quantiles and, no relation in the higher part of distribution. On the contrary, problem solving had a strong effect in the higher wage levels. Quantile regression also showed a larger wage gap in the upper part of the distribution of earnings. (Halapuu, 2015)

Tverdostup and Paas (2016) constructed a non-parametric Ñopo-type decomposition based on PIAAC data of Nordic countries (Estonia, Finland, Denmark, Norway and Sweden). This decomposition approach matches individuals with similar observable characteristics. It allows to compare and group individuals and explore the distribution of unexplained differences in the outcome. It also recognises that men and women are different in their endowments of characteristics. Ñopo posits that "not all males are comparable to all females" and that men are endowed in such sets of characteristics which have no female counterpart and that these sets are highly rewarded in the labour markets. (Ñopo, 2008)

Results from their study for Estonia, Tverdostup and Paas (2016) showed 44% of matched females and 48% of matched males. Non-matched males had a lower education than matched males. On the contrary, non-matched females had a higher education than matched females. However, regression results revealed that education is not statistically significant for both groups of men and that non-matched women experience a positive effect to earnings from education. It was rather numeracy that drove wages in case of men, meaning low association of formal education and skill levels. (Tverdostup & Paas, 2016)

Gender pay gap is more uneven in the unmatched gender specific sub-groups than in the matched group. It seems that unmatched female unique characteristics predict lower earnings than unmatched male group. This leads to conclusion that female-specific characteristics are less valued at the labour market. (Tverdostup & Paas, 2016)

2. ANALYSIS

In this part of the thesis, the author introduces the data, provides descriptive statistics and the methodology, conducts the analyses and summarises the results.

2.1. The Survey of Adult Skills

The Survey of Adult Skills, part of OECD's PIAAC, assesses the proficiency of adults in information-processing skills essential in the knowledge-based economies and societies of the 21st century. These skills are essential for accessing, understanding, analysing and using information. They are considered to be key information-processing skills as they are:

- necessary for fully integrating and participating in the labour market, education and training, and social and civic life.
- highly transferable, in that they are relevant to many social contexts and work situations.
- learnable and therefore subject to the influence of policy.

It is better for both individual and social growth that individuals with higher skills are more productive at work. Skills used at work can be divided into two: "own skills" that the individuals possess and, the "job skills" that are required by jobs. The extent of using skills is driven by the motivation of the workers and by the requirements of specific jobs. PIAAC measures the proficiency of skills and also their usage.

The first cycle of the survey lasted from 2008 to 2019 and included three separate rounds of data collection. In the first round in 2011-2012, data was collected from 24 countries. The second round took place in 2014-2015 when further 9 countries took part. The third round collected data from 6 countries in 2017-2018. The countries in the first round were: Australia, Austria, Belgium (Flanders region only), Canada, the Czech Republic, Cyprus, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Japan, Korea, the Netherlands, Norway, Poland, the Russian Federation, the Slovak Republic, Spain, Sweden, the United Kingdom (England and Northern Ireland only) and the United States. The round 2 countries: Chile, Greece, Indonesia (Jakarta region only), Israel, Lithuania, New Zealand, Singapore, Slovenia and Turkey. The round 3 participants were: Ecuador, Hungary, Kazakhstan, Mexico, Peru and the United States. (OECD, 2019) The second cycle of the survey has begun with results to be published in 2023. The survey is implemented by interviewing around 5,000 adults aged from 16 to 65 in each country. It is conducted either via computer or pencil-and-paper. The survey is designed to be comparable internationally and cross-culturally. (OECD, n.d.)

The survey assesses individuals and their levels of literacy, numeracy and problem solving in technology-rich environments. The results were represented on a 500-point scale which should mirror the mastery of the domain. (OECD, 2019) Literacy is defined as: “the ability to understand, evaluate, use and engage with written texts to participate in society, to achieve one’s goals, and to develop one’s knowledge and potential. Literacy encompasses a range of skills from the decoding of written words and sentences to the comprehension, interpretation, and evaluation of complex texts. It does not, however involve the production of text (writing).” (OECD, 2019)

Numeracy as: “the ability to access, use, interpret and communicate mathematical information and ideas in order to engage in and manage the mathematical demands of a range of situations in adult life. To this end, numeracy involves managing a situation or solving a problem in a real context, by responding to mathematical content/information/ideas represented in multiple ways. (OECD, 2019) Problem solving in technology-rich environments is: “the ability to use digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks. The assessment focuses on the abilities to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, and accessing and making use of information through computers and computer networks.” (OECD, 2019)

In addition to the proficiency tests, the PIAAC background questionnaire asked individuals about the use of skills in both their professional and private life. The questionnaire used a task-based approach (also named job requirement approach) to measure skill use frequency. This method was adopted to produce as unbiased as possible resulting indicators compared to the actual skills of the respondents. The questionnaire aimed to capture information about different clusters of tasks that the respondents performed in their current or last job in previous 12 months. The obtained information was broken up to different activities and tasks: information-processing (cognitive) skills, technological skills and the generic job tasks like interaction, learning, organisation and physical/motor activity. Respondents were questioned about the frequency, on the scale of one (never) to five (every day), and complexity of different activities. Then, twelve skill use indices (see Appendix 1) were derived using item response theory (IRT). These are continuous variables that represent the levels of use of the skill. All the indices were standardised across the sample of all countries to have mean equal to 2 and standard deviation equal to 1. The aim was for 90% of the observations to be between values 0 (low frequency) and 4 (high frequency). (OECD, 2016)

2.2. Data and descriptive statistics

PIAAC data for Estonia contains 7,632 observations and 1,328 variables. The author used the international PIAAC codebook (OECD, 2020) to navigate the many variables which also had coded names. The codebook provided the type (numerical or categorical) and range of expected answers if these were categorical variables. Out of the 7,632 people who participated in the study, 3,464 were males and 4,168 females. All of the data manipulation was done with the free programming language R using RStudio. Results were produced with the relevant packages.

The author chose the relevant variables (see the full list in Appendix 2) from the dataset to conduct the analyses. Many observations did not include any value or contained a value that was of no use for the analyses. Participants of the survey could skip questions or refuse to answer or could answer "Don't know". In such cases, the author equalised the answers to missing values. The chosen set of variables contained numerical and categorical variables. The descriptive statistics of the dataset with the numerical variables is presented in the Table 1. The data with categorical variables are visualised in the appendices (from Appendix 3 to Appendix 6) as bar plots.

Table 1. Descriptive statistics

	Males					Females				
Variable	**Obs.**	**Min**	**Mean**	**S.D.**	**Max**	**Obs.**	**Min**	**Mean**	**S.D.**	**Max**
Age	3,464	16.00	39.91	14.25	65.00	4,168	16.00	41.63	14.32	65.00
Work experience	3,185	0.00	19.54	13.00	54.00	3,891	0.00	19.75	13.31	51.00
Hourly earnings	1,711	0.17	7.32	39.37	1,621.80	2,340	0.47	4.75	13.61	618.12
Hours worked per week	2,448	1.00	41.35	11.77	125.00	2,891	1.00	37.98	10.82	120.00
Number of children	3,421	0.00	1.33	1.23	4.00	4,148	0.00	1.48	1.16	4.00
Influencing skills	2,452	0.06	1.95	1.02	5.58	2,878	0.06	1.98	1.03	5.58
Numeracy skills	2,239	-0.09	2.02	1.00	6.05	2,567	-0.09	1.93	0.92	6.05
Planning skills	2,535	0.25	2.23	0.97	3.73	3,010	0.10	2.15	0.94	3.73
Reading skills	2,544	-0.96	1.96	1.00	7.02	2,994	-0.96	2.01	0.98	7.02
Writing skills	1,981	0.06	1.65	0.85	5.80	2,474	0.06	1.69	0.79	5.80
ICT skills	1,464	0.16	2.17	1.11	5.46	2,049	0.01	2.07	1.00	5.46
Numeracy score	3,432	73.77	275.14	44.22	412.50	4,154	98.15	269.89	40.09	397.95
Literacy score	3,432	116.84	274.53	42.27	393.32	4,154	118.90	276.54	40.07	388.21
Problem-solving score	2,339	148.39	279.07	40.20	405.97	2,896	153.30	274.83	39.49	402.35

Source: author's calculations based on valid observations of PIAAC data for Estonia

The mean gross hourly earnings is 2.5 Euros or 52% higher for men (7.32 EUR) than for women (4.75 EUR). The average age is 1.7 years higher for women but the average years of work experience is practically the same (19.5 years for men and 19.8 years for women). On average, women tend to have more children than men among these survey participants. Men, on average, work per week 3.4 hours more than women. It seems there are men who work more than the regular

40-hour work week bringing the mean up over 40 hours. On the other hand, some women seem to work part-time weighing the mean value down to 38 hours per week. This is backed by the bar graph in Appendix 4. Larger proportions of women are part-time employed and fulfilling domestic duties than men.

The largest mean value difference in skill use indices[4], which show the frequency of skill use, between genders is in ICT. The mean index value is 2.17 for men and 2.07 for women. The median value is higher for men with the distribution skewed more towards the right (see Figure 1). The Figure 1 shows the distribution and median values of skill use indices where men seem to use numeracy as a skill more at work, with a mean value difference of 0.09 points. The distribution of numeracy index values for men is skewed more to the right than with a higher median value. Although planning mean value is higher for men, the median value is equal. There is a higher population of women around the median value but a higher population of men around values of 3 and 4. Women tend to use reading, writing and influencing skills more often than men. The largest distinction is in reading skills where the gap is 0.05 points. The distribution of women in reading skill is visibly more skewed to the right than of men. The same can be said for writing skill. The use of influencing skill is the most similar among men and women. The distributions are the most alike of all the skills and the median value is barely different.

The PIAAC dataset contained 10 plausible score results for each of the participants in literacy, numeracy and problem-solving. The author calculated a mean value for each observation based on the 10 plausible score values. The presented mean values (in Table 1) of cognitive skills show the mean values of 10 plausible score values. Cognitive score results mirror the skill differences between men and women. On average men have scored better in numeracy, problem-solving and women better in literacy[5]. Men have also scored both the lowest and the highest scores. The most prominent discrepancy in cognitive skills is in numeracy where men have surpassed women by 5.25 points on average. The distribution of numeracy scores of men has clearly shifted more to the right with a higher median value (see Figure 2). The shift can also be noticed with distribution of problem-solving skill, although the differences are less pronounced. Men scored higher by 4.24 points. On the other hand, women outdid men in literacy on average by 2.01 points. The

[4] Welch t-test concluded that the mean differences between genders in numeracy, planning, ICT are statistically significant and in influence, reading, writing are not

[5] Welch t-test concluded that the mean differences between genders in numeracy, literacy and problem-solving are statistically significant

distribution of literacy scores of women is slightly more skewed to the right. There seems, however, to be a higher population of men who scored the best results (350 – 400 points) and the worst results (100 – 175 points).

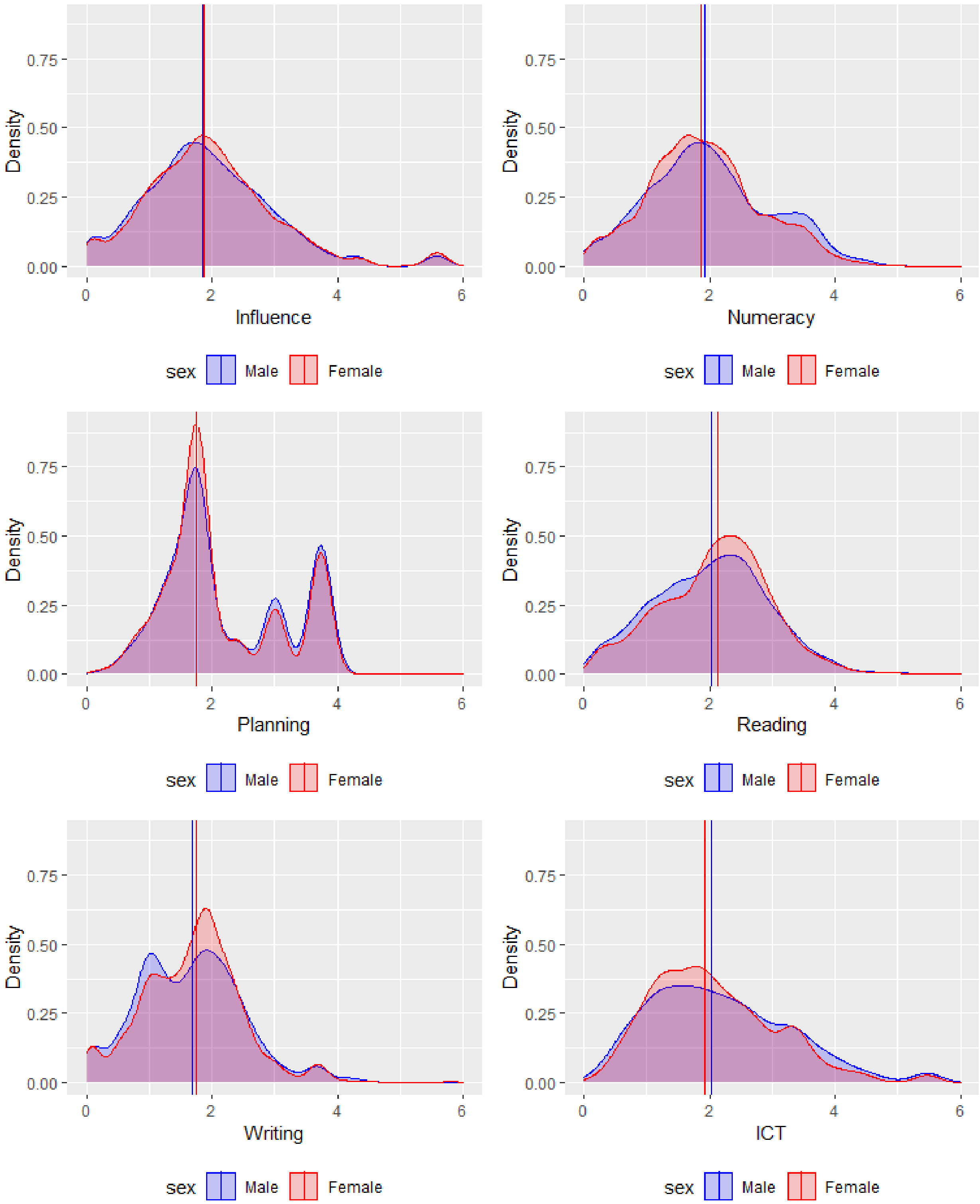


Figure 1. Kernel density estimate: gender differences in skill use at work

Source: author's calculations based on valid observations of PIAAC data for Estonia

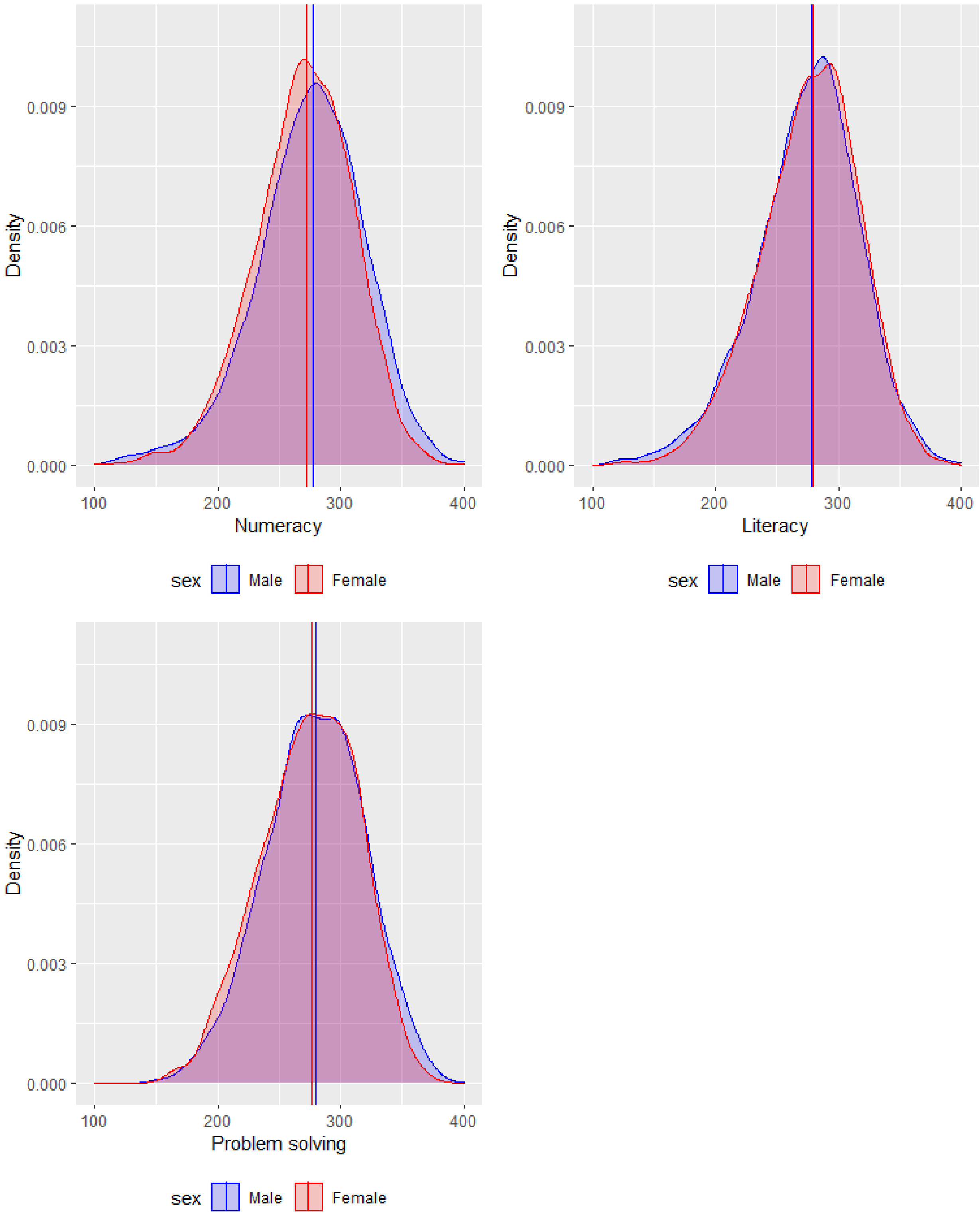


Figure 2. Kernel density estimate: gender differences in cognitive skills (score)

Source: author's calculations based on valid observations of PIAAC data for Estonia

Cognitive skill scores correlate quite strongly and positively with each other (see Figure 3). Literacy and numeracy (Lit_cog & Num_cog) have a correlation coefficient of 0.87, literacy and problem-solving of 0.83 (Probsolv_cog), numeracy and problem-solving of 0.78. The skill use indices do not correlate strongly with each other, the highest correlation coefficient is 0.44 between ICT and numeracy. The weakest relationship is between numeracy and writing (0.27). Surprisingly, there are no strong correlations between cognitive skills and skill use. Unexpectedly, literacy has an almost absent correlation with reading (0.08) and writing (0.07). Reading and writing correlate between themselves moderately (0.41) but have non-existent correlations with other cognitive skills as well. The highest correlation between skill use and cognitive skills is amidst ICT and problem-solving (0.27). ICT also has a stronger relationship with the other cognitive skills than any other skill use index. Even numeracy (skill use) and numeracy (cognitive) have a meager correlation of 0.2. Based on the correlations it seems both skill use and cognitive skill results supplement each other in terms of use in analyses as they seem to behave differently and thus, avoid the multicollinearity problem.

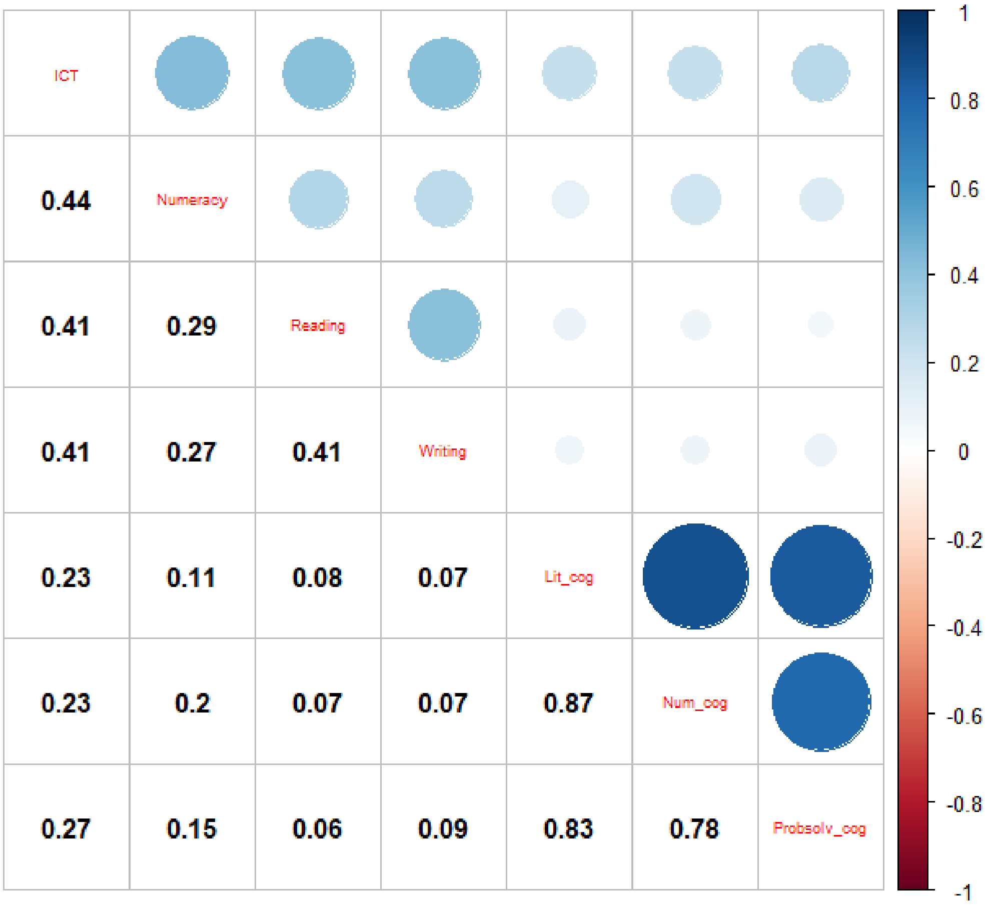


Figure 3. Correlations between skill use and cognitive skills (score)

Source: Source: author’s calculations based on valid observations of PIAAC data for Estonia

Other used variables (categorical), but are not listed in the descriptive statistics table (Table 1), are education level, number of children, employment status, skill based occupational classification, firm size, work flexibility, industry and economic sector.

The classification of levels of education is based on the UNESCO's (United Nations Educational, Scientific and Cultural Organization) standard which is abbreviated as ISCED (International Standard Classification of Education) (UNESCO, 2012). The standard divides education into nine levels:

- Level 0: early childhood education
- Level 1: primary education
- Level 2: lower secondary education
- Level 3: upper secondary education
- Level 4: post-secondary non-tertiary education
- Level 5: short-cycle tertiary education
- Level 6: bachelor's or equivalent
- Level 7: master's or equivalent
- Level 8: doctoral or equivalent

On the overall, women are more highly educated (see Appendix 3). A larger share of women than men have acquired either master's, bachelor's degrees or have completed short-cycle tertiary education or post-secondary non-tertiary education. Over one third of women have higher (tertiary) education versus one fourth of men. The number of children variable was already demonstrated in the descriptive statistics table (Table 1) but it is also presented in the appendices (see Appendix 3) as a categorical variable. The author derived the variable from number of children and whether the participants had children at all. If participants had no children then, the fact is relayed as "0". This approach of combining number of children and the fact whether participants even had children provides an easy way to compare the two different populations. A larger population (35%) of men have no children than women (25%).

In Appendix 4, the employment statuses of the participants are visualised. The most noticeable differences are in the full-time employed, part-time employed and fulfilling domestic tasks categories. Men are more full-time employed whilst more women are part-time employed and fulfilling domestic tasks or looking after children/family. However, there is not much of a difference in the distributions of hourly gross wages between the whole population and only full-time employed sub-set (see Figure 4). This probably due to that non-employed participants did not report to have any income and thus both distributions are mostly populated by full-time employed people. The author decided not to use the employment status as a variable in the analyses since the distributions are similar and using the variable most likely does not give any value. Nearly half of the women in the sample are employed as skilled workers and half of men as blue-collar workers (see Appendix 4). One quarter of women are white-collar workers but only around 10% of men are employed in white-collar jobs.

Men and women work almost in the same proportions of different firm sizes (see Appendix 5). The minute differences come from the smallest firm size (1 to 10 people) where more of men work than women and in medium size (51 to 250 people) where it is vice versa. Most of the individuals are employed in firms up to 250 employees. Employees in firms with over 251 workers make up only 10% of the population. It seems men can enjoy flexible working hours to a higher extent (see Appendix 5). Around 13% of men work flexible hours to a very high extent and 6% of women. One fourth of men and one third of women cannot work flexible hours. PIAAC data included four variables on work flexibility: sequence of tasks, how to do the work, speed of work and working hours. For this thesis, the latter of the four was chosen because it has been shown that flexible workings hours is one of the most widely used flexible working arrangement. (Wheatley, 2017)

Both genders are employed in different industries (see Appendix 6). Men seem to work in more physically demanding and/or technical industries like agriculture, forestry, fishing; mining and quarrying; manufacturing; electricity, gas, steam and air conditioning supply; water supply, sewerage, waste management and remediation activities; construction. Women, on the other hand, seem to work more in people-oriented or abstract fields like activities of extraterritorial organizations and bodies; other service activities; arts, entertainment and recreation; human health and social work activities; education. Over 80% of men work in the private sector whilst 63% of women do so (see Appendix 6). Nearly twice more women work in the public sector than men. Only a few percent of men and women are employed in non-profit organisations.

2,186 men (72%) live with their spouse and 2,577 (70%) of women; 983 men (31%) reported to have had a on-the-job training in the last year. 1,512 women (47%) had an on-the-job training in the previous year.

On the overall, men have higher wages as is evident from the Kernel density plot (Figure 4) and from descriptive statistics (Table 1). The male distribution of earnings is more shifted to the right with higher mean and median values. Notable visual difference between the overall population and the full-time employed sub-set, is that the tip of the male population in the full-time sub-set is higher than of the overall population. This is indicating a larger number of observations around the mean in the full-time sub-set than in the overall set. There seems to be slightly larger population of women, than men, around the higher end of the distribution, near the 32 euro wage level.

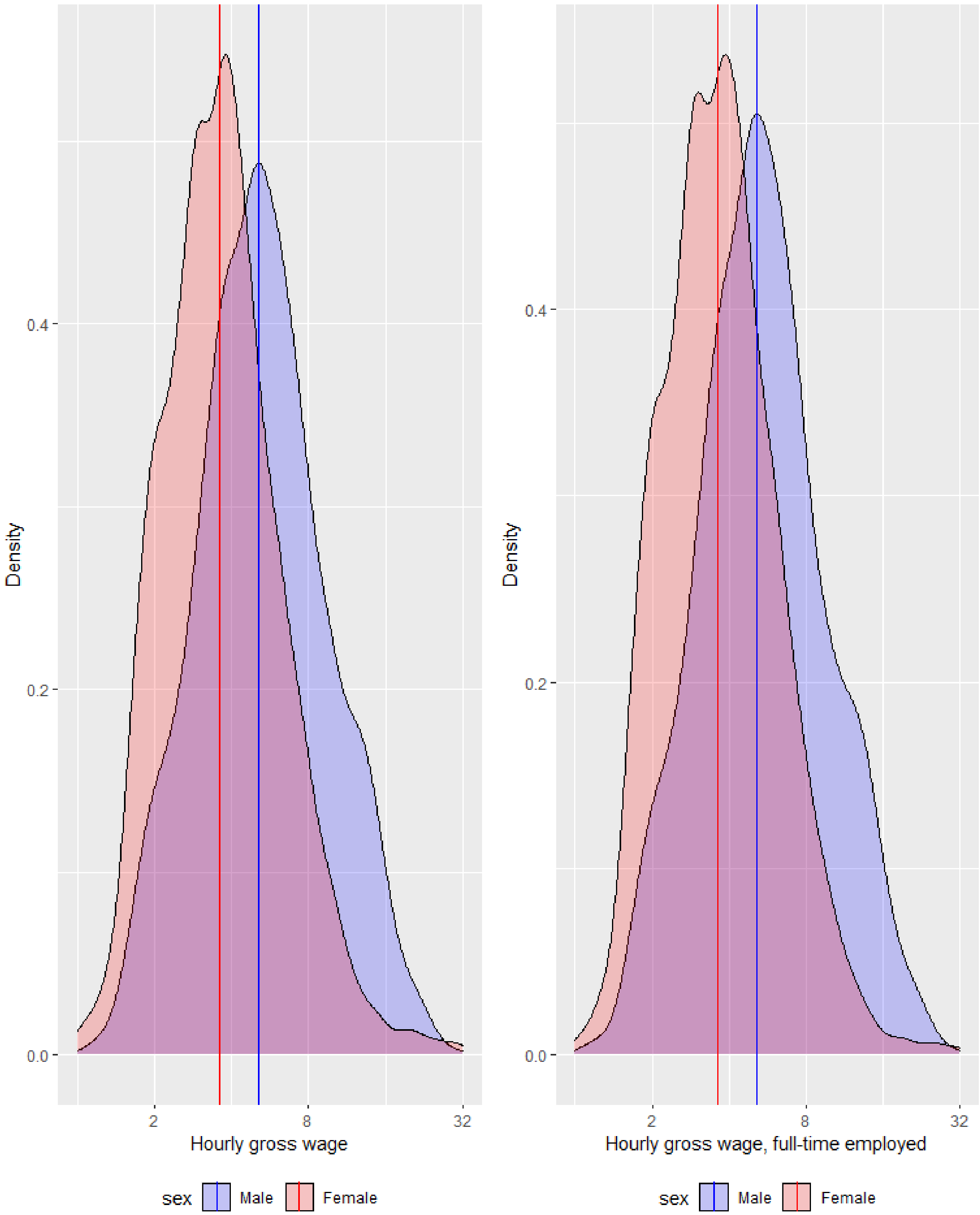


Figure 4. Kernel density estimate of hourly gross wages: all observations and full-time employed sub-set

Source: Source: author's calculations based on valid observations of PIAAC data for Estonia

2.3. Methodology

In order to assess the gender wage gap and the impact of skills to the gap, the author uses three different methods. Firstly, the ordinary least squares regression, then the quantile regression and thirdly, the Oaxaca – Blinder decomposition.

The author has considered the variables used in the analyses, based on the Mincerian equation, where the dependent variable is natural logarithm of wage and independent variables are gender dummy, experience (working years) and education, and the pertinent available variables from the PIAAC dataset. The OLS regression equation, based on the Mincerian equation and control variables, is as follows:

$$log\ wage = \alpha + \beta_0\ male\ dummy + \beta_1\ working\ years + \beta_2\ working\ years^2 + \beta_3\ education + \beta_4\ control\ variables + \varepsilon_i \quad (1)$$

log wage: natural logarithm of hourly gross wage,
α: intercept,
β: coefficients,
ε_i: error term.

In the regression analyses part, the author expands the set of variables by adding the focus variables: skill use and cognitive skills. The final OLS equation is as follows:

$$log\ wage = \alpha + \beta_0\ male\ dummy + \beta_1\ working\ years + \beta_2\ working\ years^2 + \beta_3\ education + \beta_4\ control\ variables + \beta_5\ focus\ variables + \varepsilon_i \quad (2)$$

where
log wage: natural logarithm of hourly gross wage,
α: intercept,
β: coefficients,
ε_i: error term.

By adding the focus variables separately allows to assess the relevancy and the impact of the chosen variables to the gender wage gap analysis. The models include the dummy variable for gender which is “1” in case of males and “0” in case of females. This dummy variable allows to assess the impact gender has on earnings. The author is expecting that by expanding the models, the explanatory power of the models grows and the coefficient of the gender dummy variable shrinks. The latter is important to focus on since the coefficient of the dummy variable shows how much men are paid more, on average, than women in the models just based on their gender. It shows how much the other variables in the models are not able to explain the impact of gender on wages.

The OLS regression estimates the mean coefficients and so does the Oaxaca – Blinder decomposition. Namely, it can only estimate the average differences but not the distribution (Dolton & Makepeace, 1985). To address this limitation of the OLS regression and the Oaxaca – Blinder decomposition, the author also conducts quantile regression analyses. It allows to see how focus and control variables may affect earnings throughout the earnings distribution. The quantile regression allows to divide the sample into segments and estimate the coefficients of the variables at different quantiles. Firstly, the author performs quantile regression including only the focus variables with the following equation:

$$log\ wage = \alpha_\theta + \beta_\theta\ focus\ variables + \varepsilon_\theta,$$

$$Q_\theta(w|focus\ variables) = \alpha_\theta + \beta_\theta\ focus\ variables \quad (3)$$

where
log wage: natural logarithm of hourly gross wage,
α: intercept at θ-quantile,
β_θ: coefficient at θ-quantile,
Q_θ: conditional θ-quantile of the distribution of w,
ε_i: error term at θ-quantile.

Secondly, the author performs quantile regression including the focus and control variables following the equation:

$$log\ wage = \alpha_\theta + \beta_\theta\ focus\ variables + \beta_{1\theta}\ control\ variables + \varepsilon_\theta,$$

$$Q_\theta(w|focus\ variables + control\ variables) = \alpha_\theta + \beta_\theta\ focus\ variables +$$

$$\beta_\theta\ control\ variables \quad (4)$$

where
log wage: natural logarithm of hourly gross wage,
α: intercept at θ-quantile,
β_θ: coefficient at θ-quantile,
Q_θ: conditional θ-quantile of the distribution of w,
ε_i: error term at θ-quantile.

Lastly, the author uses the Oaxaca – Blinder decomposition, introduced by (Oaxaca, 1973), & (Blinder, 1973). It is probably the most used method in gender wage gap studies. It allows to compare the outcome of a variable of interest, that is the log of gross hourly earnings, between two groups. The method also divides the difference in the outcome into two distinct parts: the explained part and the unexplained part. The explained part shows how much of the wage gap is due to the differences in x's (endowments). The unexplained part shows how much of the wage gap is due to the differences in β's (the effects of endowments). In the literature, the unexplained part is sometimes viewed as discrimination against one group but mostly it is seen as the phenomenon that the set of variables at hand is not able explain the gap enough and including other variables could help.

Since it is known, from the descriptive statistics part, that there are differences in the wages of men and women, and variation in the endowments, then it can be formulated that the difference in wages is due to distinctions in levels of endowments and the effects that the levels of endowments have on wages. The gap in mean wages is equal to:

$$\bar{W}^{males} - \bar{W}^{females} = \beta^{males}\bar{x}^{males} - \beta^{females}\bar{x}^{females} \quad (5)$$

where
$\bar{W}^{males}$: mean wages of males,
$\bar{W}^{females}$: mean wages of females,
$\bar{x}^{males}$: mean endowments of males,
$\bar{x}^{females}$: mean endowments of females,
β^{males}: coefficient estimates for males,
$\beta^{females}$: coefficient estimates for females.

The author follows the modified Oaxaca – Blinder decomposition proposed by Neumark (1988) using the coefficient estimates from pooled data regression. Neumark called the pooled estimators as “no-discrimination wage structure”. The equation is as follows:

$$\bar{W}^{males} - \bar{W}^{females} = \Delta x\beta^{p} + [\bar{x}^{males}(\beta^{males} - \beta^{p}) + \bar{x}^{females}(\beta^{p} - \beta^{females})] \quad (6)$$

where
$\bar{W}^{males}$: mean wages of males,
$\bar{W}^{females}$: mean wages of females,
$\bar{x}^{males}$: mean endowments of males,
$\bar{x}^{females}$: mean endowments of females,
β^{males}: coefficient estimates for males,
$\beta^{females}$: coefficient estimates for females,
Δx: differences in endowments,
β^{p}: pooled coefficient estimates.

2.4. Results

In this section, the author presents the empirical results from the OLS regression models, quantile regressions and Blinder-Oaxaca decompositions.

2.4.1. Results of OLS regression models

The author starts with a simple OLS model that is based on the Mincer equation and, includes all the other added control variables (see Table *2*).

Table 2. OLS regressions

	Model 1	Model 2	Model 3	Model 4
Male (dummy)	0.302*** (0.021)	0.291*** (0.023)	0.275*** (0.024)	0.250*** (0.026)
Age	0.001 (0.010)	0.007 (0.011)	0.010 (0.013)	0.010 (0.014)
Age^2	-0.0002 (0.0001)	-0.0002 (0.0001)	-0.0002 (0.0002)	-0.0002 (0.0002)
Working years	0.019*** (0.005)	0.020*** (0.006)	0.011* (0.007)	0.012* (0.007)
Working years^2	-0.0003** (0.0001)	-0.0003** (0.0001)	-0.0002 (0.0001)	-0.0002 (0.0002)
ISCED 3	0.076** (0.030)	-0.012 (0.037)	0.027 (0.065)	-0.001 (0.068)
ISCED 4	0.101** (0.042)	-0.013 (0.051)	0.0004 (0.074)	-0.039 (0.078)
ISCED 5	0.175*** (0.035)	0.040 (0.043)	0.063 (0.067)	0.007 (0.070)
ISCED 6	0.274*** (0.052)	0.136** (0.058)	0.147* (0.075)	0.085 (0.079)
ISCED 7	0.443*** (0.036)	0.248*** (0.046)	0.265*** (0.068)	0.183** (0.072)
Children (1)	0.045 (0.030)	0.038 (0.032)	0.045 (0.035)	0.060* (0.036)
Children (2)	0.071** (0.031)	0.053 (0.034)	0.044 (0.037)	0.057 (0.038)
Children (3)	0.095** (0.037)	0.085** (0.042)	0.088* (0.045)	0.111** (0.047)
Children (4)	0.114** (0.045)	0.051 (0.053)	0.184*** (0.059)	0.165** (0.064)
Living with a spouse	0.071*** (0.022)	0.056** (0.025)	0.028 (0.028)	0.021 (0.029)
Working hours	-0.011*** (0.001)	-0.011*** (0.001)	-0.014*** (0.001)	-0.014*** (0.001)
Private sector	0.182*** (0.033)	0.171*** (0.037)	0.202*** (0.040)	0.185*** (0.042)
On-the-job training	0.162*** (0.018)	0.137*** (0.020)	0.096*** (0.022)	0.094*** (0.024)
Blue collar	-0.060** (0.027)	-0.056* (0.032)	0.043 (0.046)	0.060 (0.049)
White collar	-0.219*** (0.025)	-0.244*** (0.028)	-0.195*** (0.031)	-0.186*** (0.033)
Small firm (11-50)	0.075*** (0.022)	0.090*** (0.025)	0.128*** (0.027)	0.137*** (0.029)
Medium firm (51-250)	0.098*** (0.026)	0.113*** (0.029)	0.172*** (0.032)	0.171*** (0.033)
Large firm (251-1,000)	0.127*** (0.037)	0.149*** (0.040)	0.162*** (0.044)	0.175*** (0.045)
Largest firm (1,000-…)	0.176*** (0.054)	0.181*** (0.057)	0.253*** (0.065)	0.266*** (0.067)
Little flexibility	0.112*** (0.024)	0.113*** (0.028)	0.091*** (0.031)	0.094*** (0.033)
Some flexibility	0.148*** (0.023)	0.133*** (0.026)	0.119*** (0.029)	0.118*** (0.031)
High flexibility	0.238*** (0.028)	0.233*** (0.031)	0.213*** (0.034)	0.215*** (0.036)
Very high flexibility	0.226*** (0.040)	0.189*** (0.043)	0.142*** (0.045)	0.146*** (0.047)
Cognitive literacy		0.001 (0.001)		0.0005 (0.001)
Cognitive numeracy		0.001* (0.001)		0.002** (0.001)
Cog. problem-solving		0.0001 (0.001)		-0.0005 (0.001)
Planning			0.003 (0.015)	0.004 (0.016)
Reading			0.037* (0.019)	0.050** (0.020)
Influence			0.038*** (0.015)	0.041** (0.016)
Numeracy			0.002 (0.013)	-0.0001 (0.014)
Writing			-0.005 (0.016)	-0.006 (0.018)
ICT			0.052*** (0.013)	0.043*** (0.014)
Constant	1.100*** (0.182)	0.638*** (0.227)	1.027*** (0.258)	0.618** (0.291)
Controlled for industry	Yes	Yes	Yes	Yes
Observations	3,519	2,646	1,917	1,708
R^2	0.365	0.378	0.403	0.420
Adjusted R^2	0.357	0.366	0.387	0.400

Note: *p**p***p<0.01

This Model 1 will be a baseline used in evaluating the impact of skills in the upcoming OLS analyses and Oaxaca – Blinder decompositions. In Model 1, men are getting paid 30.2% more than women and the coefficient is statistically significant. Age is not statistically significant, while working years is. Each working year predicts 1.9 percent higher salary, on average. On the other hand, age squared variable has a negative impact on the earnings and it refers to the diminishing returns to productivity with age.

All of the education levels are statistically significant and predict higher wages over lower levels of education. The positive impact increases with the level of education. Upper secondary education is predicting 7.6 percent higher wages, post-secondary education 10.1%, tertiary education 17.5%, bachelor's degree 27.4% and master's degree 44.3%. Having children is associated with higher pay than not having them. Having a child is not statistically significant but having two to four or more children are statistically significant. Having two children is affiliated with 7.1 percent higher pay, three children with 9.5 percent and four or more children with 11.4%. On the other hand, (Mincer & Polachek, 1974) found that usually the number of children variable is negative but not statistically significant. However, it may become statistically significant among women with stronger labour-force attachment or with small groups of highly educated women. Living with a spouse variable is statistically significant and has a beneficial effect on wages (7.1 percent). Surprisingly, every working hour is associated with 1.1 percent lower earnings. This could stem from that lower earning employees are working longer, for example due to longer shifts.

Being employed in the private sector is associated with 18.2% higher pay. On-the-job training predicts 16.2% higher hourly earnings. Blue and white collar variables are both statistically significant and relates with lower earnings: -6 percent and -21.9% accordingly. It is explained by the fact that the base against which the collars variables are compared is populated mainly by skilled occupation observations and elementary occupation have mostly dropped out. The hourly wages, on average, increase with the size of the firm. Workers employed in a firm with 11 to 50 people get paid 7.5 percent more, workers in firms with 51 to 250 employees get 9.8 percent more, workers in firms with 251 to 1,000 employees get 12.7 % more and workers in firms with over 1,000 employees get 17.6% more than employees of firms with 1 to 10 people.

Flexibility of working hours is statistically significant and higher amount of flexibility is associated with higher earnings. This is in line with the findings of (Shepard III et al., 1996) who found that flexible working hours can contribute to increased productivity. Even very little amount of flexibility is associated with 11.2% higher earnings. Flexibility to some extent predicts 14.8% higher wages, flexibility to a high extent predicts 23.8% and flexibility to a very high extent predicts 22.6% higher wages. In addition, the model is controlled for industry. However, the list of industries is not added to Table 2 to save space. The full list of industries can be seen in the appendices (see Appendix 6). The adjusted coefficient of determination in Model 1 is 35.7% and the number of observations is 3,519.

In Model 2, the author adds the three cognitive skill scores: literacy, numeracy and problem-solving. The male dummy coefficient drops to 29.1%. However, only numeracy is statistically significant (at 0.1 significance level) out of the three added variables. Each 10 points, out of 500 in the numeracy score, is associated with one percent earnings increase. The adjusted coefficient of determination grew slightly to 36.6% but the number of observations dropped to 2,646.

In Model 3, the author adds the five skill use indicators to Model 1: planning, reading, influence, numeracy, writing and ICT. The male dummy coefficient drops to 27.5%. Out of the six skills, influence and ICT are statistically significant at 0.01 significance level and reading at 0.1 significance level. All of these three skills are associated with a premium to wages. One point increase in the skill use frequency index of ICT skill offers a 5.2 percent increase in wages, one point in influence 3.8 percent increase and one point in reading 3.7 percent increase. The adjusted coefficient of determination grew slightly to 38.7% but the number of observations dropped to 1,917.

Skill use indicators explain better the earnings structure than cognitive scores do. Although, the number of observations contracted, Model 3 was a better fit than Model 1 and Model 2. Model 3 had the smallest male dummy coefficient and the highest adjusted coefficient of determination. In the next step, the author adds both cognitive skills and skill use to the OLS to see how the model might change.

Piling together all of the control variables and the focus variables makes for the best OLS model: Model 4. Its adjusted coefficient of determination is the highest – 40% and the male dummy coefficient is the smallest – 25%. However, the number of observations dropped to 1,708. Half of the focus variables remained statistically significant: ICT at 0.01 significance level; reading, influence and cognitive numeracy at 0.05 significance level. Each 10 points of cognitive numeracy score offer two percent wage premium. Reading is associated with 5 percent higher wages, ICT with 4.3 percent and influence with 4.1 percent. Most of the other variables are still statistically significant, except for age and living with a spouse. Only the highest level of education is statistically significant. Having one, three or four children are statistically significant but for some reason having two children is not. Having children still predicts higher wages over having no children. Working years is statistically significant at 0.1 significance level. Higher firm size predicts higher earnings. Higher flexibility also predicts higher wages but the highest bump in wages is associated with a high extent of flexibility and not the highest level of flexibility.

2.4.2. Results of quantile regressions

The author conducted quantile regression analyses, with the R package "quantreg", in order to assess the impact that the focus and control variables have on wages throughout the earnings distribution.

Firstly, the author presents the quantile regression analyses results on the focus variables: skill use (see Figure 5) and cognitive skills (see Figure 6). The coefficients in the figures are presented as round dots if they are statistically significant at the 0.05 level, otherwise they are presented as triangles. Secondly, the author presents two quantile regressions with both focus and control variables included. One regression is on the male population (see Appendix 7) and the other is on the female population (see Appendix 8).

The coefficients of influence and numeracy are not statistically significant throughout the earnings distribution (see Figure 5). Planning, for men, is statistically significant only at the 20th percentile level where it is associated with a premium of 8 percent per one unit of skill use frequency. On the other hand, for women, planning is associated with a penalty of 4 percent at the 20th and 25th percentile level. Surprisingly, reading is only statistically significant for women and is so throughout the distribution. Reading skill use offers a premium of around 10% throughout the earnings distribution and in the 90th percentile over 15%. Writing skill is not statistically for men and mostly not for women either. The 35th, 40th and 45th percentiles are statistically significant and associated with around 5 percent higher income. The ICT skill use is statistically significant throughout the distribution for women and mostly for men as well, except 10th, 80th, 85th and 90th percentile. The ICT skill is rewarded increasingly higher going up the earnings distribution of women. In the lower parts of the earnings distribution, the skill is associated with 8 to 11% higher earnings. In the middle part it is associated with 13 to 16% higher earnings and in the higher earnings part with up to 18%. On the overall, men are rewarded less for the ICT skill throughout the earnings distribution. In the lower part, men get a premium of 5 to 10%. In the middle part the earnings advantage is around 10% and it tops at 12% in the 70th percentile.

Cognitive numeracy skill is mostly statistically significant throughout the earnings distribution, except for women in the 10th to 20th percentiles and for men in the 90th percentile (see Figure 6). On the overall, men gain more from this skill than women. On the other hand, men are decreasingly rewarded for the numeracy skill going up the earnings distribution. In the lower parts of the distribution, men gain benefit of around 6 percent to income per 10 points on the cognitive skill score. In the middle part, this benefit shrinks to around 5 percent and in the higher part, to around 4 percent. Women tend to gain around 2.5 percent throughout the earnings distribution and 3 to 4 percent in the higher part of the distribution increase in wages from each 10 points in numeracy score. Literacy is mostly not statistically significant except for in the lower to middle part of the distribution of earnings of women. The skill is associated with a premium of 2.5 to 4 percent to income. Men have only two statistically significant points which are 25th and 45th percentile where the skill is associated with a 2.5% penalty on wages. Problem-solving skill is not statistically significant, except in the 60th and 65th percentile of women's earnings distribution. There it is associated with a 1.25 percent increase in wages for each 10 points in the score.

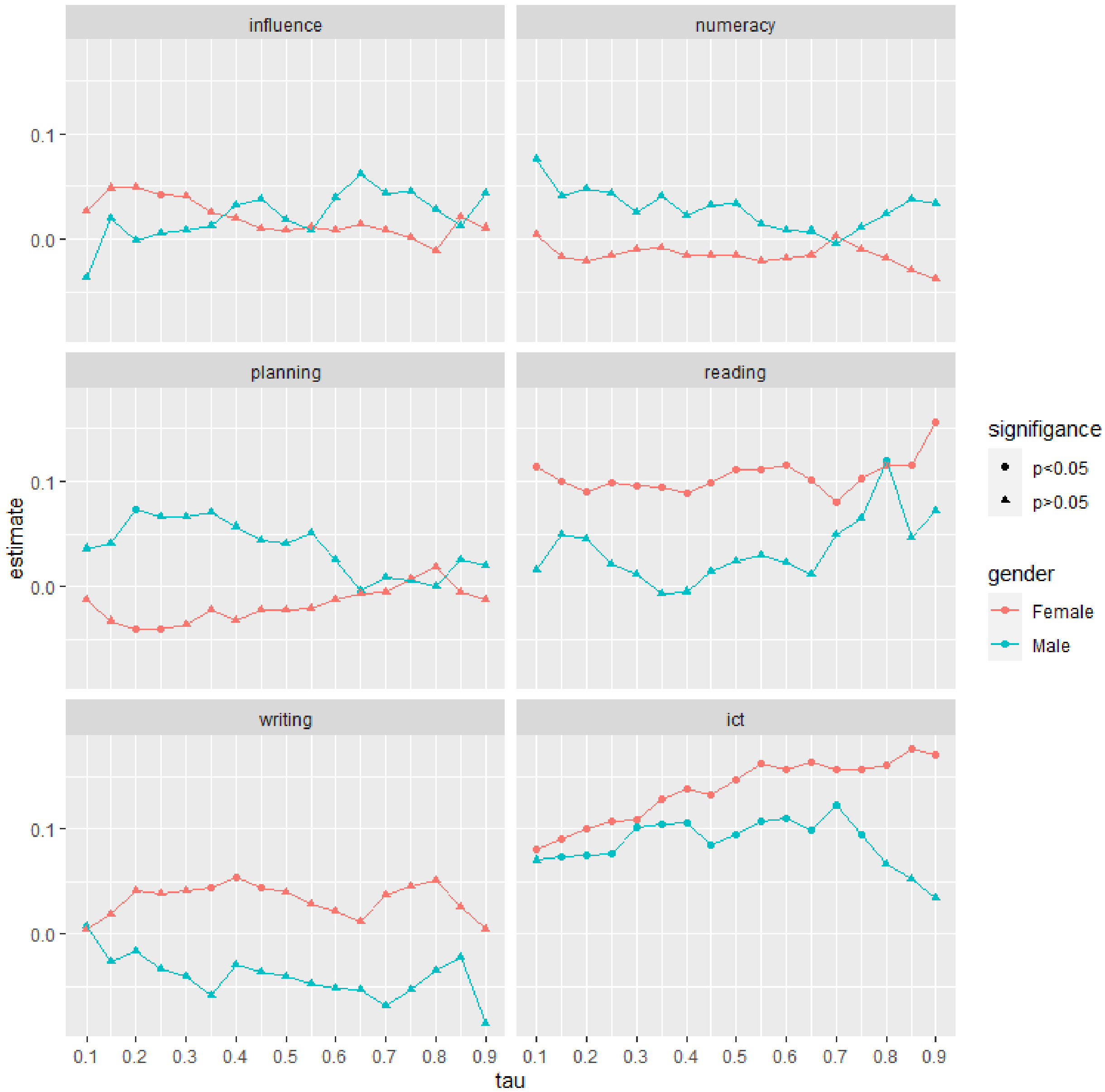


Figure 5. Wage regression coefficients of skill use

Source: author's calculations based on valid observations of PIAAC data for Estonia

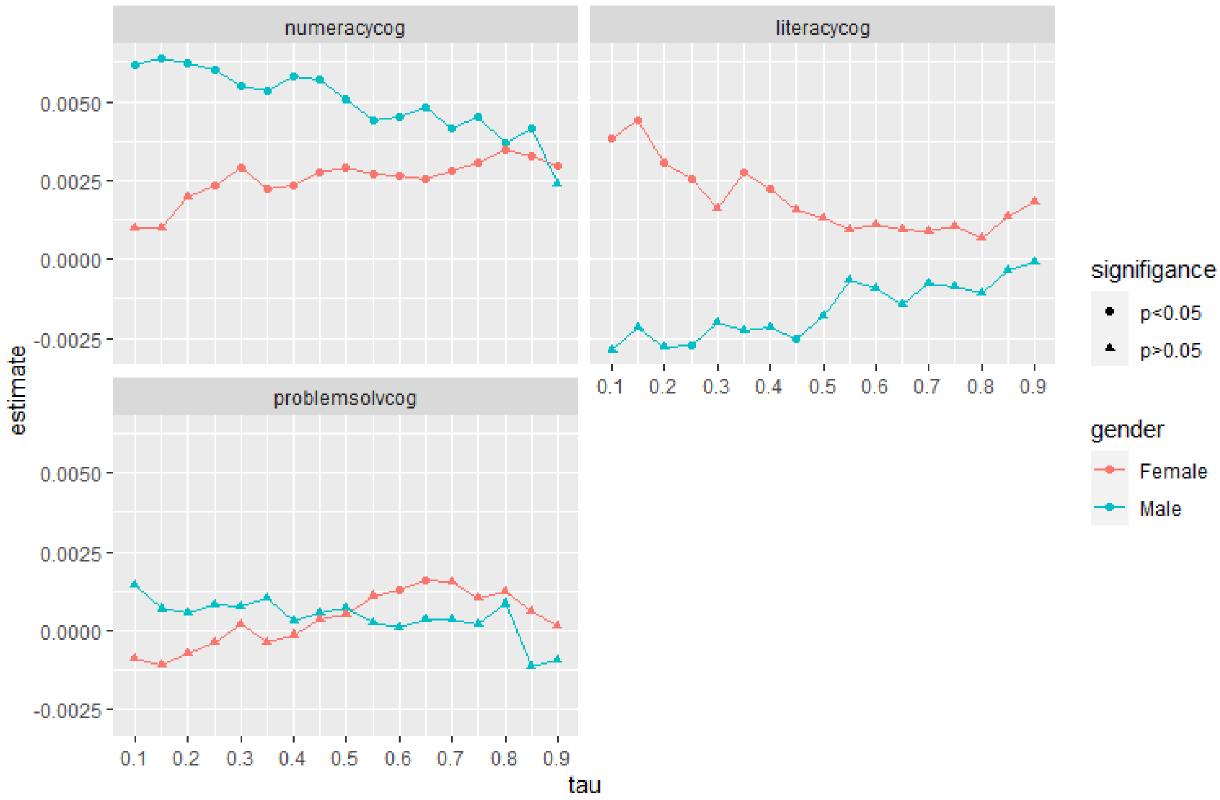


Figure 6. Wage regression coefficients of cognitive skills

Source: author's calculations based on valid observations of PIAAC data for Estonia

Quantile regression results for men (see Appendix 7) show that age is not statistically significant but working years is, in some parts of the income distribution. Only the highest level of education is rewarded with higher earnings at different earning levels. Other education levels predict lower wages at higher earnings. Having children is mostly associated with higher earnings. Working in the private sector is relevant throughout the earnings distribution and affects wages up to 37.2%, in the 90th percentile. Working hours variable predicts lower earnings at all income levels. White collar occupations also predict lower income. The firm size variable is mostly statistically significant in the 25th to 50th percentiles where firm size correlates positively with income. Working hours flexibility is statistically significant in the 25th to 90th percentiles. Higher flexibility is mostly associated with higher earnings.

Quantile regression results for women (see Appendix 8) show that age and age squared are only statistically significant in the 75th percentile. The first predicts higher and the latter lower earnings. Education is mostly relevant only in the upper part of income group where it predicts increasingly higher income with each level of education. Having children is only statistically significant at the 10th percentile where having one or three children associate with higher earnings. Living with a spouse is only statistically significant at the 10th percentile and predicts higher income. Working in a private sector is associated with higher wages, up to 15.8%. Working hours predict lower income. Training on the job predicts up to 17.8% higher earnings. White collar occupations are statistically significant in the 10th to 50th percentile and associate with lower wages. Firm size is statistically significant at all levels of income. Earnings seem to increase with the size of the firm at every income level. Flexibility of working hours is mostly statistically significant at the ends of the income distribution. Higher flexibility is associated with higher earnings, with the exception of the highest level of flexibility at the 10th percentile where it predicts 29% lower income.

2.4.3. Results of Oaxaca – Blinder decompositions

The author conducted two different twofold Oaxaca – Blinder decompositions, with the R package "oaxaca" (Hlavac, 2018), in order to assess the impact of skills to gender wage gap. The reference coefficients come from pooled regressions without the group indicator variable included as a covariate based on (Neumark, 1988). The Decomposition 1 is based on the same variables as OLS Model 1 (see Table 2) and it estimates the wage gap based on human capital variables and other variables that were used in the analysis. The Decomposition 2 is based on OLS Model 4 (see Table 2) and this second decomposition will also include the focus variables: skill use and cognitive skills. Comparing the results of the two decompositions allows to rate the influence of the focus variables on gender wage gap.

The Decompositon 1 divides 2,626 observations in two groups, based on gender (see Table 3). Men earn on average 39.5% more than women. Women log gross hourly earnings are 1.38 and men earn 1.78. The difference of 39.5% in average earnings between the two genders can be divided into two parts: explained part and unexplained part. Explained part of the difference is 20.24% and unexplained part is 19.26% (see Table 4). The Decomposition 1 is able to explain 51.24% of the gender wage gap and leaves 49.76% of it unexplained.

Table 3. Oaxaca – Blinder decomposition results

	Decomposition 1		**Decomposition 2**	
Group	Males	Females	Males	Females
Group size	1,080	1,566	639	1,069
Group result (log hourly gross wage)	1.78	1.38	1.89	1.52
Difference	39.5%		36.94%	

Source: author's calculations based on valid observations of PIAAC data for Estonia

The analysis of Decomposition 2 is based on 1,708 observations. Men on average earn 36.94% more than women. Male log hourly earnings are 1.89 and female earnings 1.52. The explained part of the difference in earnings is 19.58% and the unexplained part is 17.36%. Decomposition 2 is thus able to explain 53% of the difference in earnings and 47% stay unexplained.

Assuming that the unexplained part of the decomposition is of discriminatory nature, and the explained part is not, the coefficients of the unexplained part could show where this discrimination originates from. The unexplained part can be broken into two: unexplained *A* and unexplained *B*. In Decomposition 1, the unexplained part of 19.26% can be broken into two: 11.4% and 7.86%. The unexplained *A* coefficient of 11.4% shows discrimination in favour of men and the unexplained *B* coefficient of 7.86% shows discrimination against women. In Decomposition 2, the discrimination component in favour of men is 10.87% and against women is 6.5%. The discrimination in favour of men and against women dropped with Decomposition 2.

Table 4. Oaxaca – Blinder decomposition coefficients and standard errors

	Decomposition 1	**Decomposition 2**
Coefficient (explained)	20.24%	19.58%
SE (explained)	1.74%	1.98%
Coefficient (unexplained)	19.26%	17.36%
SE (unexplained)	1.46%	1.68%
Coefficient (unexplained A)	11.4%	10.87%
SE (unexplained A)	0.98%	1.23%
Coefficient (unexplained B)	7.86%	6.5%
SE (unexplained B)	0.6%	0.67%

Source: author’s calculations based on valid observations of PIAAC data for Estonia

Out of the cognitive skills and skill use, only cognitive numeracy has a marginal, and statistically significant, effect in the explained part of the gender wage gap (Appendix 9). It seems men, being more endowed in cognitive numeracy skill, are getting higher wages partly due to this skill. None of the skills had a statistically significant effect in the unexplained part of the gender wage gap.

3. CONCLUSION

The aim of the thesis was to investigate the differences in skills among Estonian individuals and between genders to understand how these differences might drive wages and contribute to explaining the gender wage gap. To achieve the aim, answers to the following questions were sought: “how large is the difference in skills between genders?”, “how might different skills predict wage levels?” and “how might skills affect the gender wage gap?”.

To achieve the aim and answer the raised questions, the author provided descriptive statistics of the PIAAC dataset; visualised the distributions of wages, skill use and cognitive skills; conducted OLS regressions, quantile regressions and Oaxaca – Blinder decompositions.

The gender wage gap was evident in the PIAAC dataset. The difference in gross hourly earnings between men and women was 52%. Men earned, on average, 2.5 euros more per hour. The distribution of male wage earners was clearly shifted more to high earnings than the female population. Men worked three hours more per week than women, on average. On average, women tended to have more children than men. There was a larger proportion of women employed skilled and white-collar occupations, more men were employed as blue-collar workers. In terms of work flexibility, the flexibility of working hours, men tended to enjoy more flexibility than women, on average. Women, on the other hand were more highly educated. Both genders were living with a spouse in similar proportion. Larger share of women, than men, were working part-time or were fulfilling domestic tasks. A larger population of women had on-the-job training, than men. The horizontal segregation of men and women across industries was evident. A larger proportion of women was working in the public sector than men.

What comes to skill use and cognitive skills, there were differences in the distributions but the mean nor median values were not far off between the two genders. The statistically significant differences in means, between genders, were in numeracy, planning and ICT skill use. These skills were more frequently used by men than women. The differences in means were up to 0.1 points on the indices with means around two points. These indices are able to show which of the genders uses which skill more frequently, on average, but they are not able to show the difference in the actual skill use frequency. The mean differences between genders in cognitive scores were all statistically significant. Here, men scored better in numeracy and problem-solving and women in literacy. The differences in endowments are in line with the findings of (Halpern & LaMay, 2000) who found men to be better in mathematics and, women better in reading and writing. The score means were around 270 points out of the 500 points maximum. The mean differences were around two to six points between genders.

Reading, influence and ICT were the only statistically significant skill use variables and numeracy was the only statistically significant cognitive skill in the latest OLS model. The skill use variables predicted a premium to wages, up to five percent per one point of skill use. Cognitive numeracy predicted a premium of two percent, per 10 points in score, to earnings. Adding skills as variables improved the model based on the Mincerian equation by raising the adjusted coefficient of determination, from 36% to 40%, and lowering the male dummy coefficient value, from 30% to 25%. Surprisingly, having children was associated with higher pay than not having. This in stark contrast with the findings of (Dolton & Makepeace, 1987; Waldfogel, 1997; Kleven et al., 2019). Only the highest level of education is statistically significant, predicting 18% higher earnings than primary and lower secondary education.

Based on the OLS regression, all the statistically significant skill use and cognitive skill variables predicted higher wages. Men used ICT skills more frequently were more endowed in numeracy than women. Women used influencing and reading skill more frequently, but the differences in skill use were not statistically significant between genders. Based on the above, the author accepts the hypothesis H_1: The skills that predict higher wages are more prominent in men than women.

The quantile regressions results showed that ICT was the only variable of skill use statistically significant throughout the most part of the earnings distribution of both sexes. ICT skill showed increasingly higher effect on wages going up the earnings distribution of women, up to 18%. Surprisingly, men are rewarded less for the ICT skill throughout the earnings distribution. ICT skill predicts up to 12% higher earnings for men. Reading skill was statistically significant throughout the earnings distribution but only for women. It offers a premium of around 10% to 15% throughout the earnings distribution. Out of the cognitive skills, numeracy was the only skill mostly statistically significant throughout the earnings distribution. On the overall, men gain more from this skill than women. On the other hand, men are decreasingly rewarded for the numeracy skill going up the earnings distribution. The differences between genders close at the 80^{th} percentile. The author also conducted quantile regression with together with focus and control variables, based on the same setup as the latest OLS model. Here, also, the results showed that having children, where statistically significant, was associated with a higher pay. Education was more relevant for women throughout the earnings distribution and predicted higher earnings than for men. On-the-job training was mostly only relevant for women. This is different of what Mincer (1962) indicated. He emphasised the signifigance of on-the-job training to men, being even as important as education. Firm size was statistically significant at all income levels for women but only significant at the 25^{th} and 50^{th} percentiles for men. For the most part, the flexibility of working hours was associated with higher earnings. Peciuliarly, flexibility to a very high extent was associated with 29% lower earnings for women in the 10^{th} quantile, which is the opposite of the findings of Shepard III et al.(1996).

The results of the Oaxaca – Blinder decompositions showed that controlling for skill use and cognitive skills helped to increase the explained part of the wage gap. The first decomposition had a gender wage gap of 39.5%, of which the explained component was 51.24%. The second decomposition, which included all focus variables, had a gender wage gap of 36.94%, of which the explained part was 53%. Adding skill use and cognitive skills to the decomposition, increased the explained part of the gender wage gap by 1.76 percentage points. On the other hand, the overall gender wage gap decreased with addition of the focus variables. Also, the mean wages for both men and women increased. This is an indication that the sample changed and shifted more to the higher earning individuals. In the second decomposition, cognitive numeracy was the only focus variable to have a material effect on the explained component of the wage gap. The effect of numeracy was in favour of men which is in line with the fact that men had higher numeracy scores than women. Based on the above, the author accepts the hypothesis H_2: Controlling for skills reduces the unexplained part of the wage gap.

Using PIAAC dataset allows for a better explanation than more general datasets on the labour market. For example, an analysis based on Annual Survey of Hours and Earnings, which is the UK's most comprehensive survey of individual pay, concluded that 36.1% of the wage gap could be explained (Office for National Statistics, 2018). An analysis, by Anspal (2015), based on the Estonian Labour Force Survey data for 2005 to 2014 concluded the explained part of the gender wage gap to be 32%. However, the author calls for caution with generalising the results of this thesis based on the PIAAC dataset. The issue lies with the limited number of observations with what the analyses were done. The number of observations dropped gradually with the increase of variables. Firstly, not all of the participants reported their earnings. Secondly, the participants of the survey could answer “don’t know” or could skip or even refuse to answer questions. These are probably the reasons behind the smaller number of observations on some variables. Thirdly, the samples probably changed in between the analyses along with the drop in observations. Seems that, the author comes to the same conclusion as did Halapuu (2015) cautioning with the interpretation of the findings.

All in all, skill use and cognitive skills seem do have impact on wages. These variables seem do help to explain the gender wage gap by lowering the unexplained part of the gap.

KOKKUVÕTE

OSKUSTE MÕJU SOOLISELE PALGALÕHELE EESTIS

Kevin Orion

Käesoleva lõputöö eesmärk on uurida erinevusi oskustes Eesti indiviidide ning meeste ja naiste vahel. Sellga püütakse aru saada kuidas need erinevused oskustes võivad mõjutada palgataset ja soolist palgalõhet. Eesmärgi saavutamiseks otsitakse lõputöös vastuseid järgnevatele küsimustele: "kui suured on soolised erinevused oskustes?", "kuidas võivad erineva oskused ennustada palgataset?" ja "kuidas võivad oskused mõjutada soolist palgalõhet?".

Lõputöös püstitati järgnevad hüpoteesid:
H_1: Oskused mis ennustavad kõrgemat palgataset on rohkem levinud meeste seas kui naiste seas.
H_2: Oskused aitavad vähendada selgitamata soolise palgalõhe osa.

Töö esimeses peatükis tutvustatakse inimkapitali teooriat, tuuakse välja uuringuid soolise palgalõhe teemal, oskuste ja ülesannete teemal ning soolise palgalõhe teemal Eestis. Teises peatükis tutvustab autor andmestikku, esitleb kirjeldavat statistikat, viib läbi analüüsid ja toob välja analüüside tulemused. Andmestik pärineb OECD läbiviidud rahvusvahelisest täiskasvanute oskuste uuringust. Töös kasutatud meetodid on vähimruutude meetod, kvantiilregressiooni meetod ja Oaxaca – Blinderi dekompositsioon. Kolmandas peatükis autor võtab kokku tulemused ja võrdleb neid teiste töödega.

Andmestikust tuleb välja, et mehed teenivad keskmiselt 2,5 eurot ehk 52% rohkem kui naised. Mehed töötavad keskmistel kolm tundi rohkem kui naised nädalas. Rohkem mehi töötab sinikrae tegevusaladel, rohkem naisi samas valgekrae tegevusaladel. Meestel on keskmiselt suurem vabadus töötundide paindlikkuse kohalt. Suurem proportsioon naisi töötas kas osakoormusega või tegeles koduste ülesannetega, kui mehi. Meeste ja naiste horistontaalne segregeerumine erinevatele tegevusaladele oli ilmne. Suurem vahekord naisi töötas avalikus sektoris kui mehi.

Vähimruutude meetodi põhjal oli oskusteste kasutamisest statistiliselt oluline lugemine, mõjutamine ning info- ja kommunikatsioonioskused (IKT). Kognitiivsetest oskustest oli statistiliselt oluline vaid matemaatiline kirjaoskus. Kvantiilregressiooni tulemused näitasid, et IKT oli ainuke statistiliselt oluline näitaja mõlema soo puhul. Mehed saavad IKT-oskuste kasutamise eest väiksema palgalisa kui seda saavad naised ja seda üle kogu palgajaotuse. Lugemisoskuse kasutamine oli ainult naiste puhul statistiliselt oluline ja ennustas kõrgemat palka üle kogu palgajaotuse. Kognitiivsetest oskustest oli relevantne vaid matemaatiline kirjaoskus. Meeste palgalisa selle oskuse eest on suurem kui naiste puhul, kuid mida kõrgema palgataseme poole liikuda, seda väiksem mõju palgale. Oaxaca – Blinderi dekompositsiooni tulemused näitasid, et oskuste kasutamise ja kognitiivsete oskuste lisamine dekompositsiooni aitas vähendada selgitamata soolist palgalõhet pea kahe protsendipunkti võrra (51,24% pealt 53% peale). Matemaatiline kirjaoskus oli ainuke muutuja oskustest, mis aitas märkimisväärselt kaasa selgitatud palgalõhe komponendile. Matemaatiline kirjaoskus selgitas palgalõhet meeste kasuks.

Autor tahab juhtida tähelepanu antud magistritöö tulemuste üldistamisele. Nimelt analüüsid tehti väheste vaatluste põhjal: 1 708. Selline piirang tuli andmestiku iseärasustest ja kasutades suuremal hulgal muutujaid. Esiteks, kõikidel uuringus osalejatel polnud palka märgitud. Teiseks, osalejatel oli võimalus vastata "ei tea", keelduda vastamast või jätta vastamata küsimustele. Seega osadel muutujatel oli väiksem arv vastuseid. Kolmandaks, valimid arvatavasti muutusid, analüüside vahel, koos vaatluste arvu vähenemisega.

Võttes kõike eelnevat kokku, oskuste kasutamisel ja kognitiivsetel oskustel on mõju soolisele palgalõhele. Need aitavad suurendada selgitatud osa palgalõhest.

LIST OF REFERENCES

Acemoglu, D., & Autor, D. (2010). Skills, Tasks and Technologies: Implications for Employment and Earnings. *NBER Working Papers*.

Aksoy, C. G., Özcan, B., & Philipp, J. (2020). *Robots and the Gender Pay Gap in Europe.* Bonn: IZA – Institute of Labor Economics.

Anspal, S. (2015). *Essays on gender wage inequality.* Tartu: University of Tartu Press.

Anspal, S., & Rõõm, T. (2010). *Gender Pay Gap in Estonia: Empirical Analysis.* Tallinn: Praxis Center for Policy Studies.

Anspal, S., Biin, H., Kallaste, E., Karu, M., & Kraut, L. (2009). *Sooline palgalõhe: Teoreetilise ja empiirilise kirjanduse ülevaade.* Tallinn: Eesti Rakendusuuringute Keskus CENTAR; Poliitikauuringute Keskus PRAXIS.

Autor, D., Levy, F., & Murnane, J. R. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics, 118*(4), 1279-1333.

Bacolod, M. P., & Blum, B. S. (2010). Two Sides of the Same Coin: U.S. "Residual" Inequality and the Gender Gap. *The Journal of Human Resources, 45*(1), 197-242.

Beaudry, P., & Lewis, E. (2012). *Do Male-Female Wage Differentials Reflect Differences in the Return to Skill? Cross-city Evidence from 1980-2000.* Cambridge: National Bureau of Economic Research.

Blackburn, R. M., Brooks, B., & Jarman, J. (2001). The Vertical Dimension of Occupational Segregation. *Work, Employment & Society, 15*(3), 511-538.

Blair, F. (2018). The Trouble With Human Capital Theory. *Real-WorldEconomics Review*(86), 15-32.

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

Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources, Vol. 8, No. 4*, 436-455.

Burton-Jones, A., & Spender, J.-C. (2012). *The Oxford Handbook of Human Capital.* Oxford: Oxford University Press.

Card, D., Cardoso, A. R., & Kline, P. (2015). *Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women.* Cambridge: National Bureau of Economic Research.

Christl, M., & Köppl–Turyna, M. (2020). Gender wage gap and the role of skills and tasks: evidence from the Austrian PIAAC data set. *Applied Economics, Vol. 52, Issue 2*, 113-134.

Deming, D. J. (2017). *The Growing Importance of Social Skills in the Labor Market.* Cambridge: National Bureau of Economic Research.

Deming, D., & Kahn, L. B. (2018). Skill Requirements across Firms and Labor Markets: Evidence from JobPostings for Professionals. *Journal of Labor Economics, 36*(1), 337-369.

Dias, M. C., Joyce, R., & Parodi, F. (2018). *The gender pay gap in the UK: children and experience in work.* Institute for Fiscal Studies.

Dolton, P. J., & Makepeace, G. H. (1987). Marital Status, Child Rearing and Earnings Differentials in the Graduate Labour Market. *The Economic Journal, 97,* 897-922.

Dolton, P., & Makepeace, G. H. (1985). The statistical measurement of discrimination. *Economics Letters*, *18*(4), 391-395.

Goldin, C., & Katz, L. W. (2016). A Most Egalitarian Profession: Pharmacy and the Evolution of a Family-Friendly Occupation. *Journal of Labor Economics, 34*(3).

Halapuu, V. (2015). *Infotöötlusoskuste roll soolise ja keelelise palgalõhe selgitamisel Eestis: PIAAC uuringu temaatiline aruanne nr 4.* Tartu: Haridus- ja Teadusministeerium.

Halapuu, V., & Valk, A. (2013). *Täiskasvanute oskused Eestis ja maailmas: PIAAC uuringu esmased tulemused.* Tartu: Haridus- ja Teadusministeerium.

Halpern, D. F., & LaMay, M. L. (2000). The Smarter Sex: A Critical Review of Sex Differences in Intelligence. *Educational Psychology Review, 12*(2), 229-246.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review 73*, 103-130.

Heckman, J. J., & Kautz, T. D. (2012). *Hard Evidence on Soft Skills.* Cambridge: National Bureau of Economic Research.

Hlavac, M. (2018). *oaxaca: Blinder-Oaxaca Decomposition in R.* Retrieved from https://CRAN.R-project.org/package=oaxaca

Jarman, J., Blackburn, R. M., & Racko, G. (2012). The Dimensions of Occupational Gender Segregation in Industrial Countries. *Sociology, 46*(6), 1003-1019.

Kleven, H., Landais, C., & Søgaard, J. E. (2019). Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, *11*(4), 181-209.

Marshall, A. (1920). *Principles of Economics*, *Eighth edition.* London: Palgrave Macmillan.

Masso, J., Meriküll, J., & Vahter, P. (2020). *The role of firms in the gender wage gap.* Tartu: The University of Tartu FEBA.

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy, 66*(4), 281-302.

Mincer, J. (1962). On-the-Job Training: Costs, Returns, and Some Implications. *Journal of Political Economy*, *70*(5), 50-79.

Mincer, J. (1974). *Schooling*, *Experience and Earnings.* New York: Columbia University Press.

Mincer, J., & Polachek, S. (1974). Family Investments in Human Capital: Earnings of Women. *Journal of Political Economy, 82*(2), 76-108.

Neumark, D. (1988). Employers' Discriminatory Behavior and the Estimation of Wage Discrimination. *The Journal of Human Resources, 23*(3), 279-295.

Nobel Media AB 2021. (2021, March 15). *The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 1979.* Retrieved from NobelPrize.org: https://www.nobelprize.org/prizes/economic-sciences/1979/summary/

Nobel Media AB 2021. (2021, May 2). *The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 1992*. Retrieved from NobelPrize.org: https://www.nobelprize.org/prizes/economic-sciences/1992/summary/

Ñopo, H. (2008). Matching as a Tool to Decompose Wage Gaps. *The Review for Economics and Statistics, 90*(2), 290-299.

Oaxaca, R. L. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review Vol. 14, No.*, 693-709.

OECD. (2007). *Human Capital: How what you know shapes your life.*

OECD. (2016). How skills are used in the workplace. *Skills Matter: Further Results from the Survey of Adult Skills*, 95-119.

OECD. (2019). *The Survey of Adult Skills : Reader's Companion, Third Edition, OECD Skills Studies.* Paris: OECD Publishing.

OECD. (2020, April 17). *Public data and analysis.* Retrieved from OECD Web site: http://www.oecd.org/skills/piaac/publicdataandanalysis/

OECD. (n.d.). *About PIAAC*. Retrieved from OECD Skills Surveys: https://www.oecd.org/skills/piaac/about/

Office for National Statistics. (2018). *Understanding the gender pay gap in the UK.*

Paull, G. (2008). Children and Women's Hours of Work. *The Economic Journal, 118,* 8-27.

Rõõm, T., & Kallaste, E. (2004). *Men and Women in the Estonian Labour Market: An Assessment of the Gender Wage Gap.* Tallinn: Praxis Center for Policy Studies.

Schultz, T. W. (1960). Capital Formation by Education. *Journal of Political Economy*, 571-583.

Shepard III, E. M., Clifton, T. J., & Kruse, D. (1996). Flexible Work Hours and Productivity: Some Evidence from the Pharmaceutical Industry. *Industrial Relations, 35*(1), 123-139.

Tverdostup, M., & Paas, T. (2016). Assessment of Labour Market Returns in the Case of Gender Unique Human Capital. *ERSA conference papers.* European Regional Science Association.

Tverdostup, M., & Paas, T. (2016). The Gender Wage Gap in the Human Capital Framework: A Cross-Nordic Assessment Based on PIAAC. *Estonian Discussions on Economic Policy, Vol. 24, No. 2*, 137-160.

UNESCO. (2012). *International Standard Classification of Education ISCED 2011*. Montreal: UNESCO Institute for Statistics.

Vahter, P., & Masso, J. (2018). *The Contribution of Multinationals to Wage Inequality: Foreign Ownership and the Gender Pay Gap.* Tartu: The University of Tartu FEBA.

Waldfogel, J. (1997). The Effect of Children on Women's Wages. *American Sociological Review, 62*(2), 209-217.

Wheatley, D. (2017). Employee satisfaction and use of flexible working arrangements. *Work, Employment & Society, 31*(4), 567-585.

Yamaguchi, S. (2016). Changes in Returns to Task-Specific Skills and Gender Wage Gap. *The Journal of Human Resources, 53*(1), 32-70.

APPENDICES

Appendix 1. Skill use indicators

Indicator	Group of tasks
Information Processing skills	
Reading	Reading documents (directions, instructions, letters, memos, e-mails, articles, books, manuals, bills, invoices, diagrams, maps).
Writing	Writing documents (letters, memos, e-mails, articles, reports, forms).
Numeracy	Calculating prices, costs or budgets; use of fractions, decimals or percentages; use of calculators; preparing graphs or tables; algebra or formulas; use of advanced maths or statistics (calculus, trigonometry, regressions).
ICT skills	Using e-mail, Internet, spreadsheets, word processors, programming languages; conducting transactions online; participating in online discussions (conferences, chats).
Problem solving	Facing hard problems (at least 30 minutes of thinking to find a solution).
Other generic skills	
Task discretion	Choosing or changing sequence of job tasks, the speed of work, working hours; choosing how to do the job.
Learning at work	Learning new things from supervisors or co-workers; learning-by-doing; keeping up to date with new products or services.
Influencing skills	Instructing, teaching or training people; making speeches or presentations; selling products or services; advising people; planning others' activities; persuading or influencing others; negotiating.
Co-operative skills	Co-operating or collaborating with co-workers.
Self-organising skills	Organising time.
Physical skills (gross)	Working physically for a long period.
Dexterity	Using skill or accuracy with hands or fingers.

Source: (OECD, 2019)

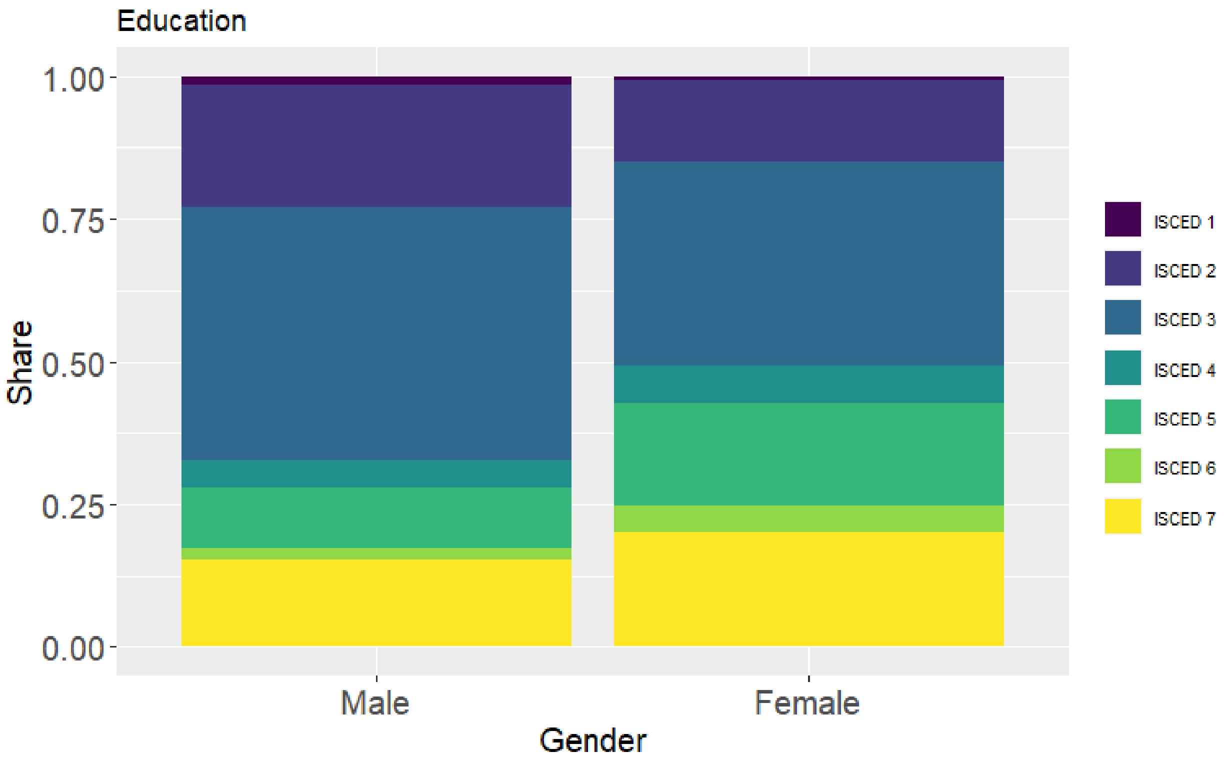
Appendix 2. Used variables

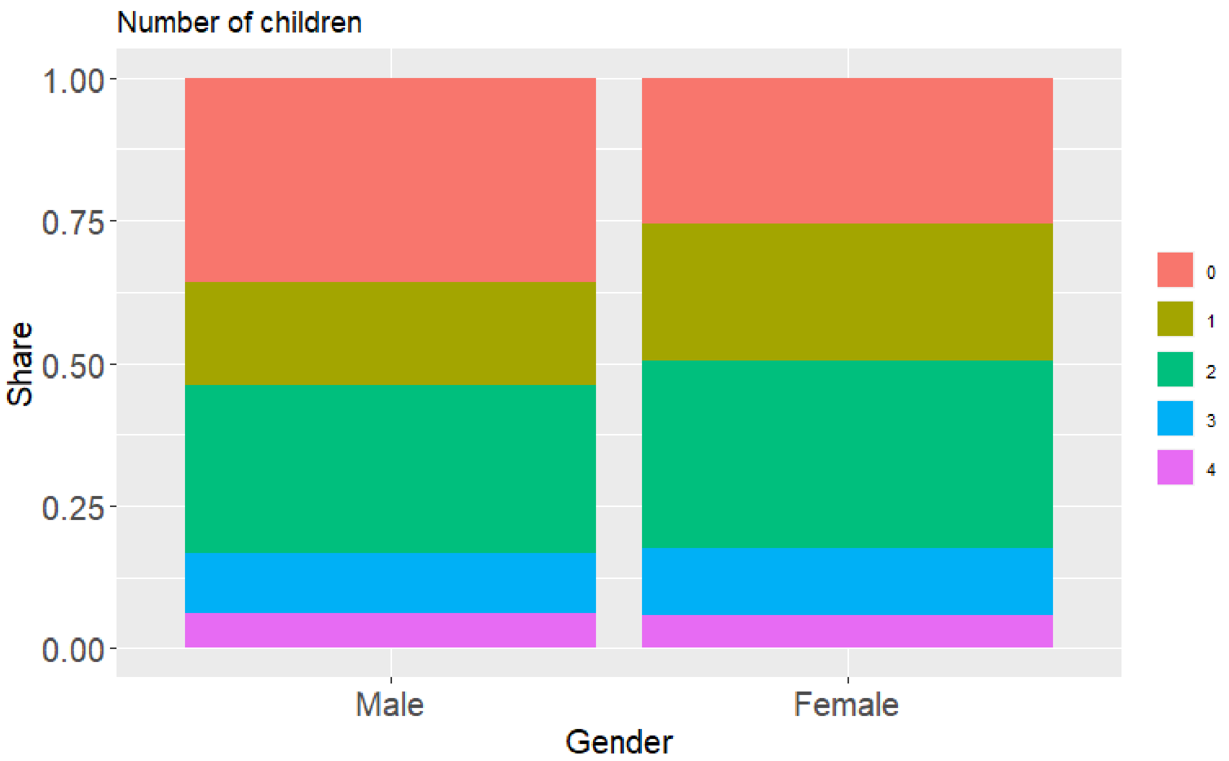
Variable	Variable name in PIAAC	Label
Dependent variable		
Gross hourly earnings	EARNHR	Hourly earnings excluding bonuses for wage and salary earners (derived)
Independent variables		
Focus variables		
Cognitive skills		
Numeracy score	num1, num2, num3, num4, num5, num6, num7, num8, num9, num10	Numeracy scale score - Plausible values 1 to 10
Literacy score	lit1, lit2, lit3, lit4, lit5, lit6, lit7, lit8, lit9, lit10	Literacy scale score - Plausible values 1 to 10
Problem-solving score	psl1, psl2, psl3, psl4, psl5, psl6, psl7, psl8, psl9, psl10	Problem-solving scale score - Plausible values 1 to 10
Skill use		
Influencing skills	INFLUENCE	Index of use of influencing skills at work (derived)
Numeracy skills	NUMWORK	Index of use of numeracy skills at work (basic and advanced - derived)
Planning skills	PLANNING	Index of use of planning skills at work (derived)
Reading skills	READWORK	Index of use of reading skills at work (prose and document texts - derived)
Writing skills	WRITWORK	Index of use of writing skills at work (derived)
ICT skills	ICTWORK	Index of use of ICT skills at work (derived)
Control variables		
Gender		
Gender	GENDER_R	Person resolved gender from BQ and QC check (derived)
Age		
Age	AGE_R	Person resolved age from BQ and QC check (derived)
Work experience		
Years of work experience	C_Q09	Current status/work history - Years of paid work during lifetime
Hours worked per week		
Hours worked per week	D_Q10	Current work - Hours/week
Children		
Children	J_Q03a	Background - Children
Number of children	J_Q03b_C	Background - Number of children (top-coded at 4)
Number of children (0) – used as the basis	Derived by the author based on Children and Number of children	
Number of children (1)		
Number of children (2)		
Number of children (3)		
Number of children (4)		
Occupational classification		
Skilled occupations – used as the basis	ISCOSKIL4	Occupational classification of respondent's job (4 skill based categories), last or current (derived)
Semi-skilled white-collar occupations		
Semi-skilled blue-collar occupations		
Elementary occupations – used as the basis		

Appendix 2 continued.

Firm size		
1 to 10 people – used as the basis	D_Q06a	Current work - Amount of people working for employer
11 to 50 people		
51 to 250 people		
251 to 1,000 people		
More than 1,000 people		
Work flexibility (working hours)		
Not at all – used as the basis	D_Q11d	Current work - Work flexibility - Working hours
Very little		
To some extent		
To a high extent		
To a very high extent		
Education		
ISCED 1 – used as the basis	EDCAT7	Highest level of formal education obtained (7 categories - derived)
ISCED 2 – used as the basis		
ISCED 3		
ISCED 4		
ISCED 5		
ISCED 6		
ISCED 7		
Living with a spouse/partner		
Spouse	J_Q02a	Background - Living with spouse or partner
On-the-job training		
Training	B_Q12c	Activities - Last year – On-the-job training
Economic sector		
The private sector	D_Q03	Current work - Economic sector
The public sector – used as the basis		
A non-profit organisation – used as the basis		
Industry		
Industry (please see the full list in Appendix 6)	ISIC1C	Industry classification of respondent's job at 1-digit level (ISIC rev 4), current job (derived)
Employment status		
Full-time employed	C_Q07	Current status/work history - Subjective status
Part-time employed		
Unemployed		
Pupil, student		
Apprentice, internship		
In retirement or early retirement		
Permanently disabled		
In compulsory military or community service		
Fulfilling domestic tasks or looking after children/family		
Other		

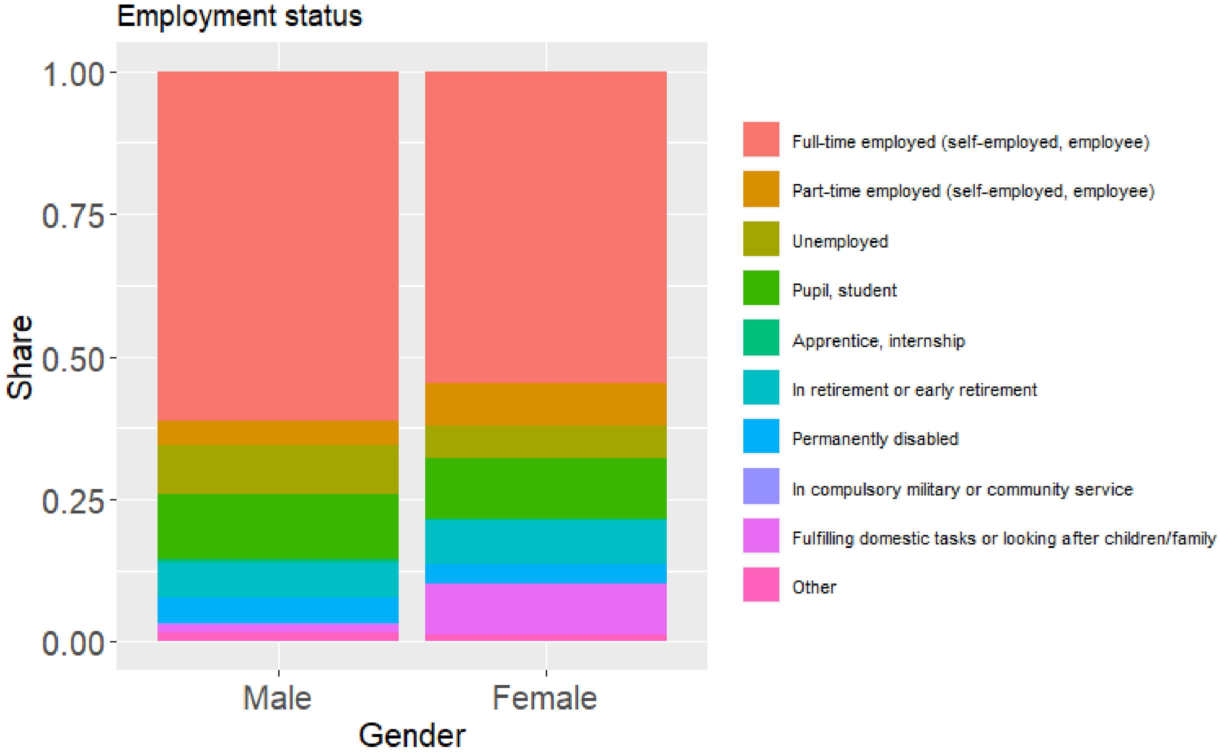
Appendix 3. Education & Number of children: gender comparison

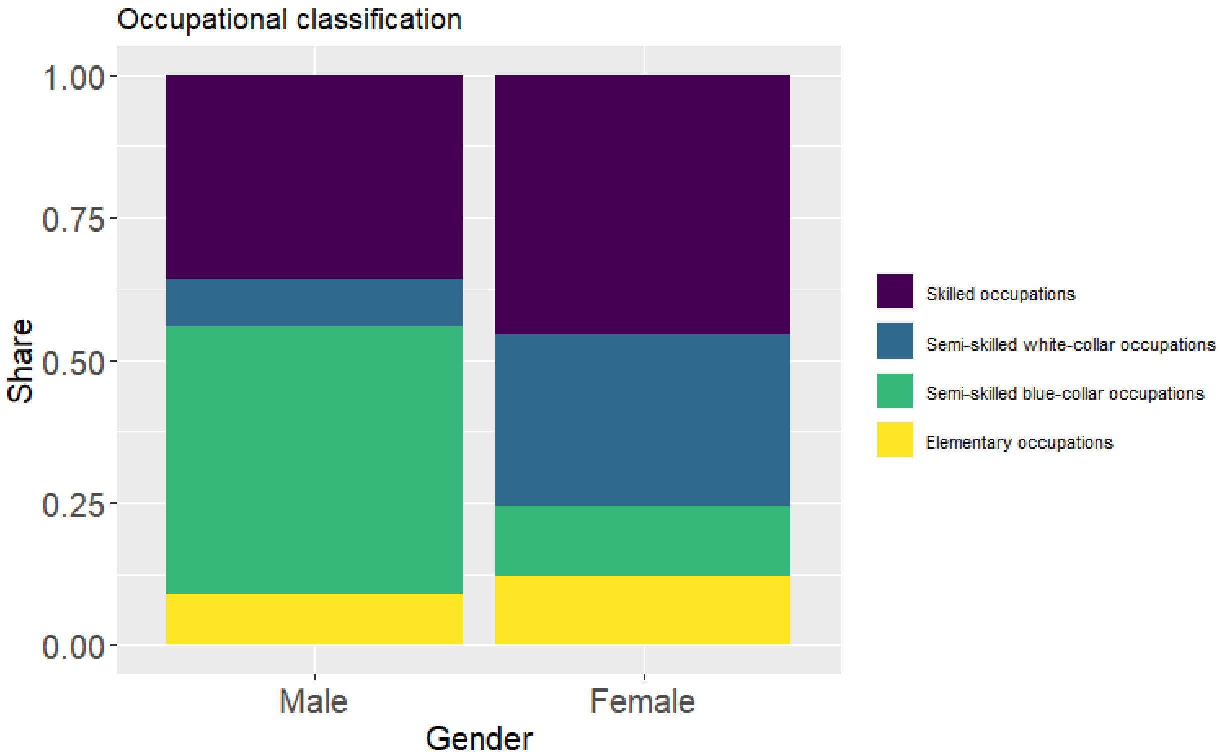




Source: author's calculations based on valid observations of PIAAC data for Estonia

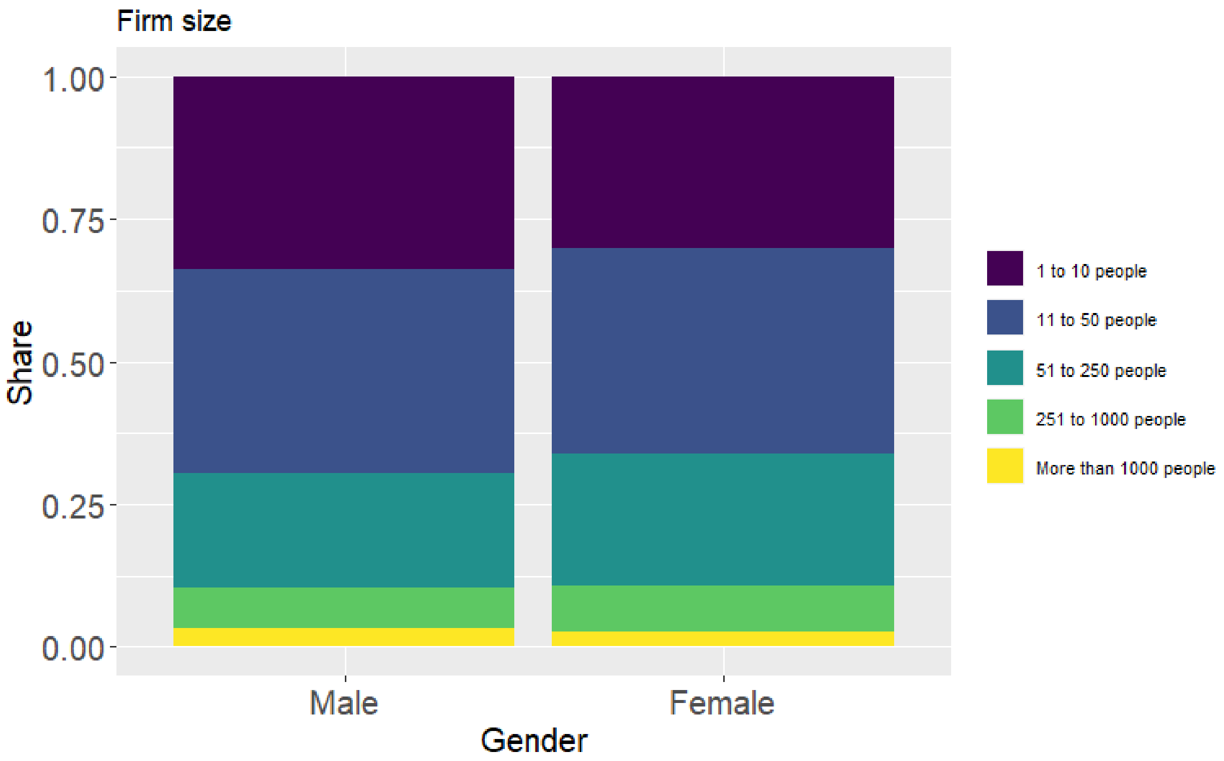
Appendix 4. Employment status & Occupational classification: gender comparison

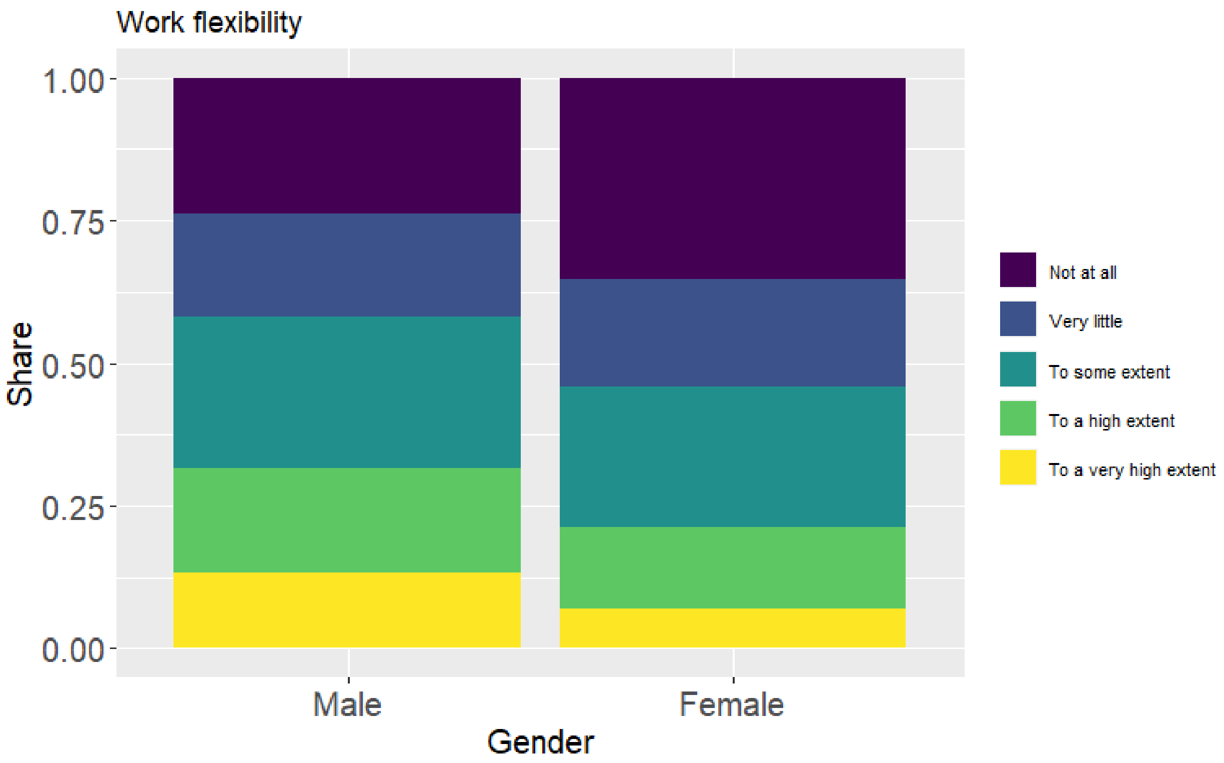




Source: author's calculations based on valid observations of PIAAC data for Estonia

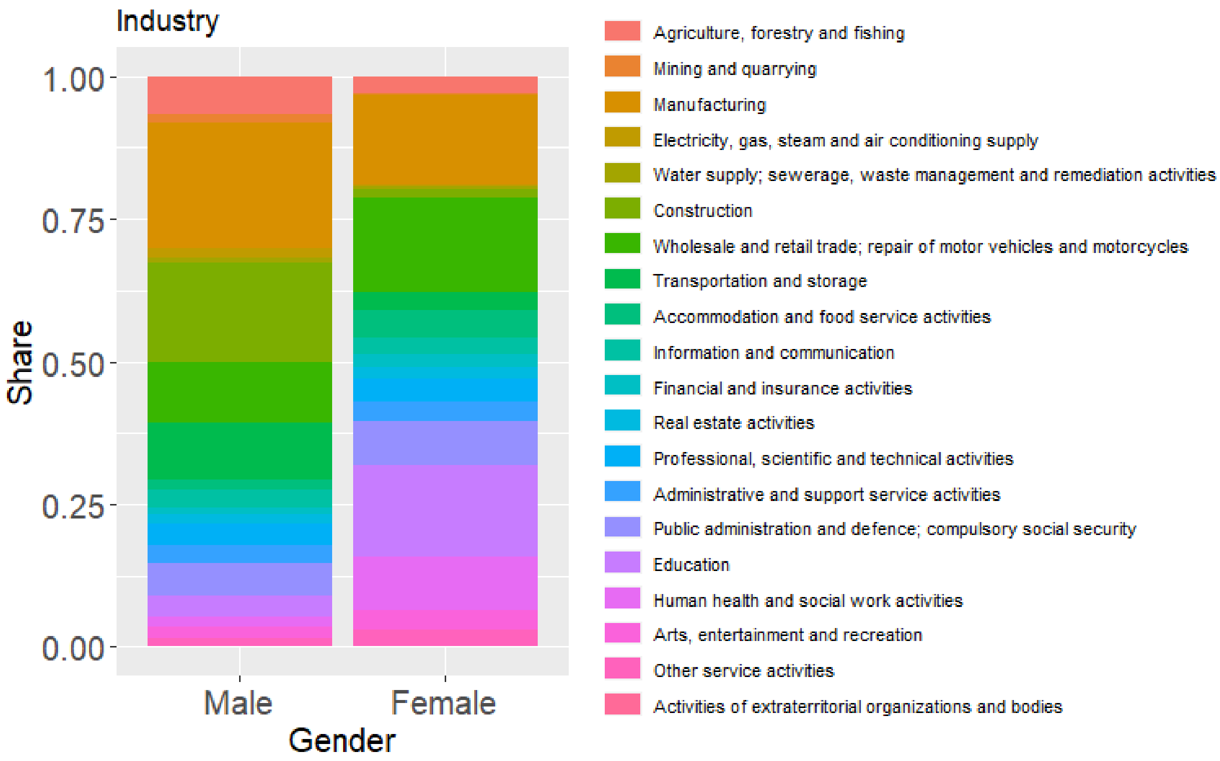
Appendix 5. Firm size & Work flexibility: gender comparison

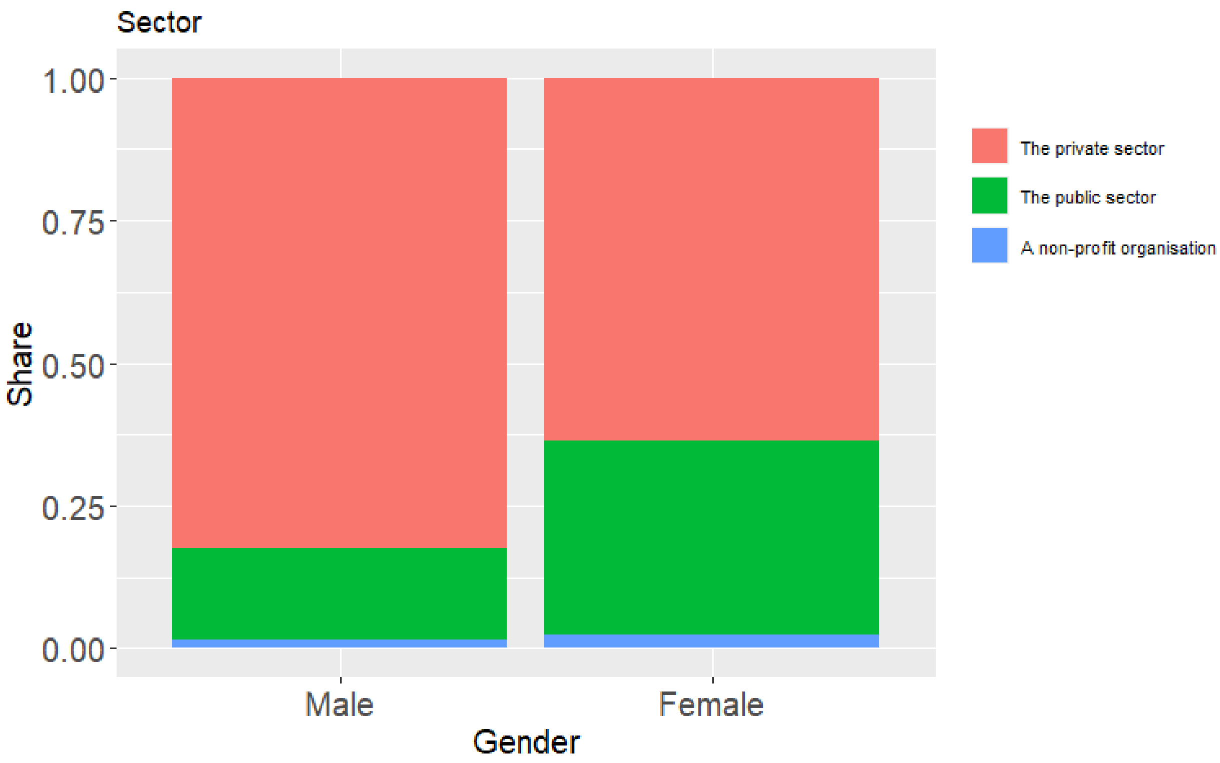




Source: author's calculations based on valid observations of PIAAC data for Estonia

Appendix 6. Industry & Sector: gender comparison





Source: author's calculations based on valid observations of PIAAC data for Estonia

Appendix 7. Quantile regression results for males

	10%	25%	50%	75%	90%
Age	0.032	-0.016	0.001	-0.004	0.031
	(0.046)	(0.035)	(0.028)	(0.041)	(0.027)
Age^2	-0.001	0.00002	-0.0001	0.00001	-0.0005
	(0.001)	(0.0004)	(0.0003)	(0.0005)	(0.0003)
Working years	0.032	0.049***	0.022	0.025	0.024*
	(0.024)	(0.018)	(0.015)	(0.021)	(0.014)
Working years^2	-0.0001	-0.001**	-0.0004	-0.001	-0.0003
	(0.001)	(0.0004)	(0.0003)	(0.0005)	(0.0003)
ISCED 3	0.232	0.146	0.068	-0.039	-0.187**
	(0.163)	(0.124)	(0.101)	(0.145)	(0.095)
ISCED 4	0.057	0.129	0.156	-0.084	-0.323***
	(0.199)	(0.151)	(0.123)	(0.177)	(0.116)
ISCED 5	0.268	0.144	0.072	-0.012	-0.102
	(0.176)	(0.134)	(0.108)	(0.156)	(0.102)
ISCED 6	0.215	0.290*	0.073	0.180	-0.186
	(0.211)	(0.160)	(0.130)	(0.187)	(0.123)
ISCED 7	0.343*	0.275**	0.253**	0.087	0.013
	(0.183)	(0.139)	(0.113)	(0.163)	(0.107)
Children (1)	-0.017	0.067	0.095	0.152	-0.018
	(0.106)	(0.081)	(0.065)	(0.094)	(0.062)
Children (2)	-0.021	0.152*	0.132**	0.206**	-0.001
	(0.106)	(0.080)	(0.065)	(0.094)	(0.062)
Children (3)	0.092	0.161	0.150*	0.181	0.027
	(0.136)	(0.103)	(0.084)	(0.121)	(0.079)
Children (4)	-0.055	0.153	0.292***	0.347**	0.157
	(0.168)	(0.128)	(0.104)	(0.149)	(0.098)
Living with a spouse	0.133	0.013	-0.018	-0.002	0.020
	(0.105)	(0.080)	(0.065)	(0.093)	(0.061)
Private sector	0.231***	0.281***	0.322***	0.321***	0.372***
	(0.080)	(0.061)	(0.049)	(0.071)	(0.046)
Working hours	-0.007**	-0.010***	-0.015***	-0.013***	-0.004**
	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
Training at work	0.118*	0.083	0.067	-0.020	-0.008
	(0.068)	(0.052)	(0.042)	(0.060)	(0.039)
Blue collar	0.064	0.065	0.060	0.052	0.033
	(0.105)	(0.080)	(0.064)	(0.093)	(0.061)
White collar	-0.070	-0.213**	-0.190***	-0.186*	-0.123*
	(0.111)	(0.084)	(0.068)	(0.098)	(0.064)
Small firm (11-50)	0.101	0.189***	0.190***	0.022	-0.006
	(0.081)	(0.062)	(0.050)	(0.072)	(0.047)

Appendix 7 continued.

Medium firm (51-250)	0.181**	0.179***	0.152***	0.024	-0.008
	(0.090)	(0.068)	(0.055)	(0.080)	(0.052)
Large firm (251-1,000)	0.091	0.236**	0.354***	0.115	-0.006
	(0.130)	(0.099)	(0.080)	(0.115)	(0.075)
Largest firm (1,000-…)	0.036	0.241*	0.434***	0.239	0.0001
	(0.166)	(0.126)	(0.102)	(0.147)	(0.096)
Little flexibility	0.091	0.169**	0.186***	0.106	0.120*
	(0.110)	(0.084)	(0.068)	(0.098)	(0.064)
Some flexibility	0.040	0.152*	0.211***	0.208**	0.147**
	(0.103)	(0.078)	(0.063)	(0.091)	(0.060)
High flexibility	0.091	0.100	0.278***	0.341***	0.405***
	(0.114)	(0.086)	(0.070)	(0.101)	(0.066)
Very high flexibility	-0.003	0.185*	0.269***	0.319***	0.399***
	(0.131)	(0.100)	(0.081)	(0.116)	(0.076)
Cognitive literacy	-0.0001	0.001	-0.001	-0.0002	0.002
	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)
Cognitive numeracy	0.004*	0.001	0.003**	0.003	0.004***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)
Cog. problem-solving	0.001	0.001	0.0002	-0.001	-0.004***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
Influence	-0.056	-0.064*	0.012	0.004	0.032
	(0.046)	(0.035)	(0.028)	(0.041)	(0.027)
Numeracy	-0.007	-0.005	-0.011	-0.056	-0.064***
	(0.040)	(0.031)	(0.025)	(0.036)	(0.023)
Planning	0.061	0.072**	0.026	0.065*	0.029
	(0.043)	(0.032)	(0.026)	(0.038)	(0.025)
Reading	0.057	0.057	0.013	0.054	0.112***
	(0.053)	(0.040)	(0.033)	(0.047)	(0.031)
Writing	0.059	0.005	0.007	0.019	-0.044
	(0.051)	(0.039)	(0.031)	(0.045)	(0.030)
ICT	0.021	0.039	0.026	0.070*	0.033
	(0.041)	(0.031)	(0.025)	(0.036)	(0.024)
Constant	-1.280	0.155	0.754	1.077	0.797*
	(0.826)	(0.628)	(0.508)	(0.733)	(0.480)
Controlled for industry	Yes	Yes	Yes	Yes	Yes
Observations	639	639	639	639	639
Note:	*p**p***p<0.01				

Source: author’s calculations based on PIAAC data, dependent variable is natural logarithm of hourly gross wage

Appendix 8. Quantile regressions results for females

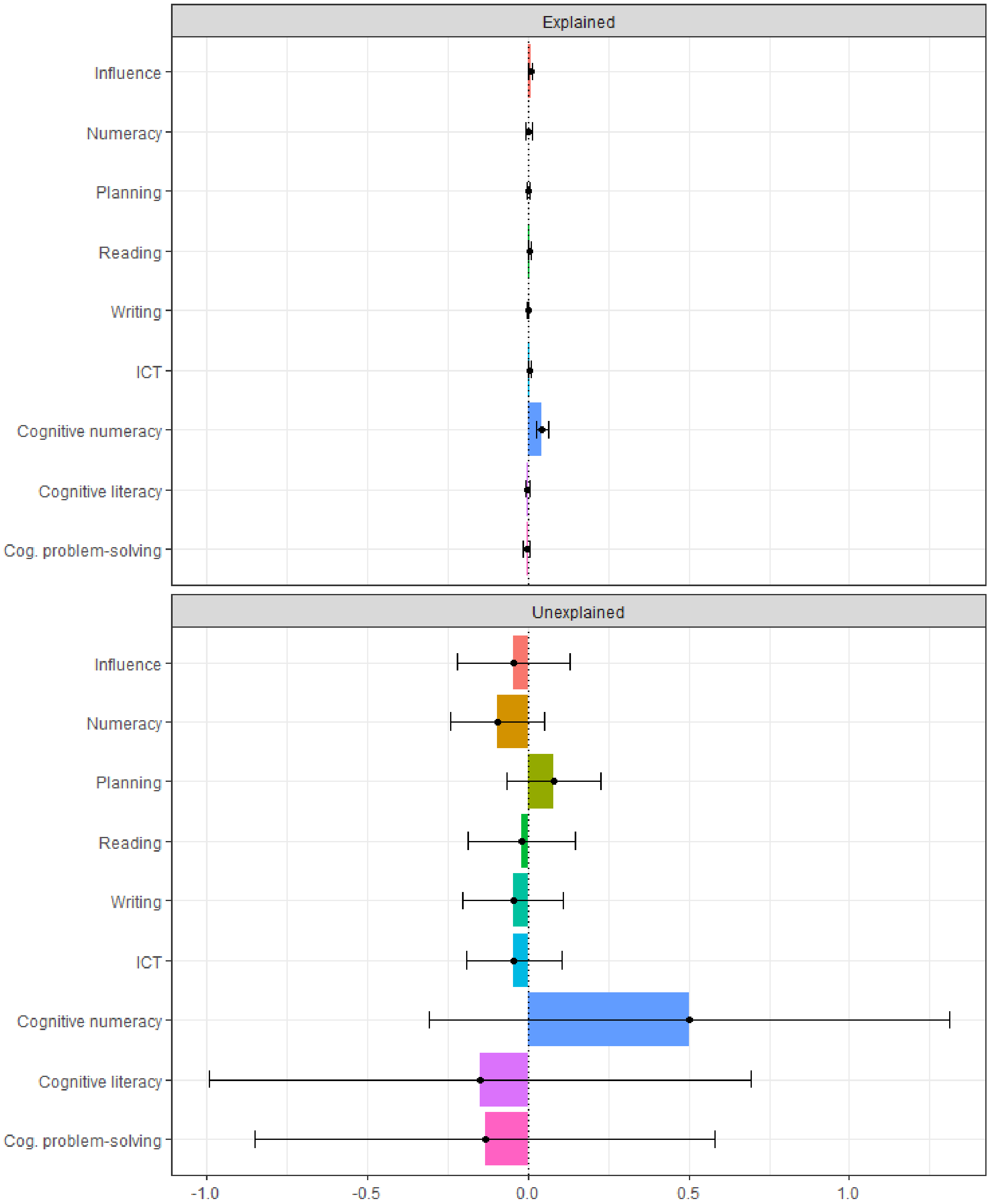
	10%	25%	50%	75%	90%
Age	-0.007	-0.005	0.026	0.040**	0.027
	(0.024)	(0.020)	(0.019)	(0.020)	(0.022)
Age^2	0.0001	0.0001	-0.0003	-0.0005*	-0.0003
	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0003)
Working years	0.005	0.011	0.003	-0.006	0.006
	(0.011)	(0.010)	(0.009)	(0.009)	(0.011)
Working years^2	-0.0001	-0.0003	-0.00001	0.0002	-0.0001
	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
ISCED 3	-0.050	0.007	0.015	0.212*	0.256*
	(0.142)	(0.119)	(0.112)	(0.117)	(0.132)
ISCED 4	0.013	0.003	-0.056	0.145	0.241*
	(0.156)	(0.131)	(0.124)	(0.128)	(0.146)
ISCED 5	-0.029	0.023	0.079	0.295**	0.321**
	(0.143)	(0.120)	(0.114)	(0.118)	(0.134)
ISCED 6	0.094	0.073	0.152	0.320**	0.374**
	(0.156)	(0.131)	(0.124)	(0.128)	(0.146)
ISCED 7	0.126	0.166	0.246**	0.548***	0.574***
	(0.146)	(0.123)	(0.116)	(0.120)	(0.137)
Children (1)	0.160**	0.073	-0.017	-0.048	-0.036
	(0.065)	(0.055)	(0.051)	(0.053)	(0.061)
Children (2)	0.064	0.022	-0.015	-0.034	-0.018
	(0.071)	(0.060)	(0.056)	(0.058)	(0.066)
Children (3)	0.152*	0.055	0.063	-0.016	-0.113
	(0.086)	(0.072)	(0.068)	(0.071)	(0.080)
Children (4)	0.166	-0.016	-0.076	-0.087	0.011
	(0.124)	(0.104)	(0.099)	(0.102)	(0.116)
Living with a spouse	0.087*	0.058	0.021	0.034	-0.002
	(0.047)	(0.040)	(0.037)	(0.039)	(0.044)
Private sector	0.103**	0.089**	0.158***	0.154***	0.124***
	(0.045)	(0.038)	(0.036)	(0.037)	(0.042)
Working hours	-0.013***	-0.010***	-0.013***	-0.017***	-0.021***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
On-the-job training	0.178***	0.106***	0.062*	0.062*	0.102***
	(0.042)	(0.035)	(0.033)	(0.034)	(0.039)
Blue collar	-0.048	-0.025	-0.126	-0.028	0.204
	(0.147)	(0.124)	(0.117)	(0.121)	(0.138)
White collar	-0.234***	-0.195***	-0.128***	-0.045	-0.050
	(0.053)	(0.045)	(0.042)	(0.044)	(0.050)
Small firm (11-50)	0.184***	0.188***	0.183***	0.130***	0.154***
	(0.049)	(0.042)	(0.039)	(0.041)	(0.046)

Appendix 8 continued.

Medium firm (51-250)	0.223***	0.218***	0.233***	0.228***	0.233***
	(0.056)	(0.047)	(0.044)	(0.046)	(0.052)
Large firm (251-1,000)	0.269***	0.338***	0.346***	0.281***	0.333***
	(0.076)	(0.064)	(0.060)	(0.062)	(0.071)
Largest firm (1,000-...)	0.548***	0.376***	0.376***	0.320***	0.488***
	(0.131)	(0.110)	(0.104)	(0.108)	(0.122)
Little flexibility	0.109**	0.044	0.022	0.066	0.026
	(0.055)	(0.046)	(0.043)	(0.045)	(0.051)
Some flexibility	0.114**	0.088**	0.064	0.082*	0.166***
	(0.051)	(0.043)	(0.041)	(0.042)	(0.048)
High flexibility	0.145**	0.119**	0.168***	0.213***	0.288***
	(0.062)	(0.052)	(0.050)	(0.051)	(0.058)
Very high flexibility	-0.290***	-0.093	-0.064	0.157**	0.397***
	(0.089)	(0.075)	(0.071)	(0.073)	(0.083)
Cognitive literacy	0.001	0.0004	-0.0002	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Cognitive numeracy	0.001	0.001	0.002*	0.001	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Cog. problem-solving	0.001	0.0005	0.001	0.0001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Influence	0.038	0.033	0.014	0.012	-0.007
	(0.028)	(0.023)	(0.022)	(0.023)	(0.026)
Numeracy	0.006	-0.002	-0.007	-0.007	0.009
	(0.023)	(0.020)	(0.019)	(0.019)	(0.022)
Planning	-0.028	-0.042*	-0.021	0.014	-0.017
	(0.029)	(0.024)	(0.023)	(0.024)	(0.027)
Reading	0.018	0.074**	0.041	0.011	0.043
	(0.036)	(0.030)	(0.029)	(0.030)	(0.034)
Writing	0.005	0.032	0.032	0.049*	0.040
	(0.031)	(0.026)	(0.025)	(0.026)	(0.029)
ICT	0.071***	0.077***	0.092***	0.081***	0.041*
	(0.023)	(0.020)	(0.019)	(0.019)	(0.022)
Constant	0.101	0.242	0.043	0.099	0.705
	(0.489)	(0.412)	(0.388)	(0.403)	(0.457)
Controlled for industry	Yes	Yes	Yes	Yes	Yes
Observations	1,069	1,069	1,069	1,069	1,069
Note:	*p**p***p<0.01				

Source: author’s calculations based on PIAAC data, dependent variable is natural logarithm of hourly gross wage

Appendix 9. Oaxaca-Blinder Decomposition 2



Source: author's calculations based on PIAAC data, dependent variable is natural logarithm of hourly gross wage

Appendix 10. Non-exclusive licence

A non-exclusive licence for reproduction and for granting public access to the graduation thesis[1]

I, Kevin Orion,

1. Give Tallinn University of Technology a permission (non-exclusive licence) to use free of charge my creation

“Impact of Skills to Gender Wage Gap in Estonia”,

supervised by Simona Ferraro,

1.1. to reproduce with the purpose of keeping and publishing electronically, including for the purpose of supplementing the digital collection of TalTech library until the copyright expires;

1.2. to make available to the public through the web environment of Tallinn University of Technology, including through the digital collection of TalTech library until the copyright expires.

2. I am aware that the author will also retain the rights provided in Section 1.

3. I confirm that by granting the non-exclusive licence no infringement is committed to the third persons’ intellectual property rights or to the rights arising from the personal data protection act and other legislation.

---

[1] The non-exclusive licence is not valid during the access restriction period with the exception of the right of the university to reproduce the graduation thesis only for the purposes of preservation.