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

Miina Hõbenael

RETURNS TO DIGITAL SKILLS ON WAGES: A MICRO-LEVEL ANALYSIS ACROSS EUROPEAN COUNTRIES USING PIAAC DATA

Bachelor's thesis

Programme Applied Economics, specialisation Economic Analysis

Supervisor: Senior Lecturer Simona Ferraro, PhD

Tallinn 2020

I hereby declare that I have compiled the bachelor’s 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 11,918 words from the introduction to the end of conclusion.

Miina Hõbenael ……………………………
(signature, date)

Student code: 179329TAAB
Student e-mail address: miinahobenael@mail.ee

Supervisor: Senior Lecturer 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

This thesis aims to find how much an employee gains from digital skills - defined as the problem-solving skills in technology-rich environments estimated by the Programme for the International Assessment of Adult Competencies (PIAAC). Besides returns to digital skills on wages, the relationship between returns to digital skills and work experience and the impact of digital skills on gender wage gap are also examined. The data from PIAAC is investigated with a log-linear regression for some selected European Union (EU) countries. The results show that higher problem-solving skills are systematically linked with higher wages in eleven out of the thirteen selected EU countries. One standard deviation increase in problem-solving skills in the pooled sample of full-time employees aged 35 to 54 is associated with an average increase in gross hourly wage by 14.9 per cent. Across all the countries in the sample of full-time employees aged between 25 to 65, the returns to skills for the group of more experienced workers - aged from 35 to 54 - are 4 percentage points higher than for the group of inexperienced workers (aged between 25 to 34). The returns to problem-solving skills for women in the main sample of prime-aged full-time workers ages 35 to 54 are 2.1 per cent higher than for men, however, the country analyses show that there is no significant difference in returns to problem-solving skills between men and women and some of the gender wage gap is traceable to the differences in digital skills between genders.

Keywords: Digital skills, labour market outcomes, international assessment data, gender wage gap

INTRODUCTION

Technology has an increasing influence on society as a whole. From self-learning programs to the ever-growing amount of tasks requiring interaction with sophisticated machinery, it is apparent there is no fleeing from the changes that come with it. With the growing impact of technology in people's day to day lives, one might start to wonder, what are the implications of technology skills on wages.

Harmon *et al*. (2003) define learning as an investment of current resources in trade for prospective repayments while Hanushek and Woessmann (2008) summarize in their article that the formation of an individuals' wage is associated with human capital. Defining human capital through labour-market skills, it appears feasible to consider technological skills beneficial in today's competitive labour market. Still, there is a limited amount of evidence about what returns on digital skills mean and how individual characteristics like age, gender and experience might affect the acquisition and accumulation of skills.

This graduation thesis aims to clarify to what degree an individual's digital skills might raise his/her wage for selected countries within the European Union (EU) area. Handled main research question is what are the returns of digital skills on earnings namely, how much an employee gains from digital skills. Three hypotheses are explored in this paper.

First, the thesis hypothesises that digital skills increase an individual's wage in the selected sample which consists of thirteen EU countries. The second hypothesis is that employees' work experience is positively correlated with returns to digital skills. It means that experienced workers gain more from digital skills than inexperienced ones. Finally, the thesis author hypothesises that the difference in digital skills can explain some of the inequality of wages between genders which implies that the gender wage gap in the sample is traceable to the variation of digital skills between men and women.

To fulfil the aim of the paper, the quantitative approach is used to conduct a regression analysis using the Organization for Economic Cooperation and Development's (OECD) Programme for the International Assessment of Adult Competencies (PIAAC) data. The OECD-PIAAC data is publicly available, covering individuals from ages 16 to 65 with different backgrounds and skill-sets. The survey measures skills in three areas: literacy, numeracy and problem-solving in technology rich environments. In the context of the thesis, problem-solving skills as well as information about individuals' background are used to estimate the returns to digital skills.

The sample covers individuals from thirteen EU countries: Belgium, the Czech Republic, Denmark, Estonia, Finland, Greece, Ireland, Lithuania, Netherlands, Poland, Slovak Republic, Slovenia, and the United Kingdom. The main sample used for estimating the returns of problem-solving skills on wages and the impact of the digital skills on the gender wage gap consists of full-time, not self-employed workers aged between 35 to 54. Another sample of full-time employees aged between 25 to 65 is used for determining the relationship between the actual work experience and returns to skills.

The added value of this thesis is in the choice of selected EU countries, using the OECD-PIAAC data from two out of all three rounds of the First Cycle of the survey, and in examining the effect on labour market outcome: individual's wage. The thesis follows the methods presented by Hanushek *et al.* (2015) but it adds differences in terms of main measurement for skills and more PIAAC rounds – Hanushek *et al.* (2015) uses data only from the first round of the First Cycle.

The structure of the thesis is as follows. The first chapter presents the theoretical background for the topic at hand and gives a solid base for the following parts. This section of the thesis explains the importance of human capital, explores the returns of formal schooling and the relationship between skills and the labour market. The last subchapter of the first chapter describes briefly the essence of digital skills and their importance today and in the future.

The second part of the thesis contains an overview of methodology and data. It gives an introduction of PIAAC data, describes the research design for checking the validity of the three hypotheses, decribes how multicollinearity and causal relationship between wages and skills is checked and delves into the criteria of the selection of the thesis sample countries. The last two parts of the second chapter outline the steps for further restricting the sample and provide the descriptive statistics of the baseline regression variables.

The third chapter of the thesis discusses and summarizes the empirical results. The first subchapter of the third part outlines the baseline regression outcomes on problem-solving skills wage returns and the second subchapter addresses the problems of causality. The third subchapter presents the returns to digital skills from regression where three control variables are added to avoid bias from omitted variables. The last two subchapters discuss the returns of skills relationship with work experience and digital skills impact on the gender wage gap. Final chapter concludes the whole thesis.

1. THEORETICAL OVERVIEW

1.1. The importance of human capital

Gained skills and experience are a type of capital, which is a result of a conscious investment. Investment in human capital is the plausible reason why the growth in national output has been greater than the growth of land, man-hours, and physical capital. (Schultz 1961)

Becker (1962) defines human capital as investing in schooling, on-the-job-training, and medical care that changes expected real income by the instilling of resources in individuals. Despite differences in the impact on earnings, all investments into human capital enhance people capabilities which in turn increase real income probabilities. The study additionally explains that the different effect of human capital on perceived earnings between the older and younger generations is caused by the fact that the profits from accumulated human capital are added to earnings at an older age as opposed to a younger age when the expenses for said human capital are subtracted.

The empirical approach of labour-market skills was significantly altered by Jacob Mincer (1970), who started measuring human capital by the amount of education received. He observes that the study of income allocation is closely linked with the human capital approach and supports this observation by noting that differences in earnings are firstly associated with returns to investments in human capital. Hanushek and Kimko (2000) further advocate the importance of schooling by determining that the labour force quality, which is associated with schooling, has a solid connection with economic growth.

The importance of investing in human capital is additionally supported by Barro (2001), whose analysis covering the growth effects of education in about 100 countries from 1965 to 1995, concludes that the long-term economic growth is positively connected with the starting level of schooling of adult males at the secondary and higher educational levels. He states that these results imply educated employees have a significant role in the distribution of technology considering workers with the aforementioned academic knowledge would match with technological innovation.

1.2. The returns on formal education

Besides other factors such as gender and age, labour market possibilities are also defined by schooling but the reason why a solid educational background is essential in getting a good job with a satisfactory wage is still unclear. Various models have been offered to clarify why people with lower levels of education are disadvantaged in the labour market compared to individuals with higher levels of education. (Barone, Van de Werfhorst 2011)

Harmon *et al.* (2003) concentrate on schooling as an individual's decision to invest in human capital and examine the returns to education. Their study explains that since the positive impact on earnings of an individual from engagement in schooling are unreasonably large compared to other investments with similar risk, there exists some sort of market failure that impedes individuals from fulfilling their personally best possible strategies. A primary issue here is that education may be valued by employers as a signal of ability rather than a variable that influences productivity. The question is whether there is a positive association between academic schooling and earnings and whether education is a sign of present productivity or it magnifies productivity.

Cunha and Heckman (2007) emphasize that it is widely recognized that individuals have distinguishable abilities, which present an important part in explaining the difference in people’s socio-economic achievements. They state that the level of ability is relevant because of its connection to wages, participation in crime and social and economic growth. An individual with inferior abilities has a lower probability of being employed in higher-wage jobs, whilst the chance of that individual committing a crime is higher. Likewise, society or country in which citizens have superior abilities, experience both social and economic growth. Large ability gaps across children develop before they begin formal learning in school which implies that education’s importance in defining labour market possibilities might be overstated (*Ibid.*).

Green and Riddell (2013) observe in their study that formal learning is the main driving force of adult literacy skills. They find that the direct influence of the other characteristics of the individual’s backround, such as the features of the parents on literacy is comparatively small in size as opposed to the notable impressions of the schooling. The results of their research also show a weak negative connection between literacy skills and age beginning in the 20s. These findings imply that literacy skills are essentially defined by formal education.

1.3. The skills and labour market

The Fourth Industrial Revolution in combination with ageing populations and workforces make skills an essential element for economic prosperity and private and social welfare (Martin 2018). According to McIntosh and Vignoles (2001) several approaches can be used for estimating employees' skill levels, including a person's education level, their qualifications and the amount of received on-and-off-the-job training. The empirical research endeavouring to quantify the returns to schooling since the 1970s is based on the fundamental work by Jacob Mincer (1974).

McIntosh and Vignoles (2001) explain that the scarce literature of skill returns on earnings is mainly due to difficulties with skill measurement and data availability. They hold the position that in the time when the demand for skills is surging, it is critical to have information about the primary skills advantages in the labour market. They assess literacy and numeracy skills influence on employees' labour market results and emphasize that there is a distinct proof of wage return to such elementary skills.

Skills can be roughly divided into cognitive and non-cognitive skills although as Borghans *et al.* (2008) point out in their paper, this type of compartmentalization should be regarded with caution since few skills are purely one or the other. Both cognitive and non-cognitive skills change and overlap during an individuals lifetime, however, unlike cognitive skills non-cognitive skills such as aspirations, personality traits and motivations can not be sufficiently measured by achievement tests (Kautz *et al.* 2014).

Green and Riddell (2003) note that cognitive skills have a notable contribution to earnings. In their work, they observe that including cognitive skills in a regression leads to a decrease in the impact of education on earnings. Their conclusion is that cognitive and non-cognitive skills are both productive but that adding one does not raise the productivity of the other and that experience does not assist the creation of cognitive skills of the variety seized in a literacy test. Their findings are also supported by Hanushek and Woessmann (2008), according to whom cognitive skills have a substantial influence on individual profits, distribution of income and economic growth in the labour market.

Hanushek *et al.* (2015) state that existing evaluations of the labour-market returns to human capital provide a malformed understanding of the purpose of skills in diverse economies, while the existing literature on the returns to skills in the labour market is restricted covering very specific data of early-career workers in the United States. They point out that the difference in skills is not entirely created by the variation in schooling and employ a different identification strategy, which is similar to the Mincer equation, except that they use cognitive skills rather than years of schooling. This approach is supported by Martin (2018) who emphisizes that from substituting skills by educational fulfilment, while controlling for a wide range of other factors considered to influence returns to investments in schooling, arises a problem as to whether cognitive and non-cognitive skills obtained outside formal learning influence returns, beyond those connected to schooling.

1.4. About the digital skills

Digital technology's gaugeable influence on multiple markets started with the emergence of the Internet, which made commercial inexpensive computer-to-computer communication widely possible. The use of digital technology decreases the cost of storage, computation and delivery of data because it is in essence information packed into small chunks. (Goldfarb, Tucker 2019)

One of today's toughest challenges is the fast evolution of digital technologies, which presents people with circumstances that require them to use an increasing variety of cognitive skills to perform and resolve problems in digital contexts (Aviram, Eshet-Alkalai 2006). Autor *et al.* (2003) observed that automation is linked with diminished labour input of routine hand-operated and cognitive tasks and increased labour input of nonroutine cognitive tasks. They underline that although the remarkable association between the adoption of computer-based technologies and the growing usage of college-educated labour is often portrayed as confirmation of the skill-biased technical change, the computers ability to replace workers in accomplishing cognitive tasks is limited.

The steps for fulfilling cognitive tasks are not completely comprehensible. Computer technology replaces employees in completing regular responsibilities which can be easily defined with programmed commands while having a supporting function in carrying out nonroutine tasks requiring flexibility, creativity, generalized problem-solving skills and complex interactions. Improvements in the quantity and quality of routine informational inputs shift the supply curve in the production function framework outward, which raises the marginal productivity of workers doing nonroutine tasks that require these inputs. Advancing computer technology influences skill demand the most through the decreasing price of computer capital because it boosts the relative demand for employees who have a relative advantage in nonroutine tasks, typically college-educated workers. (*Ibid.*)

According to Aviram and Eshet-Alkalai (2006), a mixture of technical-procedural, cognitive and emotional-social skills customarily summarizes digital literacy. Ferraro (2018) shows in her article that using digital technology such as computer, tablet or laptop enhances student performance through better learning and knowledge attainment which in turn increase students' test scores.

Hatlevik *et al.* (2015) emphasize that in literature different theories are utilised to distinguish what, and how, learners maintain, apply and learn with technology. They remark that a shared trait of approaches such as digital competence, digital skills and digital literacy is the digital core to which a knowledge aspect is added. In their work, digital competence is defined as the skills, knowledge and attitudes that enable students to use digital media separately and jointly with others in an analytical, reliable and imaginative way for participation, work and problem-solving, which means that according to them digital competence is more than skills as it also contains knowledge and attitude.

Mohammadyari and Singh (2015) hold the position that digital literacy merits more recognition in e-learning and other contexts as it includes the concept of information technology use as a skill that can be evolved. Their study purposes that the level of digital literacy affects productivity through its influence on performance and effort expectations, which suggests that individual digital literacy promotes the use of e-learning and should be recognised when analysing the impact of the latter on performance.

There is no particular determinant or action that can disclose digital competence alone. A large part of digital competence can be foreseen by looking at cultural integration, self-efficacy, strategic information use and previous academic achievements which indicates that both home environment and individual preferences appear to have an imperative role in digital competence. (Hatlevik *et al.* 2015)

The importance of digital skills has been even more pronounced in the light of the ongoing coronavirus pandemic referred to also as COVID-19. The need for social distancing and other regulations meant to prevent further spreading and damaging effects of the virus have made people turn to technological solutions to continue such everyday tasks as studying and going to work. Even information exchange with the outside world and communication with friends have largely moved to digital channels. The adaption process is more challenging for some individuals due to the excistance of a digital divide which manifests in both intellectual and physical preparedness for accessing the digital world (Schleicher 2020).

2. METHODOLOGY AND DATA

2.1. About PIAAC’s Survey of Adult Skills

Devised to be valid cross-culturally, PIAAC's Survey of Adult Skills contains a vast amount of information measuring key cognitive and workplace skills. It enables policymakers in different countries to observe the growth of fundamental aspects of human capital. Data from the survey provides an ample evidence base for policy-relevant analysis through facilitating a better understanding among others of connections between key cognitive skills and a range of variables.

The Survey of Adult Skills is conducted every 10 years for individuals from ages 16 to 65 from over forty different countries, and the data is publicly available in three different formats from the OECD Skills Survey webpage[1]. There have been two cycles of the survey: the First Cycle taking place between 2011-2018 and the Second Cycle from 2018 onward. The results of the Second Cycle will be announced in 2023.

The First Cycle consisted of three rounds of data collections. Round 1 in 2011-2012 covered twenty-four countries. These countries are Australia, Austria, Belgium (Flanders), Canada, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Japan, Korea, Netherlands, Norway, Poland, Russian Federation, Slovak Republic, Spain, Sweden, United Kingdom (England and Northern Ireland), and the United States. Data from nine more countries, including Chile, Greece, Indonesia, Israel, Lithuania, New Zealand, Singapore, Slovenia, and Turkey, were collected in the second round in 2014-2015. Round 3 in 2017 consisted of Ecuador, Hungary, Kazakhstan, Mexico, Peru, and once again the United States. The Second Cycle is going to cover among the beforementioned five new countries: Croatia, Korea, Latvia, Portugal, and Switzerland.

[1] OECD Skills Survey webpage can be found at https://www.oecd.org/skills/piaac/.

The Survey of Adult Skills consists of two parts: the background questionnaire and the direct assessment of cognitive skills. The background questionnaire includes several questions related to the individual's characteristics such as age, wage, labour status, years of schooling and working experience. The direct assessment component of the survey estimates the cognitive skills of adults in literacy, numeracy and problem-solving in technology rich environments, which present a framework for the other, higher-order cognitive skills and are necessary for an extensive array of circumstances. Several new elements such as social outcomes, quality of work environment, socio-emotional skills, adaptive problem-solving and employer survey are introduced in the Second Cycle, while problem-solving in technology rich environments is only available for the First Cycle.

The term "literacy" refers to the reading of the written texts not involving either the comprehension or production of spoken language or the production of text, while numeracy is defined as the ability to access, use, interpret, and communicate mathematical information and ideas. Problem-solving in technology rich environments combines the capacity to use a computer with the cognitive skills required to solve problems to evaluate how adults utilise digital tools for dealing with information efficiently. (OECD 2013)

2.2. Research design and methodology

This thesis follows Hanushek *et al.* (2015) and Anghel and Balart (2017) in estimating the effects of problem-solving skills to the earnings of an individual. Both works use numeracy scores as an estimator for cognitive skills in their main analysis, Anghel and Balart (2017) however mainly concentrate on the effects of non-cognitive skills on earnings.

The empirical model used in this thesis is based on the work by Hanushek *et al.* (2015) and it is as following:

$$ln y_i = \beta_0 + \gamma C_i + \beta_1 E_i + \beta_2 E_i^2 + \beta_3 G_i + \varepsilon_i$$

where
$y_i -$ the hourly wage of an individual i,
$C_i -$ individual's i mean score value from plausible values in the PIAAC test,
$E_i -$ years of labour market experience, measured either as actual or potential work experience,
$G_i -$ gender indicator,
$\varepsilon_i -$ stochastic error term.

The focus in on the effect of gamma γ on the output. The methodology applied in the current thesis is a secondary regression analysis based on the detailed database. PIAAC data is intended to be a panel data in the future when the Second Cycle provides information about the variables in multiple points of time, however in the context of this paper the cross-sectional data from the First Cycle is used.

Regarding the question about the appropriate functional form of the earnings relationship Anghel and Balart (2017) point out that the log-linear model has been a commonly used specification. The work by Mincer (1974), Heckman and Polachek (1974) and Welland (1978) indicate that the best functional form of the relationship between earnings and human capital is log-linear. Furthermore, Hanushek *et al.* (2015) also applied in their analysis the linear specification which did not pick up the nonlinearities which were caught by the log-linear form.

The underlying PIAAC names, labels and the expected signs of the variables used for the variables in regression analyses are indicated in Table 1. Hourly earnings is the dependent variable and it is derived excluding bonuses for wage and salary earners and corrected with the purchasing power parity (PPP) in US dollars. This is needed to ensure the comparability of wages across countries. Labour market experience expressed in years is the actual work experience. The survey participants are instructed to count only those years in work experience where they have worked full-time or part-time at least for 6 months.

Another variable which could be used for estimating the labour market experience is potential work experience calculated as age minus years of schooling minus six. The main challenge with the actual work experience is that it might in part be endogenous to skill levels as one of the channels by which skills influence the earnings could be experience (Hanushek *et al.* 2015; Anghel, Balart 2017). The use of actual and potential work experience is analysed by both Hanushek *et al.* (2015) and Anghel and Balart (2017) and neither note a significant change in the results.

The thesis hypothesises that digital skills increase an individual’s wage in the selected sample of EU countries. Digital skills are determined by the individual’s test scores in the problem-solving domain. PIAAC’s test scores are marked on the 500-point scale and an impartial assessment of an individual’s skills is provided by the average of ten plausible values available for all three domains. Plausible values are a statistical means to estimate the scores based on the answers to the subset of skill evaluation questions and the respondent's background information since not all the questions pertaining the skill are administered to the respondents (Caro, Biecek 2017). It is important to note that due to the complex nature of PIAAC data, it is not advisable to use only one plausible value for estimations and it pays to be careful while analysing this data.

Table 1. Underlying variables names and labels in PIAAC and the expected signs for variables used in the regression analyses

Name in PIAAC	Label in PIAAC	Expected association
Logarithm of earnings lny_i		
EARNHRPPP	Hourly earnings excluding bonuses for wage and salary earners, PPP corrected $US (derived)	Dependent variable
Individual's i mean score value in the PIAAC test C_i		
PVPSL1	Problem-solving scale score - Plausible value 1	+
PVPSL2	Problem-solving scale score - Plausible value 2	+
PVPSL3	Problem-solving scale score - Plausible value 3	+
PVPSL4	Problem-solving scale score - Plausible value 4	+
PVPSL5	Problem-solving scale score - Plausible value 5	+
PVPSL6	Problem-solving scale score - Plausible value 6	+
PVPSL7	Problem-solving scale score - Plausible value 7	+
PVPSL8	Problem-solving scale score - Plausible value 8	+
PVPSL9	Problem-solving scale score - Plausible value 9	+
PVPSL10	Problem-solving scale score - Plausible value 10	+
Labour market experience E_i and labour market experience squared E_i^2		
C_Q09	Current status/work history - Years of paid work during lifetime	+
Gender indicator G_i		
GENDER_R	Person resolved gender from BQ and QC check (derived)	−

Notes:

1. Variable names and labels in PIAAC are accessible from the International Codebook for PIAAC Public Use Files, which can be found on the OECD Skills Surveys webpage under documentation section of Data and Tools.
2. "+" indicates positive association, "−" indicates negative association.
3. "BQ" -Background Questionnaire, "QC" - Quality Control.
4. GENDER_R expected association is negative since the gender indicator implies that the individual i is female.

PIAAC's computer-based estimation has an adaptive multistage design, which means that the questions depend on the answers that test takers have given before (Anghel, Balart 2017). PIAAC among the Programme for International Student Assessment (PISA), the Trends in International Mathematics and Science Study (TIMSS), and the Progress in International Reading Literacy Study (PIRLS), is an international large-scale assessment, which design includes complex sampling and testing methods that influence the analysis of the registered data. Caro and Biecek (2017) emphasize that this type of design makes it challenging to determine the standard error since the introduced two sources of error cannot be computed with conventional routines of analytical software.

The total standard error for variables observed once is equal to the sampling error. The computation of the standard errors of plausible values, however, require replication techniques (Caro, Biecek 2017). Anghel and Balart (2017) use one replication technique for two-stage sampling called "jackknife" replicate procedure developed by Maurice Quenouille in 1949 for computing the standard errors. Current thesis utilizes for analysis the free software environment R which has a package called "intsvy" meant for working with the international assessment data like PIAAC. This package and its merits are the main topics of the paper by Caro and Biecek (2017).

The second hypothesis of the paper is that employees' work experience is in positive correlation with returns to digital skills which means that experienced employees gain more from digital skills than inexperienced ones. Correlation analysis between actual work experience and age is conducted to confirm the positive relationship between the two since while potential work experience increases with the employee's age by default it is not given with the actual work experience. The sample of full-time employees aged 25-65 is then divided into three categories as per Hanushek *et al.* (2015): entry-age workers aged 25-34, prime-age workers aged 35-54 and exit-age workers aged 55-65 and the returns to problem-solving skills in technology rich environments are allowed to differ between these categories.

Finally, the third hypothesis of the paper that the difference in digital skills can explain some of the inequality of wages between genders is analysed similarly to the work of Hanushek *et al.* (2015). This hypothesis implies that some of the gender wage gap in the sample is traceable to the difference in digital skills between men and women. The validity of the third hypothesis is evaluated by checking the heterogeneity in wage determination since, given that the returns to skills between men and women are alike, it is possible to estimate the part of cognitive skills in explaining the gender wage gap by comparing the gender wage gap with and without conditioning on skills (Hanushek *et al.* 2015).

2.3. Dealing with multicollinearity and the causal relationship

Multicollinearity in the regression analysis has the potential to influence estimated coefficients since the least square estimators of coefficients have large variances (Mansfield, Helms 1982). All the regression analyses in this thesis are checked for multicollinearity with either Farrar-Glauber test or looking at the variance inflation factor (VIF). Independent variables actual work experience and experience squared are expected to be multicollinear, however, since they are variables modelling a concave relationship between experience and earnings, the multicollinearity between them is for the most part ignored. Allison (2012) explains that multicollinearity can be ignored when the variables of interest have low variance indicator factors and when the multicollinear variables include powers.

Hanushek *et al.* (2015) bring out three potential threats to the causal relationship between skills and earnings: measurement error in skills variable, reverse causation and omitted variables. The current thesis is going to use similar approaches to the *Ibid.* to address these threats. Measurement error in skills variable is handled with instrumental variable approach since in PIAAC context it is possible to use numeracy skills as an instrument for problem-solving skills. The relevant cognitive dimension is, therefore, the variation shared in both skill variables.

Issues with reverse causation caused by, for example, the possibility that the people with better jobs have better skills can be addressed by instrumental variables such as schooling which is a skills estimator observed before truly entering the labour market. The final concern of omitted variables is addressed adding independent variables to the regression analysis of returns to digital skills to see the robustness of the variables of interest. Table 2 displays the underlying variables names and labels in PIAAC and the expected signs of the added independent variables.

As Hanushek *et al.* (2015) point out none of the before-mentioned approaches solve the problem of causality. The various strategies used for dealing with possible biases, however, indicate if there is a bias caused by measurement error, reverse causality, and omitted variables and is it upward or downward bias.

Table 2. Underlying variables names and labels in PIAAC and the expected signs for the extra variables used in the extended regression analyses for controlling for omitted variables

Name in PIAAC	Label in PIAAC	Expected association
Schooling S_i		
YRSQUAL	Highest level of education obtained imputed into years of education (derived)	+
Computer-use at work CW_i		
ICTWORK	Index of use of ICT skills at work (derived)	+
Native speaker N_i		
NATIVESPEAKER	Respondent is a native speaker (derived by CAPI)	+

Notes:

1. Variable names and labels in PIAAC are accessible from the International Codebook for PIAAC Public Use Files, which can be found on the OECD Skills Surveys webpage under documentation section of Data and Tools.
2. "+" indicates positive association.
3. "ICT skills"- information communication technology skills
4. "derived by CAPI"- variable derived through the computer-assisted personal interview (CAPI)
5. For underlying variable names and labels for logarithm of earnings, mean score value, labour market experience, labour market experience squared, and gender indicator see Table 1.

2.4. Criteria of thesis sample selection

The sample of this thesis consists of the PIAAC data of thirteen countries from the First Cycle. The European area map in Figure 1 illustrates the three-fold selection process. Only the countries, which are in the EU are chosen since the EU countries' goals for the labour market are relatively similar. The article 3 point 3 of the Treaty on the European Union states that "the Union shall work for the sustainable development of Europe based on balanced economic growth and price stability, a highly competitive social market economy, aiming at full employment and social progress, and a high level of protection and improvement of the quality of the environment".

EU countries covered by PIAAC are divided into three categories and are either green or yellow on the figure. The only exception is Cyprus which has not been included in the selection process even though it is part of the European Union. OECD Skills Outlook (2013) stresses that the information about Cyprus relates only to the southern part of the island as there is no single authority representing all the people on the island. The PIAAC data about Cyprus was not publicly available from OECD webpage at the time of writing this thesis.

The main research question of the thesis is how much does an employee gain from digital skills. Digital skills estimation is based on the data of the problem-solving in technology rich environments scores. Problem-solving domain, however, was optional and this excludes three more countries from the sample since France, Italy and Spain did not participate in the problem-solving skills survey. Figure 1 shows countries included in the sample for this thesis. Similarly to Anghel and Balart's (2017) approach, countries for which continuous data on earnings is not available are also excluded. That constraint applies for Austria, Germany, Sweden and Hungary which are displayed as light green on the map.

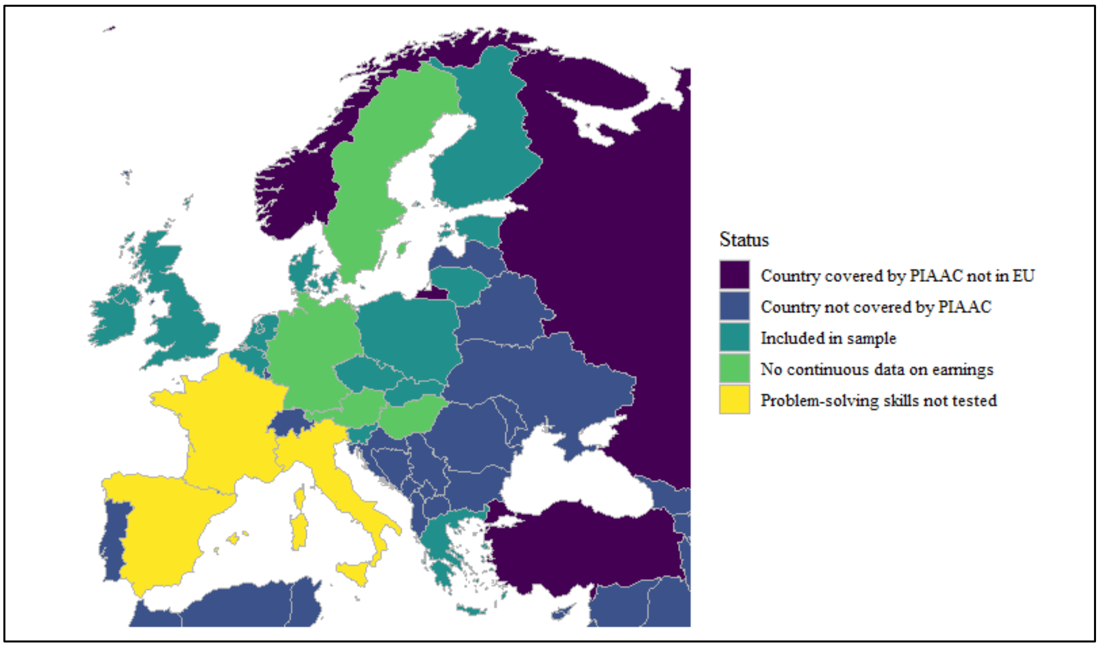

Figure 1. The map of the European area filled in by the status of the country for the thesis
Source: Created by author in R based on information obtained from the PIAAC

The final sample consists of thirteen countries (Figure 1) represented in bluish green (Belgium, the Czech Republic, Denmark, Estonia, Finland, Greece, Ireland, Lithuania, Netherlands, Poland, Slovak Republic, Slovenia and the United Kingdom). The United Kingdom is included in the sample since it was a member country of EU at the time of the testing as it withdrew from the union in January 2020.

2.5. Working with the dataset

As mentioned before, in the research design and methodology, this thesis utilizes an R package called “intsvy” which takes into consideration the nature of the survey. Caro and Biecek (2017) discuss several different ways for importing data into R one of which is accessing the data from PIAAC study by using the R package with converted data which due to its significant size can be downloaded from the “pbiecek” account on GitHub.

It is important to note that this package meant for accessing the PIAAC data contains only the data from the first round of the First Cycle with the exclusion of Australia for which the access to the public use data file needs to be required separately. Hence, for the three of the sample countries Greece, Lithuania and Slovenia not covered with this approach the data is downloaded from the OECD Skills Surveys webpage, imported to R, converted and joined vertically with the data of the other sample countries from the package.

The PIAAC names, PIAAC labels, value labels and value types of the variables which were used to further define the sample of the thesis can be found in the Appendix 1. Observations of all the individuals who did not have the plausible values of problem-solving scores were removed from the sample by setting the variable PSLSTATUS to “has PV” and dropping all the other rows where the variable had another value. The 82,472 observations in the sample dropped to 61,027 observations after applying the constraint. Taking the computer-based form of the assessment is necessary to handle the problem-solving portion of the survey and hence the 26% drop in observations is in line with the OECD (2016) conclusion that on average the computer-based form of the assessment was taken by a 73.5% of the respondents.

The works of Hanushek *et al.* (2015) and Anghel and Balart (2017) were followed by further limiting the main thesis sample to prime-age full-time and non-self-employed workers that is employees aged between 35 and 54 years working at the time of the survey, with at least 30 hours per week. Instead of hours worked per week, the current thesis uses the question “C_Q07” whether a respondent works full-time for determining the respondent’s workload. This is needed because, over three-quarters of the values of the variable “D_Q10”, which covers data of current working hours, are missing values. These constrictions lead to a homogeneous sample with strong labour-force commitment and the best estimates of long-run returns to skills as according to Hanushek *et al.* (2015) they limit the influence of factors that might affect labour force attachment like healthlimitations or family demands and isolate the direct labour market effects on skills. The appropriate values of the variables given in Appendix 1 are used to restrict the data.

Using sample of observations for representing populations means a sample selection bias can appear (Certo *et al.* 2016). This is highly important in the context of this thesis as Hanushek *et al.* (2015) point out because skills may also influence the individual's probability of participating in the labour market. For this reason, a two-step Heckman regression model is used to eliminate the concern of selection bias in restricting the sample of respondents to individuals active in the labour market. The probit selection equation dependent variable is a dummy variable indicating if the individual is active in the labour market and independent variables age, age squared, living with a partner or spouse, having kids and schooling are included as variables that might influence the participation in the labour market. Heckman regression model's outcome equation dependent variable is logarithm of hourly wage while experience, experience squared, and schooling are independent variables.

Top and bottom of one percent of wage distribution is trimmed in each sample country and in the pooled sample to limit the influence of outliers since Hanushek *et al*. (2015) found that these outliers largely represent data entry or coding errors. Actual work experience is compared with the age of the respondent and set to match the criteria that actual work experience cannot be smaller than 1 year and bigger than 39 years in the sample of full-time employees aged between 35 to 54. These constraints prevent the unlikely observations where an individual has been working at least six months per year since before being at least 15 years old or claims to be full-time worker while having no actual working experience.

PIAAC scores in problem-solving in technology rich environments are standardized in the subsequent regression analyses for easier interpretation to have a zero mean and standard deviation equal to one. This ensures that γ or gamma is the percentage increase in wages when there is a one standard deviation increase in skills.

The process of achieving the other sample used in the thesis of full-time employees aged between 25 to 65 follows the same steps outlined previously for the main sample. The only difference between the samples is the age scope since the sample used for analysing the experience patterns needs to cover in addition to the prime-age workers both entry-age and exit-age workers.

2.6. Descriptive statistics

Descriptive statistics of the main variables across the pooled sample are provided in Table 3 that covers 11,775 observations from thirteen different countries. The table contains the mean values, standard deviations and the minimum and maximum values. Descriptive statistics per country including means and standard deviations are available for perusal in Appendix 2.

Table 3. Descriptive statistics of the pooled sample full-time employees aged between 35 to 54

Variable	Mean	Standard Deviation	Minimum	Maximum
Gross hourly wage in PPP US$	16.60	9.45	2.50	45.27
PIAAC score in problem-solving in technology rich environments	277.31	40.07	127.53	460.57
Experience (years)	22.13	6.92	2.00	39.00
Female (share)	0.49	0.50	0.00	1.00

Source: Author's calculations based on data obtained from OECD
Notes:

1. Observations, means and standard deviations per country can be found in the Appendix 2.
2. No special functions or packages used for calculating the descriptive statistics.

The descriptive statistics show that the average gross hourly wage in purchasing power parity in the United States dollars, in the pooled sample of prime-age full-time employees, is \$16.60 and the average PIAAC score in problem-solving in technology rich environments is 277.31, which is just a little above the middle of the evaluating scale of 500. Both variables' minimum and maximum values indicate that there are notable differences across the values, however looking at both the standard deviation and the minimum and maximum values it is clear that the PIAAC scores are distributed more evenly than the gross hourly wages since the mean of the gross hourly wages is closer to the minimum value, which indicates that there are fewer values in the upper part of the scale.

The average actual work experience for individuals being between age 35-54 is approximately 22 years. The minimum actual work experience is having spent 2 years for at least six months in the full-time or part-time work. Maximum experience of 39 years seems plausible in the context of the age composition since it means that at least one individual must have worked at least six months a year, starting from age 16.

Table 3 also shows that with 49% of the individuals being women, the pooled sample is evenly distributed between the genders. The minimum and maximum illustrate the coding of the variable gender which is a dummy variable with zero being assigned for males and one for females.

Figure 2 gives a brief overview of the mean and standard error of gross hourly wages in PPP US$ in both the pooled sample of prime-age full-time employees and across countries. The means are marked in the figure as black dots while standard errors are depicted by the reaches of the black lines surrounding the means. The standard error shows how far the sample mean is anticipated to be from the actual population mean while the standard deviation discussed in Table 3 and calculated across countries in Appendix 2 measures the variability from the mean.

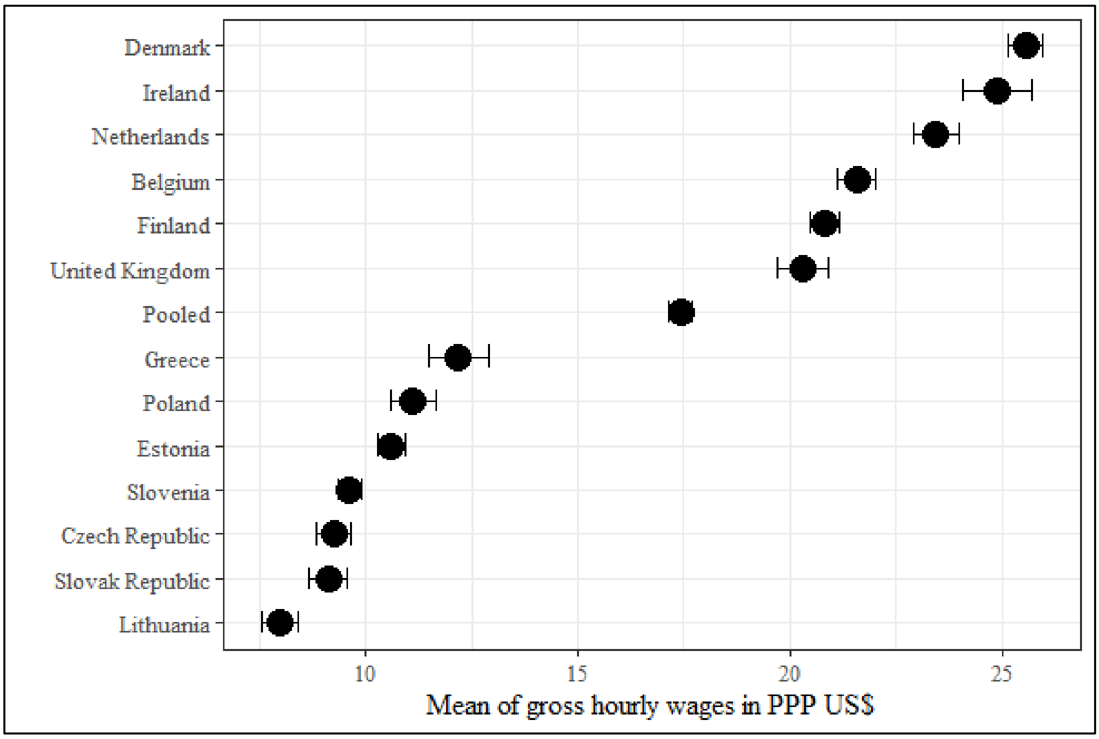

Figure 2. The mean and the standard error of gross hourly wages in PPP US$ in the pooled sample of prime-age full-time employees and across countries
Source: Author’s calculations in R using the package “intsvy” based on data obtained from OECD PIAAC results of the First Cycle

The underlying data of Figure 2 is presented in Appendix 3 since the mean of gross hourly wages is calculated in the R package “intsvy” with the function “piaac.mean” as opposed to the function used for Table 3 and Appendix 2. Comparison of the data from different calculation methods shows that there are no big differences between the results.

The highest mean gross hourly wage according to our sample 25.56 US dollars receive employees in Denmark while in Lithuania the workers get the lowest mean hourly wage of 7.98 US dollars. Pooled sample average gross hourly wage of 17.42 US dollars splits the sample clearly in half with Denmark, Ireland, Netherlands, Belgium, Finland and United Kingdom having higher mean hourly wages and Greece, Poland, Estonia, Slovenia, Czech Republic, Slovak Republic and Lithuania having lower mean hourly wages than the pooled sample average.

The mean and standard error for problem-solving skills in technology rich environments in the pooled sample and across countries are illustrated on Figure 3. The data underlying the Figure 3 can be found in Appendix 4 where the "piaac.mean.pv" function from the "intsvy" package has been used in the mean calculation process as opposed to the Table 3.

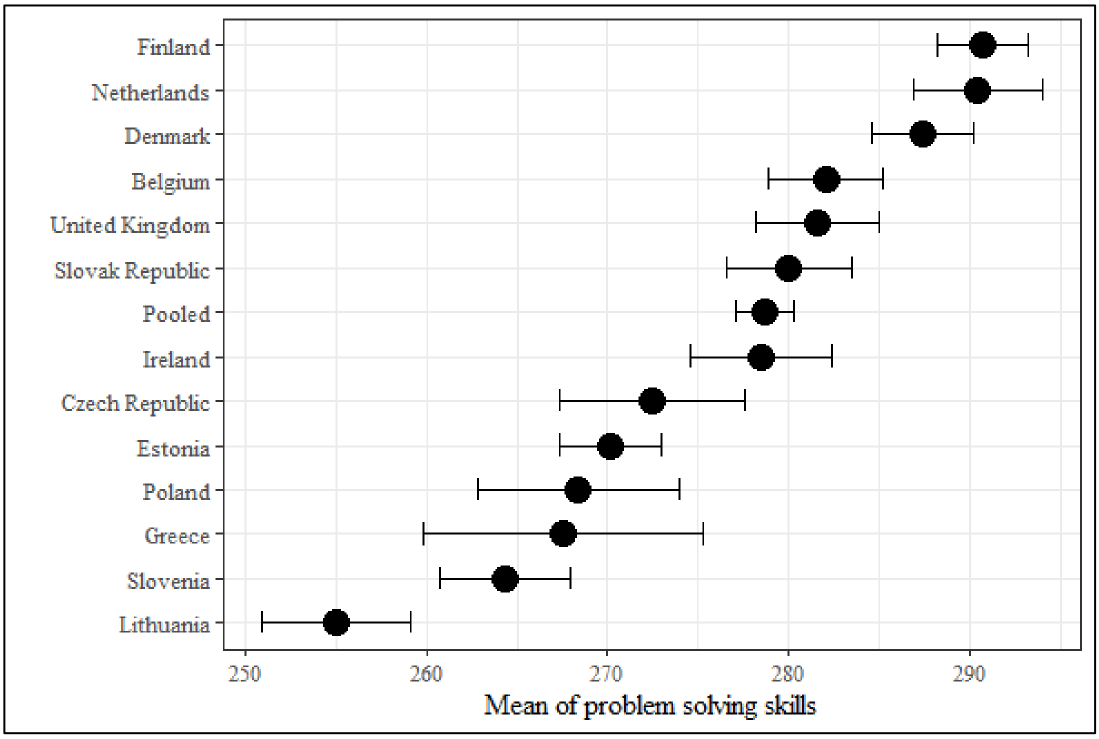

Figure 3. The mean and the standard error of problem-solving skills in technology rich environments in the pooled sample of prime-age full-time employees and across countries
Source: Author's calculations in R using the package "intsvy" based on data obtained from OECD PIAAC results of the First Cycle

The highest mean of problem-solving skills is in Finland where the average score is 290.71 out of the scale of 500. Lowest average score of problem-solving from the sample of prime-age full-time employees at 255.07 is in Lithuania which means that all the sample countries' mean of problem-solving skills are in the upper part of the scale. The pooled sample average score is 278.67 which similarly with the previous figure halves the sample countries into two groups. Comparing the Figures 2 and 3 a pattern emerges: countries which have the mean of problem-solving skills higher than the pooled average also have higher mean of gross hourly wage with the exceptions of Slovak Republic and Ireland.

3. EMPIRICAL ANALYSIS

3.1. Baseline results for returns to problem-solving skills

The relationships between the variables used for identifying returns to problem-solving skills are analysed with a correlation matrix which can be found in Table 4. Very high positive correlation between the variables experience and experience squared is expected since experience squared is directly derived from the variable experience.

Table 4. Correlation matrix of variables used in the initial regression analysis for returns to problem-solving skills

	Wage logarithm	Problem-solving	Experience	Experience squared	Female
Wage logarithm	1.000	0.396	0.057	0.053	-0.217
Problem-solving	0.396	1.000	-0.193	-0.200	-0.086
Experience	0.057	-0.193	1.000	0.984	-0.075
Experience squared	0.053	-0.200	0.984	1.000	-0.070
Female	-0.217	-0.086	-0.075	-0.070	1.000

Source: Author's calculations in R based on data obtained from OECD
Notes:

1. Underlying PIAAC variable names and labels in Table 1.
2. "Experience squared" is the squared value of experience in years divided by 1,000.
3. Observations 11,775.

The highest positive correlation 0.40 besides between the variables experience and experience squared is between the dependent variable logarithm of hourly earnings without bonuses and the standardized mean of problem-solving plausible values. The highest negative correlation -0.22 is between the dependent variable and gender indicator.

The regression results with and without the independent variable experience squared are presented in Table 5. It is clear from the first row of the Table 5 that the experience squared variable is not statistically significant. Farrar-Glauber test as expected from the correlation matrix in Table 4 detects multicollinearity between the variables experience and experience squared, however since the results of the regression do not change drastically when the variable experience squared is discarded from the model in the second row of the Table 5 and the variables model the concave relationship between experience and wages the results of the model with experience squared are discussed.

Table 5. Pooled sample initial regression with and without the variable experience squared in the sample of full-time employees aged between 35 to 54

Regression type	Problem-solving	Experience	Experience squared	Female
With experience squared	0.253*** (0.005)	0.009** (0.004)	0.038*** (0.092)	-0.216*** (0.010)
Without experience squared	0.253*** (0.005)	0.011*** (0.001)	–	-0.216*** (0.010)

Source: Author's calculations in R based on data obtained from OECD
Notes:

1. No special R packages used for calculations.
2. Dependent variable is the logarithm of hourly wages adjusted to PPP in $US, problem-solving score standardized to standard deviation of 1 and experience squared divided by 1,000.
3. Standard errors in parenthesis.
4. Significance Level values *, **, *** indicate significance at 10%, 5% and 1% level.
5. Both regression models cover 11,775 observations, multiple R-squared is 0.205, adjusted R-squared is 0.205 and are statistically significant on 1% level.

The results given in Table 5 show that better digital skills estimated by problem-solving scores are significantly related to higher labour market earnings. Problem-solving coefficient in the pooled estimation across 13 European Union countries in a sample of prime-age full-time employees implies that a standard deviation increase in the problem-solving skills is associated with an average increase in hourly wages, adjusted to PPP, of 25.3 per cent. The influence of independent variables on dependent variable match with the expected associations from Table 1 with the increase in problem-solving skills and experience also increasing the hourly wage and gender indicator showing female gender bias. According to the results in Table 5, the earnings disadvantage of women is *ceteris paribus* 21.6 per cent.

The individual wage impact of 25.3 per cent of problem-solving skills is very high compared to the Hanushek *et al.* (2015) estimation of 14.3 per cent. It is possible that the model is picking up the impact of other cognitive skills such as literacy and numeracy under problem-solving skills coefficient or there are problems with causal identification of the relationship between skills and earnings which means that the model could be overestimating the influence of problem-solving skills in technology rich environments on wages.

3.2. Addressing the potential problems in causality

One of the potential threats to the causality is the measurement error. Table 6 contains the results of the approach for addressing this problem described in detail in the third subchapter of the second part. The first row of Table 6 describes the pooled sample ordinary least squares (OLS) regression results which illustrate the first stage of the two-stage least squares (2SLS) given in the second row.

Table 6. Measurement error in the sample of full-time employees aged between 35 to 54

Method	Dependent variable	Problem-solving	Numeracy	Experience	Experience squared	Female
OLS	Problem-solving	–	0.782*** (0.006)	-0.018*** (0.004)	0.087 (0.100)	0.026** (0.011)
2SLS	Wage logarithm	0.256*** (0.007)	–	0.009** (0.004)	0.040 (0.092)	-0.216*** (0.010)

Source: Author's calculations in R based on data obtained from OECD
Notes:

1. Methods: "OLS"- ordinary least squares, "2SLS"- two-stage least squares
2. Each row describes the results of its own regression analysis.
3. Standard errors in parenthesis.
4. Significance Level values *, **, *** indicate significance at 10%, 5% and 1% level.
5. Observations 11,775.

The first row of the Table 6 implies that numeracy with the point estimate of 0.78 is a strong instrument for problem-solving skills. Comparing the results between Table 5 and the second row of the Table 6, the estimates on problem-solving skills are very similar which indicates that there is no bias related to the domain-specific measurement error.

Reverse causation is the second potential danger to the causal relationship between skills and wages. Table 7 uses years of schooling as an instrument for problem-solving skills in the two-stage least squares where the estimate of problem-solving skills is approximately 2.5 times higher than in Table 5. According to Hanushek *et al.* (2015), the two-stage least squares results should not be interpreted as causal effects, however, the estimates in Table 7 imply that the reverse causation may not create upward bias in ordinary least squares estimates of returns to skills.

Table 7. Reverse causation in the sample of full-time employees aged between 35 to 54

Method	Dependent variable	Problem-solving	Schooling	Experience	Experience squared	Female
OLS	Problem-solving	–	0.137*** (0.003)	0.010 (0.007)	-0.644*** (0.150)	-0.253*** (0.017)
2SLS	Wage logarithm	0.652*** (0.018)	–	0.005 (0.005)	0.404*** (0.114)	-0.137*** (0.013)

Source: Author's calculations in R based on data obtained from OECD
Notes:

1. Methods: "OLS"- ordinary least squares, "2SLS"- two-stage least squares
2. Each row describes the results of its own regression analysis.
3. Standard errors in parenthesis.
4. Significance Level values *, **, *** indicate significance at 10%, 5% and 1% level.
5. Observations 11,710.

The final threat for identification of causation adressed in this thesis are the omitted variables. This problem is assessed by adding three variables to the regression analyses: years of schooling, computer-use at work, native language. The correlation matrix between the variables used in the regression in Table 5 and the added three variables can be found in Appendix 5. The highest correlation 0.98 is again as expected between the variables experience and experience squared and the second highest positive correlation of 0.32 is between the dependent variable and problem-solving skills. The relationship between the dependent variable and gender indicator has the highest negative correlation -0.24 while there is also surprisingly a negative correlation between dependent variable and the native speaker indicator. The negative relationship between the native speaker indicator and wages in the pooled sample could be explained with the migration of skilled workers to and within European Union countries since according to European Commission (2018) most movers are of prime working age. Another thing to note from the *Ibid.* is that the skilled movers are not equally distributed between the EU countries which can be roughly categorized to destination countries like Belgium, United Kingdom, Denmark, Netherlands, and outgoing countries like Poland, Slovakia, Estonia, Lihtuania.

The steps and results of adding the three control variables are illustrated in detail in Appendix 6. Returns to a one standard deviation increase in digital skills dropped almost by half from 25.3 per cent to 14.9 per cent when the control variables were added which implies that the previous regression analysis (see Table 5) of the skill returns was biased upwards due to omitted variables. The adjusted regression analysis results are in line with the results of Hanushek *et al.* (2015) hence the three control variables are added to all following regression analyses for checking the validity of the three proposed hypotheses in this thesis.

3.3. Returns to problem-solving skills

The extended regression results for returns to problem-solving skills with and without the independent variable experience squared can be found in Table 8. As before with the baseline returns to problem-solving skills, multicollinearity exists between variables experience and experience squared, however, the results from the model with the mentioned variable are discussed further.

Table 8. Pooled sample regression with and without the variable experience squared in the sample of full-time employees aged between 35 to 54

Regression type	Problem-solving	Experience	Experience squared	Female	Schooling	Computer-use at work	Native speaker
With experience squared	0.149*** (0.007)	0.008* (0.004)	0.119 (0.098)	-0.237*** (0.011)	0.051*** (0.002)	0.063*** (0.006)	-0.177*** (0.022)
Without experience squared	0.149*** (0.007)	0.014*** (0.001)	–	-0.237*** (0.011)	0.051*** (0.002)	0.063*** (0.006)	-0.178*** (0.022)

Source: Author's calculations in R based on data obtained from OECD
Notes:

1. No special R packages used for calculations.
2. Dependent variable is the logarithm of hourly wages adjusted to PPP in $US, problem-solving score standardized to standard deviation of 1 and experience squared divided by 1,000.
3. Standard errors in parenthesis.
4. Significance Level values *, **, *** indicate significance at 10%, 5% and 1% level.
5. Both regression models cover 9,282 observations, multiple R-squared is 0.224, adjusted R-squared is 0.223 and are statistically significant on 1% level.

Similarily with Table 5, results in Table 8 imply that higher problem-solving skills in technology rich environments are significantly related to higher wages and there is no statistically significant concave relationship between hourly wage and actual work experience in the pooled sample. A standard deviation increase in problem-solving skills in the sample of full-time employees aged between 35 to 54 from thirteen EU countries is associated with an average increase of 14.9 per cent, in hourly wages. All the signs of independent variables match with the expectations in Tables 1 and 2, except the sign of the native speaker indicator which is negative. The potential reason for the seeming disadvantage of being a native speaker in the pooled sample was briefly explained in the previous subchapter about addressing the problems in causality. Comparing the results between the Tables 5 and 8, the earnings disadvantage of women was downward biased by the absence of the three control variables in Table 5. Gender wage gap 23.7 per cent in Table 8 is 2.1 percentage points higher than 21.6 per cent displayed in Table 5.

Looking at the regession results by country in Table 9, it is clear that the impact of problem-solving skills on wages varies considerabely between the countries. The highest returns to problem-solving skills varying approximaltely between 7 and 8 per cent are in Estonia, Netherlands, Slovak Republic and United Kingdom while in both Czech Republic and Greece the impact of problem-solving skills on wages is not statistically significant.

Table 9. Regression by country with the variable experience squared

Country	Problem-solving	Experience	Experience squared	Female	Schooling	Computer-use at work	Native speaker
Belgium	0.054***	0.013	-0.025	-0.076***	0.044***	0.038***	-0.032***
Czech Republic	0.002	0.006	-0.160	-0.202***	0.056***	0.061***	-0.024
Denmark	0.037***	0.005	-0.050	-0.100***	0.031***	0.062***	0.059***
Estonia	0.069***	0.008	-0.181	-0.427***	0.052***	0.102***	0.061
Finland	0.037***	0.010*	-0.051	-0.198***	0.051***	0.073***	0.052
Greece	0.034	0.022	-0.005	0.018	0.085***	0.007	-0.158
Ireland	0.038**	-0.001	0.303	-0.035	0.050***	0.020	0.046
Lithuania	0.055***	0.012	-0.276	-0.238***	0.067***	0.087***	0.038
Netherlands	0.065***	-0.002	0.200	-0.033	0.054***	0.052***	0.049
Poland	0.059**	0.041**	-0.736*	-0.186***	0.095***	0.047*	-0.394***
Slovak Republic	0.069***	0.011	-0.190	-0.301***	0.064***	0.060***	0.009
Slovenia	0.049***	0.021**	-0.324	-0.125***	0.123***	0.034**	0.069
United Kingdom	0.082***	-0.008	0.322*	-0.111***	0.055***	0.079***	-0.060

Source: Author’s calculations in R based on data obtained from OECD
Notes:

1. No special R package used for calculations.
2. Dependent variable is the logarithm of hourly wages adjusted to PPP in $US, problem-solving score standardized to standard deviation of 1 by country and experience squared divided by 1,000.
3. Significance Level values *, **, *** indicate significance at 10%, 5% and 1% level.
4. For further information about the regression by country see Appendix 7.

The regression results by the country in Table 9, on average suggest a male-female earnings gap. There are, however, noticeable differences between the countries. In the sample of full-time employees aged 35 to 54 there is no significant gender wage difference in Greece, Ireland and Netherlands. The other ten countries can be roughly divided in two: four countries with wage difference of 15 per cent and under and six countries with male-female earnings gap of 19 per cent and over. Estonia's gender wage gap of 43 per cent is the biggest in the current sample and the second biggest gender wage gap of 30 per cent is in Slovak Republic.

The results in Table 9 are calculated with the independent variable of experience squared divided by 1,000 even though the experience squared is not statistically significant in over the half of sample countries. Results calculated without the independent variable of experience squared can be found in Appendix 7.

The wage impact of problem-solving skills variation across countries is further illustrated on Figure 4. Returns to skills in Czech Republic and Greece are not shown since they are not statistically significant. The data underlying the figure is available for perusal in Table 8 for pooled sample and in Table 9 or Appendix 7 for across countries.

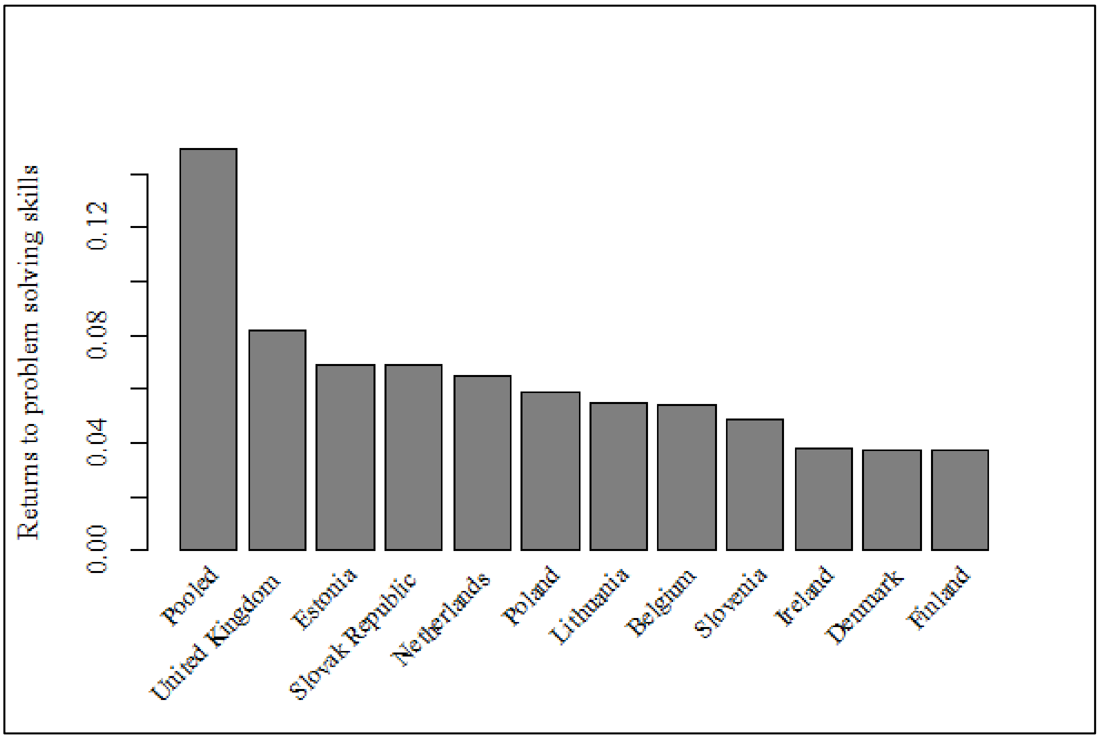

Figure 4. Returns to problem-solving skills in technology rich environments in a sample of prime-age full-time employees calculated without speacial functions
Source: Author's calculations in R based on data obtained from OECD PIAAC results of the First Cycle

3.4. Work experience and returns to digital skills

The first step of evaluating the second hypothesis of the thesis that employees' work experience is in positive correlation with returns to digital skills, is to show the correlation between actual work experience and age. Strong positive correlation between these two variables is important for the analysis because it makes possible to use age categories as a proxy for experience since it indicates that age and experience change together in the same direction.

Figure 5 shows the correlation between actual work experience and age in the sample of full-time employees between ages of 25 to 65. Pearson's statistically significant correlation coefficient of 0.94 as well as the layout of the data points in Figure 5 indicate that the correlation between the two variables is strong and positive which means that age categories can be used as a proxy for actual work experience in the subseqeuent regression analysis.

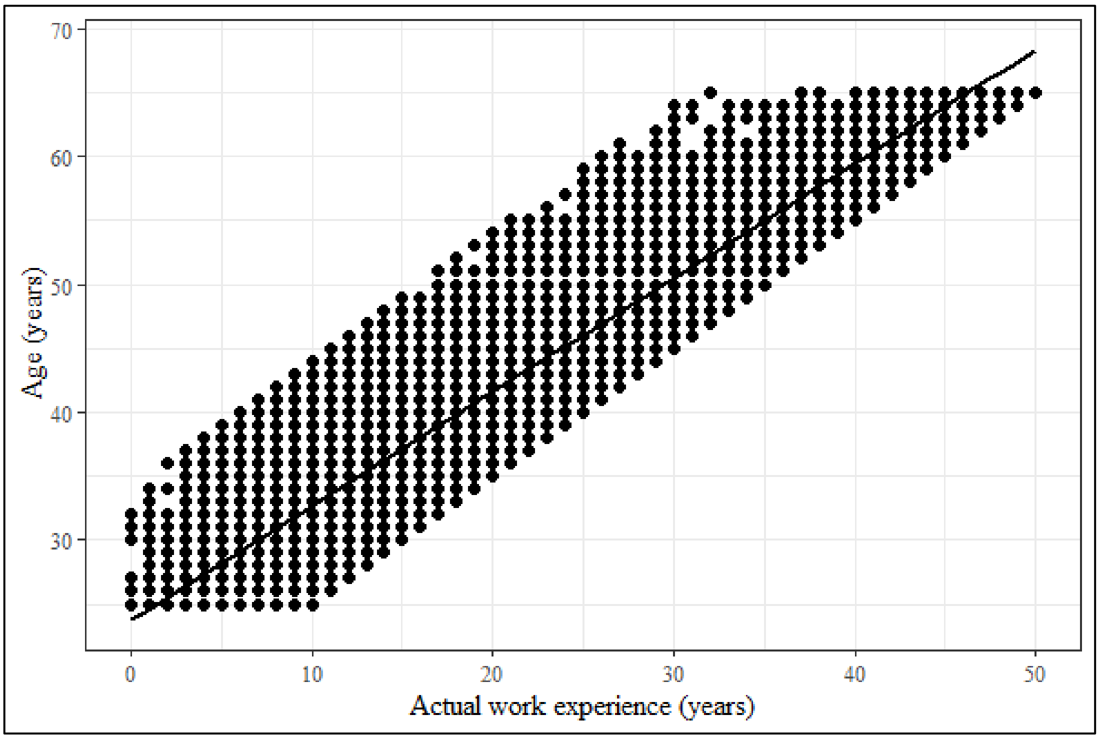

Figure 5. Correlation between the actual work experience and age in the sample of full-time employees aged between 25 to 65
Source: Author's calculations in R based on data obtained from OECD PIAAC results of the First Cycle

Tables 10 and 11 include dummy variables and interaction terms of age groups and problem-solving skills for the age categories of prime-age aged 35-54 and exit-age aged 55-65 full-time employees to test for age varying or in the context of this thesis, the actual work experience varying returns to skills. Entry-age employees aged 25-34 are set as the base category of the analysis to avoid the dummy variable trap. The results of allowing the returns to skills to differ between the age categories indicate that there are significant age and therefore experience differences in returns to problem-solving in technology rich environments skills.

Across all the countries in the extended sample of full-time employees (Table 10) aged 25-65 entry-age workers' returns to problem-solving skills are 13 per cent and the returns to skills for prime-age employees are 4 percentage points higher than the entry-age workers. The aggregate result of the returns to problem-solving skills of approximately 17 per cent for prime-age workers is a bit higher than the results of the analysis conducted previously about returns to digital skills in the pooled sample of prime-age full-time employees (see Table 8). Returns to skills for exit-age employees compared to entry-age workers are statistically insignificant.

Table 10. The experience pattern of returns to skills using age categories as proxies

Regression type	Problem-solving	Returns for prime-age	Returns for exit-age	Prime-age	Exit-age	Female
With experience squared	0.128*** (0.008)	0.038*** (0.010)	-0.000 (0.015)	-0.057*** (0.015)	-0.061** (0.024)	-0.193*** (0.008)
Without experience squared	0.123*** (0.008)	0.045*** (0.010)	0.015 (0.014)	0.005 (0.014)	-0.088*** (0.024)	-0.192*** (0.008)

Source: Author's calculations in R based on data obtained from OECD
Notes:

1. No special R packages used for calculations.
2. Dependent variable is the logarithm of hourly wages adjusted to PPP in $US, problem-solving score standardized to standard deviation of 1 and experience squared divided by 1,000.
3. Estimations control for actual work experience variables, schooling, computer-use at work and native speaker indicator.
4. Standard errors in parenthesis
5. Significance Level values *, **, *** indicate significance at 10%, 5% and 1% level.
6. Prime-age 35-54, exit-age 55-65. Entry-age 25-34 set as base.
7. Both regression models cover 16,927 observations, multiple R-squared with experience squared is 0.225 and without 0.220, adjusted R-squared with experience squared is 0.224 and without 0.219. Both models are statistically significant on 1% level.

Examining the results of the sample countries individually in Table 11, returns to problem-solving skills are insignificant for workers aged between 25 to 34 in five out of thirteen countries from which there are no significant age interactions in Greece, Ireland and Lithuania. The percentage points by which the prime-age workers in three out of thirteen sample countries have higher returns to skills vary between 4 and 7. Six countries, where no significant age and therefore experience interactions are noted but the returns to problem-solving skills are significant, are Czech Republic, Denmark, Estonia, Poland, Slovak Republic, and Slovenia. This is for the most part consistent with Hanushek *et al.* (2015) findings that in the Eastern Europe transition economies prime-age and exit-age workers may not have the skills relevant for today's labour market unlike the entry-age employees.

Table 11. The experience pattern of returns to skills by country using age categories as proxies

Country	Problem-solving	Returns for prime-age	Returns for exit-age	Prime-age	Exit-age	Female
Belgium	0.014	0.042**	0.047	-0.003	-0.031	-0.074***
Czech Republic	0.052***	-0.027	-0.031	-0.061	-0.002	-0.162***
Denmark	0.023*	0.017	0.014	0.063***	0.037	-0.080***
Estonia	0.078***	0.010	0.016	-0.037	-0.057	-0.358***
Finland	-0.018	0.066***	0.048**	0.026	0.001	-0.181***
Greece	0.029	0.009	-0.123	0.141**	0.134	0.005
Ireland	0.025	0.012	0.051	0.063**	-0.022	-0.008
Lithuania	0.031	0.035	-0.015	-0.080	-0.031	-0.193***
Netherlands	0.032*	0.043*	0.047	0.131***	0.149***	-0.056***
Poland	0.093***	-0.028	-0.081	-0.061	0.112	-0.163***
Slovak Republic	0.058***	0.026	-0.026	-0.037	0.008	-0.276***
Slovenia	0.062***	-0.009	0.042	-0.029	0.195***	-0.110***
United Kingdom	0.055***	0.029	0.058**	0.103***	0.041	-0.082***

Source: Author's calculations in R based on data obtained from OECD
Notes:

1. No special R package used for calculations.
2. Dependent variable is the logarithm of hourly wages adjusted to PPP in $US, problem-solving score standardized to standard deviation of 1 by country and Experience squared divided by 1,000.
3. Estimations control for actual work experience variables, schooling, computer-use at work and native speaker indicator.
4. Standard errors in parenthesis.
5. Significance Level values *, **, *** indicate significance at 10%, 5% and 1% level.
6. Prime-age 35-54, exit-age 55-65. Entry-age 25-34 set as base.

The experience pattern of exit-age workers across the individual countries is consistent. The returns to problem-solving skills among exit-age workers compared to entry-age employees are significantly higher only in Finland where the 4.8 percentage point increase for exit-age workers is lower than the 6.6 percentage point increase for prime-age workers. Therefore, according to the conducted analysis, prime-age employees gain more from problem-solving skills than entry-age workers in the pooled sample regression, however this relationship is not seen across individual country regression results.

Looking at the differences between Tables 9 and 11, in all sample countries where the gender wage gap is statistically significant, the wage gap decreases with using experience patterns. In Netherlands, the gender wage gap not statistically significant in Table 9 turns into significant wage gap of 5.6 per cent seen in Table 11.

3.5. Digital skills impact on gender wage gap

Gender is a well-known factor which influences the determination of an individual's wage and thus the third hypothesis of this thesis endevours that the difference in problem-solving skills can explain some of the inequality of wages between genders. Gender interaction terms with problem-solving skills are added to the baseline model to trace the gender returns to problem-solving skills in Table 12.

Table 12. Returns to problem-solving skills by gender in the pooled sample with and without experience squared

Regression type	Problem-solving	Returns for females	Experience	Experience squared	Female
With experience squared	0.138*** (0.009)	0.021* (0.012)	0.009** (0.004)	0.113 (0.098)	-0.242*** (0.011)
Without experience squared	0.138*** (0.009)	0.022* (0.012)	0.014*** (0.001)	–	-0.242*** (0.011)

Source: Author's calculations in R based on data obtained from OECD

Notes:

1. No special R packages used for calculations.
2. Estimations control for actual work experience variables, schooling, computer-use at work and native speaker indicator.
3. Standard errors in parenthesis.
4. Significance Level values *, **, *** indicate significance at 10%, 5% and 1% level.
5. Both regression models cover 9,282 observations, multiple R-squared is 0.224, adjusted R-squared is 0.224 and are statistically significant on 1% level.

Table 12 displays the returns to the problem-solving skills by gender in the pooled sample of prime-age full-time employees. These results imply that the returns to problem-solving skills in technology rich environments, in the pooled sample, differ significantly between men and women as women have 2.1 per cent higher returns to problem-solving skills than men.

Across all the sample countries individually, the skill returns for females are small and statistically insignificant (see Appendix 8). Table 13 therefore shows the gender wage gap without conditioning on problem-solving skills by country since the returns to skills for men and women for individual countries are similar, and therefore it is possible to determine the part of problem-solving skills in explaining the gender wage gap by contrasting the regression models with and without the independent variable representing problem-solving skills.

Table 13. Regression analysis by country without problem-solving skills

Country	Experience	Experience squared	Female	Schooling	Computer-use at work	Native speaker
Belgium	0.013	-0.051	-0.090***	0.049***	0.051***	-0.009
Czech Republic	0.006	-0.160	-0.202***	0.056***	0.061***	-0.023
Denmark	0.005	-0.072	-0.110***	0.034***	0.071***	0.078***
Estonia	0.007	-0.207	-0.443***	0.058***	0.118***	0.063
Finland	0.010*	-0.087	-0.207***	0.054***	0.084***	0.071
Greece	0.020	0.043	0.013	0.088***	0.015	-0.151
Ireland	-0.001	0.289	-0.050*	0.055***	0.029**	0.054
Lithuania	0.010	-0.269	-0.247***	0.072***	0.098***	0.035
Netherlands	-0.001	0.136	-0.048*	0.063***	0.067***	0.078**
Poland	0.039**	-0.714	-0.214***	0.099***	0.058**	-0.365**
Slovak Republic	0.014	-0.296	-0.316***	0.069***	0.073***	0.021
Slovenia	0.019**	-0.302	-0.130***	0.132***	0.046***	0.084
United Kingdom	-0.006	0.242	-0.134***	0.063***	0.102***	-0.028

Source: Author's calculations in R based on data obtained from OECD
Notes:

1. No special R package used for calculations.
2. Dependent variable is the logarithm of hourly wages adjusted to PPP in $US, problem-solving score standardized to standard deviation of 1 by country and experience squared divided by 1,000.
3. Significance Level values *, **, *** indicate significance at 10%, 5% and 1% level.

Comparing the gender wage gaps of individual countries between Tables 9 and 13, gender wage gap is not affected by conditioning on skills in Czech Republic. The wage gap is insignificant in both cases in Greece while gender wage gap of five per cent disappears when conditioning on problem-solving skills in Ireland and Netherlands. Looking at the remaining sample countries varying parts of the gender wage gap can be attributed to the gender differences in problem-solving skills in technology rich environments.

CONCLUSION

Basic digital skills are imperative for thriving in today's technology-rich world as more and more everyday tasks require interacting with ever-changing technology. Individuals are motivated to invest in human capital through schooling, training or skill development since diversity in the human capital is an important factor in explaining the differences in wages.

The current thesis aimed to explain to what extent digital skills might raise wages in selected countries within the EU. To fulfil the aim of the thesis, the author worked through the relevant field-specific literature and based on the acquired information used regression analyses to examine the validity of proposed hypotheses. The sample of the paper consisted of thirteen European countries represented in OECD's PIAAC data and for estimating the digital skills data regarding problem-solving skills in technology rich environments was used.

The first hypothesis that digital skills increase individuals' wage in the selected sample of prime-age full-time employees was confirmed. An empirical analysis of log-linear model where the dependent variable was gross hourly wage and independent variables included problem-solving skills, actual working experience, actual working experience squared, gender, schooling, computer-use at work, and native speaker indicator results showed that higher problem-solving skills are systematically linked with higher wages in eleven out of thirteen sample countries.

The main research question how much does an employee gain from digital skills was also answered. In the pooled sample of full-time employees aged 35 to 54 one standard deviation increase in problem-solving skills was associated with an average increase in gross hourly wage by 14.9 per cent. Across countries, the highest returns to problem-solving skills were in Estonia, Netherlands, Slovak Republic and United Kingdom while in both Czech Republic and Greece the impact of problem-solving skills on wages was not statistically significant.

The analysis of the second hypothesis that employee's work experience is in positive correlation with returns to digital skills yielded mixed results. The actual working experience was proxied in the regression model with age categories after checking that there is a strong positive correlation between age and work experience. The returns to skills were allowed to differ between three age categories with the help of two dummy variables with the base category being entry-age employees aged 25-34.

Across all the countries the returns to skills for prime-age employees were 4 percentage points higher than for the entry-age workers. Examining the results of the sample countries individually, returns to problem-solving skills were significantly higher for workers aged between 35 to 54 than for entry-age workers in three out of thirteen countries while the returns among exit-age workers compared to entry-age employees were significantly higher in two sample countries. Therefore, the second hypothesis of the thesis that experienced employees gain more from digital skills than inexperienced ones was partially confirmed since prime-age employees have higher returns to problem-solving skills than entry-age workers while the pattern of the returns to problem-solving skills between entry-age and exit-age workers is undetermined.

To estimate the validity of the third hypothesis that the difference in digital skills can explain some of the inequality of wages between genders two regression analyses were conducted. The results of the first analysis indicated that in the pooled sample the returns to problem-solving skills in technology rich environments differ significantly between men and women with women having 2.1 per cent higher returns to the problem-solving skills than men. The results for individual countries, however, were the opposite of the pooled sample which meant that it was possible to determine the part of problem-solving skills in explaining the gender wage gap by country by contrasting the regression models with and without the independent variable representing problem-solving skills.

Comparing the models by country with and without the independent variable representing problem-solving skills it was concluded that in most sample countries some of the gender wage gap was traceable to the difference of problem-solving skills between the genders. Czech Republic was the only country out of thirteen sample countries where a significant gender wage gap was unaffected by the conditioning on skills.

One of the charms of PIAAC is that it is still ongoing with the results of the Second Cycle of data collection coming in 2023. The addition of new countries and some new variables like adaptive problem-solving could prove to be a good match for the problem-solving skills limited for the First

Cycle. Therefore in addition to changing the sample design and taking a look at the returns to problem-solving skills for social subgroups it would be possible to evaluate the returns to adaptive problem-solving skills and compare them with returns to problem-solving skills.

KOKKUVÕTE

DIGITAALSETE OSKUSTE MÕJUST PALKADELE: EUROOPA RIIKIDE MIKROTASEME ANALÜÜS KASUTADES PIAAC ANDMEID

Miina Hõbenael

Tehnoloogial on kasvav mõju kogu ühiskonnas. Digitaalsed põhioskused on tänases tehnoloogirikkas maailmas edukaks toimetulekuks elulise tähtsusega, kuna üha enam igapäevaseid ülesandeid nõuavad pidevalt muutuva tehnoloogia kasutamist. Tehnoloogia kasvav mõju inimeste igapäevaelus võib tekitada küsimuse, kuidas mõjutavad sellega paratamatult kaasnevad muutused palku.

Üksikisiku palga kujunemine on seotud inimkapitaliga (Hanushek, Woessmann 2008). Tõlgendades indiviidi digitaalseid oskuseid inimkapitalina, on loogiline järeldada, et tehnoloogilised oskused on eeliseks konkurentsitihedal tööturul. Siiski on saadaval piiratud mahus kirjandust digitaalsete oskuste tähenduse ning selle kohta, kuidas nende kujunemist mõjutavad erinevad individuaalsed karakteristikud, nagu vanus, sugu ja kogemus.

Empiirilist lähenemist tööturuoskustele muutis märkimisväärselt Jacob Mincer (1970), kes asus inimkapitali mõõtma omandatud hariduse järgi. Hanushek *et al.* (2015) toovad aga Minceri töö nõrga kohana välja asjaolu, et selline lähenemine eeldab inimestevaheliste oskuste erinevuste kujunemist ainult omandatud hariduse erinevuste baasil, mistõttu kasutavad nad oma artiklis teistsugust strateegiat, mis sarnaneb Minceri mudelile, kuid omandatud hariduse asemel mõõdab inimkapitali oskuste mõõdik.

Käesoleva lõputöö eesmärgiks on välja selgitada, millisel määral mõjutavad inimeste digitaalsed oskused nende palku valitud Euroopa Liidu riikides. Peamine uurimisküsimus, millele vastust otsitakse, on, kui palju saab töötaja kasu digitaalsetest oskustest. Töös on püstitatud kolm hüpoteesi:

1. Valitud Euroopa Liidu riikides on digitaalsetel oskustel positiivne mõju inimeste palkadele.

2. Töötajate töökogemuse ja digitaalsetest oskustest saadava kasu vahel valitseb positiivne korrelatsioon.
3. Osa soolisest palgalõhest on seletatav meeste ja naiste vahelise digitaalsete oskuste erinevusega.

Töö eesmärgi täitmiseks ning hüpoteeside kontrollimiseks viiakse Hanushek *et al.* (2015) eeskujul läbi log-lineaarsed regressioonanalüüsid, kus sõltuvaks muutujaks on isiku ostujõu pariteediga korrigeeritud brutotunnipalk USA dollarites ning sõltumatuteks muutujateks probleemilahendusoskus tehnoloogiarikastes keskkondades, tegelik töökogemus ja selle ruut ning sooline tunnus. Analüüsi käigus lisanduvad veel kolm kontrollmuutujat: haridustase aastates, arvuti kasutamine tööl, riigikeelt emakeelena kõnelemise indikaator. Teise ja kolmanda hüpoteesi empiirilisse analüüsi on lisaks eelnevatele kontrollmuuutjatele lisatud ka vastavalt kas vanuselist või soolist kuuluvust kirjeldavad näivad tunnused. Probleemilahendusoskus tehnoloogiarikkas keskkonnas on antud töös digitaalseid oskusi kirjeldav näitaja, mille andmed pärinevad koos kõigi teiste muutujatega andmetega Majanduskoostöö ja Arengu Organisatsiooni (OECD) avalikult tasuta kättesaadavast rahvusvahelisest täiskasvanute oskuste uuringust PIAAC.

Töös kasutatud valim hõlmab kolmeteist PIAACi esimese tsükli andmetekogumise ajal Euroopa Liitu kuulunud riiki: Belgia, Tšehhi, Taani, Eesti, Soome, Kreeka, Iirimaa, Leedu, Holland, Poola, Slovakkia, Sloveenia, Suurbritannia. Põhivalim, mida kasutatakse digitaalsete oskuste palgamõjude ning nende soolise palgalõhe seletamise võime hindamiseks, koosneb täistööajaga töötavatest palgatöötajatest vanuses 35 kuni 54 aastat. Piirangute eesmärgiks on eraldada oskuste mõju teistest aktiivsesse tööjõudu kuulumist mõjutavatest asjaoludest nagu tervislik seisund ja perekondlikud nõudmised. Töötajate töökogemuse ja digitaalsetest oskustest saadava kasu vahel valitseva seose määramiseks kasutatakse põhivalimist erinevat 25-65 aastaste täistööajaga töötavate palgatööliste valimit.

Esimene töös püstitatud hüpotees, et valitud Euroopa Liidu riikides on digitaalsetel oskustel positiivne mõju inimeste palkadele, leidis kinnitust. Regressioonanalüüsi tulemused näitasid, et paremad probleemilahendusoskused on süstemaatiliselt seotud kõrgemate palkadega üheteistkümnes kolmeteistkümnest valimisse kuulunud riigist. Riikideüleses koondatud põhivalimis on probleemilahendusoskuste suurenemine ühe standardhälbe võrra seotud 14.9 protsendilise brutotunnipalga tõusuga. Riikide lõikes oli probleemilahendusoskustel palgale kõige suurem positiivne mõju Eestis, Taanis, Slovakkias ja Suurbritannias, kus ühele standardhälbe võrra suurenemisele vastav brutotunnipalga tõus varieerus 7 ja 8 protsendi vahel. Probleemilahendusoskused polnud kõrgemate palkadega statistiliselt seotud Tšehhis ja Kreekas.

Teine töö autori poolt püstitatud hüpotees töötajate töökogemuse ja digitaalsetest oskustest saadava kasu vahelise positiivse korrelatsiooni kohta leidis osaliselt kinnitust. Ühendatud riikideüleses valimis oli 35-54 aastaste töötajate grupi kasu 4 protsenti kõrgem võrreldes baaskategooriaks olnud 25-34 aastaste töötajate kasuga. Vanema kategooria ehk 55 kuni 65 aastaste töötajate kasu võrreldes 25-34 aastaste omaga pole antud valimi põhjal võimalik hinnata. Seega saab väita vaid seda, 25-54 aastaste töötajate poolt saadav kasu digitaalsetest oskustest on positiivses korrelatsioonis nende töökogemusega.

Kolmanda hüpoteesi paikapidavuse kontrollimiseks viidi läbi kaks regressioonanalüüsi. Esimese analüüsi tulemused näitasid, et paramatest probleemilahendusoskustest saadav kasu erineb meeste ja naiste lõikes. Riikide eraldi regressioonanalüüsides aga paramatest probleemilahendusoskustest saadav kasu meeste ja naiste lõikes ei erinenud, seega oli oskuste osa palgaerinevuste selgitamisel võimalik riikide lõikes hinnata võrreldes omavahel kahe regressioonmudeli tulemusi, millest ühes oli oskuste mõju modelleeriv sõltumatu tunnus välja jäetud.

Riikide lõikes palgalõhede analüüsimine andis tulemuseks, et suuremas osas valimisse kuulunud riikidest oli võimalik osa palgalõhest selgitada digitaalsete oskuste erinevusega meeste ja naiste vahel. Kolmeteistkümnest riigist oli Tšehhi ainus, kus statistiliselt oluline sooline palgalõhe jäi mudelist oskuste eemaldamisel samasuguseks.

PIAAC andmete kogumise jätkumine ning uute muutujate ja riikide lisamine teises tsüklis, mille tulemused avaldatakse 2023. aastal, annavad võimaluse võrrelda oskuste mõju palkadele erinevate perioodide või muutujate lõikes. Samuti on võimalik käesolevas lõputöös läbiviidud analüüsi laiendada, võrreldes probleemilahendusoskuste põhjal leitud tulemusi adaptiivsete probleemilahendusoskuste põhjal saadud tulemustega. Adaptiivne probleemilahendusoskus on üks teises tsüklis lisanduvatest muutujatest.

LIST OF REFERENCES

Allison, P. (2012, Sept 10). When Can You Safely Ignore Multicollinearity? [Blog post]. Retrieved from https://statisticalhorizons.com/multicollinearity, 3 May 2020.

Anghel, B., Balart, P. (2017). Non-cognitive Skills and Individual Earnings: New Evidence from PIAAC. – *Journal of the Spanish Economic Association*, Vol. 8, No. 4, 417-473.

Autor, D. H., Levy, F., Murnane, R.J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. – *The Quarterly Journal of Economics*, Vol. 118, No. 4, 1279-1333.

Aviram, A., Eshet-Alkalai, Y. (2006). Towards a Theory of Digital Literacy: Three Scenarios for the Next Steps. – *European Journal of Open, Distance and E-Learning*, Vol. 9, No. 1.

Barone, C., Van de Werfhorst, H. G. (2011). Education, Cognitive Skills and Earnings in Comparative Perspective. – *International Sociology*, Vol. 26, No. 4, 483-502.

Barro, R. J. (2001). Human Capital and Growth. – *American Economic Review,* Vol. 91, No. 2, 12-17.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. – *Journal of Political Economy*, Vol. 70, No. 5, 9-49.

Borghans, L., Duckworth, A. L., Heckman, J. J., Ter Weel, B. (2008). The Economics and Psychology of Personality Traits. – *Journal of Human Resources*, Vol. 43, No. 4, 972-1059.

Caro, D. H., Biecek, P. (2017). intsvy: An R Package for Analysing International Large-Scale Assessment Data. – *Journal of Statistical Software*, Vol. 81, No. 7, 1-44.

Certo, S.T., Busenbark, J.R., Woo, H.S., Semadeni, M. (2016). Sample Selection Bias and Heckman Models in Strategic Management Research. – *Strategic Management Journal*, Vol. 37, No. 13, 2639-2657.

Cunha, F., Heckman, J. (2007). The Technology of Skill Formation. – *American Economic Review*, Vol. 97, No. 2, 31-47.

European Commission. (2018*). Study of the Movement of Skilled Labour*. Luxembourg: Publications Office of the European Union.

Ferraro, S. (2018). Is Information and Communication Technology Satisfying Educational Needs at School? – *Computers & Education*, Vol. 122, 194-204.

Goldfarb, A., Tucker, C. (2019). Digital Economics. – *Journal of Economic Literature*, Vol. 57, No.1, 3-43.

Green, D. A., Riddell, W. C. (2003). Literacy and Earnings: An Investigation of the Interaction of Cognitive and Unobserved Skills in Earnings Generation. – *Labour Economics*, Vol. 10, No. 2, 165-184.

Green, D. A., Riddell, W. C. (2013). Ageing and Literacy Skills: Evidence from Canada, Norway and the United States. – *Labour Economics*, Vol. 22, 16-29.

Hanushek, E. A., Kimko, D. D. (2000). Schooling, Labor-Force Quality, and the Growth of Nations. – *American Economic Review*, Vol. 90, No. 5, 1184-1208.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., Woessmann, L. (2015). Returns to Skills Around the World: Evidence from PIAAC. – *European Economic Review*, Vol. 73, 103-130.

Hanushek, E. A., Woessmann, L. (2008). The Role of Cognitive Skills in Economic Development. – *Journal of Economic Literature*, Vol. 46, No. 3, 607-668.

Harmon, C., Oosterbeek, H., Walker, I. (2003). The Returns to Education: Microeconomics. – *Journal of Economic Surveys*, Vol. 17, No. 2, 115-156.

Hatlevik, O. E., Guðmundsdóttir, G. B., Loi, M. (2015). Digital Diversity Among Upper Secondary Students: A Multilevel Analysis of the Relationship Between Cultural Capital, Self-Efficacy, Strategic Use of Information and Digital Competence. – *Computers & Education*, Vol. 81, 345-353.

Heckman, J., Polachek, S. (1974). Empirical Evidence on the Functional Form of the Earnings-Schooling Relationship. – *Journal of the American Statistical Association*, Vol. 69, No. 346, 350-354.

Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., Borghans, L. (2014). Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success. *NBER Working Papers*, No. 20749. Cambridge: National Bureau of Economic Research.

Mansfield, E.R., Helms, B.P. (1982). Detecting Multicollinearity. – *The American Statistician*, Vol. 36, No. 3a, 158-160.

Martin, J. (2018). Skills for the 21st Century: Findings and Policy Lessons from the OECD Survey of Adult Skills. *OECD Education Working Papers*, No. 166. Paris: OECD Publishing.

McIntosh, S., Vignoles, A. (2001). Measuring and Assessing the Impact of Basic Skills on Labour Market Outcomes. – *Oxford Economic Papers*, Vol. 53, No. 3, 453-481.

Mincer, J. (1970). The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach. – *Journal of Economic Literature*, Vol. 8, No. 1, 1-26.

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

Mohammadyari, S., Singh, H. (2015). Understanding the Effect of E-learning on Individual Performance: The Role of Digital Literacy. – *Computers & Education*, Vol. 82, 11-25.

OECD. (2013). *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*. Paris: OECD.

OECD. (2016). *Technical Report of the Survey of Adult Skills (PIAAC), 2nd edition.* Paris: OECD.

Schleicher, A. (2020, Apr 1). Education Disrupted – Education Rebuilt: Some Insights from PISA on the Availability and Use of Digital Tools for Learning. [Blog post]. Retrieved from https://oecdedutoday.com/coronavirus-education-digital-tools-for-learning/, 10 May 2020.

Schultz, T. W. (1961). Investment in Human Capital. – *American Economic Review*, Vol. 51, No. 1, 1-17.

Welland, J.D. (1978). Cognitive Abilities, Schooling and Earnings: The Question of Functional Form. – *The Review of Economics and Statistics*, Vol. 60, No. 4, 622-627.

APPENDICES

Appendix 1. Names, labels, value labels and types of PIAAC variables to further define the thesis samples

Name in PIAAC	Label in PIAAC	Value Labels	Value Type
AGEG10LFS	Age in 10 year bands (derived)	24 or less	Valid
		25-34	Valid
		35-44	Valid
		45-54	Valid
		55 plus	Valid
		<16	Missing
		>65	Missing
		Not stated or inferred	Missing
PSLSTATUS	Problem-solving - PV Status	Has PV	Valid
		Literacy Related Non-Response	Valid
		CBA Non-Response	Valid
		Valid skip	Missing
		Don't know	Missing
		Refused	Missing
		Not stated or inferred	Missing
C_Q07	Current status/work history - Subjective status	Full-time employed (self-employed, employee)	Valid
		Part-time employed (self-employed, employee)	Valid
		Unemployed	Valid
		Pupil, student	Valid
		Apprentice, internship	Valid
		In retirement or early retirement	Valid
		Permanently disabled	Valid

Name in PIAAC	Label in PIAAC	Value Labels	Value Type
C_Q07	Current status/work history - Subjective status	In compulsory military or community service	Valid
		Fulfilling domestic tasks or looking after children/family	Valid
		Other	Valid
		Valid skip	Missing
		Don’t know	Missing
		Refused	Missing
		Not stated or inferred	Missing
D_Q04	Current work - Employee or self-employed	Employee	Valid
		Self-employed	Valid
		Valid skip	Missing
		Don’t know	Missing
		Refused	Missing
		Not stated or inferred	Missing

Notes: Variable names and labels and value labels and types in PIAAC are accessible from the International Codebook for PIAAC Public Use Files, which can be found on the OECD Skills Surveys webpage under documentation section of Data and Tools.

Appendix 2. Descriptive statistics per country in the sample of full-time prime-age employees

Country	Observations	Gross hourly wage in PPP, (US$)	PIAAC PSL Score	Experience, (years)	Female, (share)
Belgium	894	21.51 (6.89)	281.80 (37.22)	22.9 (6.5)	0.35
Czech Republic	745	9.70 (4.71)	276.64 (41.09)	21.3 (6.9)	0.49
Denmark	1,282	25.81 (6.88)	287.54 (36.23)	23.4 (7.2)	0.48
Estonia	1,199	10.49 (6.46)	269.50 (37.73)	21.5 (6.6)	0.60
Finland	1,206	21.13 (7.06)	292.29 (34.57)	21.1 (6.8)	0.48
Greece	429	12.25 (6.65)	263.72 (46.58)	19.0 (6.4)	0.46
Ireland	696	25.36 (9.05)	281.24 (36.62)	21.3 (6.9)	0.47
Lithuania	976	7.91 (4.49)	255.98 (40.03)	22.0 (6.7)	0.63
Netherlands	727	23.86 (7.93)	291.40 (35.05)	23.6 (6.6)	0.20
Poland	408	11.03 (6.05)	266.51 (43.03)	20.2 (6.8)	0.53
Slovak Republic	745	8.67 (4.89)	277.48 (32.57)	21.7 (6.4)	0.54
Slovenia	1,033	9.77 (4.32)	263.39 (45.88)	22.2 (7.4)	0.53
United Kingdom	1,435	20.11 (8.11)	282.19 (38.27)	23.7 (6.8)	0.49

Source: Author's calculations based on data obtained from OECD

Notes:

1. "PIAAC PSL score" stands for PIAAC score in problem-solving in technology rich environments.
2. Standard deviations for gross hourly wage in PPP, PIAAC PSL score and experience per country are in the brackets.
3. No special functions used for calculating the descriptive statistics.

Appendix 3. The mean and the standard error of gross hourly wages in PPP US$ in the pooled sample of prime-age full-time employees and across countries

Country	Frequency	Mean in PPP (US$)	Standard Error
Belgium	894	21.59	0.23
Czech Republic	745	9.25	0.21
Denmark	1,282	25.56	0.20
Estonia	1,199	10.60	0.17
Finland	1,206	20.83	0.17
Greece	429	12.18	0.36
Ireland	696	24.90	0.41
Lithuania	976	7.98	0.21
Netherlands	727	23.45	0.27
Poland	408	11.12	0.27
Slovak Republic	745	9.12	0.23
Slovenia	1,033	9.62	0.14
United Kingdom	1,435	20.31	0.30
Pooled	11,775	17.42	0.15

Source: Author’s calculations using R package “intsvy” based on the data from OECD PIAAC

Appendix 4. The mean and the standard error of problem-solving skills in technology rich environments in the pooled sample of prime-age full-time employees and across countries

Country	Frequency	Mean	Standard Error	Standard Deviation	Standard Error
Belgium	894	282.03	1.61	40.66	1.15
Czech Republic	745	272.45	2.62	44.16	1.88
Denmark	1,282	287.33	1.43	38.61	1.09
Estonia	1,199	270.13	1.44	40.81	0.94
Finland	1,206	290.71	1.27	38.20	0.94
Greece	429	267.55	3.92	51.31	2.69
Ireland	696	278.48	1.99	40.58	1.58
Lithuania	976	255.07	2.09	44.24	1.71
Netherlands	727	290.41	1.81	39.00	1.36
Poland	408	268.37	2.83	46.65	2.04
Slovak Republic	745	279.97	1.76	37.18	1.37
Slovenia	1,033	264.30	1.84	49.23	1.25
United Kingdom	1,435	281.59	1.75	43.24	1.62
Pooled	11,775	278.67	0.81	43.95	0.78

Source: Author’s calculations using R package “intsvy” based on the data from OECD PIAAC

Appendix 5. Correlation matrix of the variables used in the extended regression analyses

	lny_i	C_i	E_i	E_i^2	G_i	S_i	CW_i	N_i
lny_i	1.000	0.320	0.058	0.056	-0.242	0.270	0.233	-0.066
C_i	0.320	1.000	-0.229	-0.232	-0.139	0.286	0.322	0.055
E_i	0.058	-0.229	1.000	0.984	-0.062	-0.233	-0.057	0.052
E_i^2	0.056	-0.232	0.984	1.000	-0.058	-0.235	-0.061	0.045
G_i	-0.242	-0.139	-0.062	-0.058	1.000	0.055	-0.055	0.021
S_i	0.270	0.286	-0.233	-0.235	0.055	1.000	0.226	-0.043
CW_i	0.233	0.322	-0.057	-0.061	-0.055	0.226	1.000	-0.002
N_i	-0.066	0.055	0.052	0.045	0.021	-0.043	-0.002	1.000

Source: Author's calculations in R based on data obtained from OECD

Notes:

1. Meanings of abbreviations in Table 1 and 2.
2. "E_i^2" is the squared value of experience in years divided by 1,000.
3. "CW_i" illustrates individual's use of ICT skills at work on a scale from 0 to 10.
4. "N_i" is the native speaker indicator with value 1 if the individual is a native speaker.
5. Observations 9,282.

Appendix 6. Regression results and steps of adding the three control variables

Variable	No Added Controls	N_i Added	CW_i Added	S_i Added	CW_i and S_i Added	N_i, CW_i and S_i Added
Problem-solving	0.253*** (0.005)	0.190*** (0.006)	0.160*** (0.006)	0.149*** (0.006)	0.144*** (0.007)	0.149*** (0.007)
Experience	0.009** (0.004)	0.008* (0.005)	0.005 (0.005)	0.008* (0.004)	0.007 (0.004)	0.008* (0.004)
Experience squared	0.038 (0.092)	0.050 (0.102)	0.104 (0.101)	0.133 (0.099)	0.147 (0.098)	0.119 (0.098)
Female	-0.216*** (0.010)	-0.217*** (0.011)	-0.219*** (0.011)	-0.243*** (0.011)	-0.241*** (0.011)	-0.237*** (0.011)
Native speaker	–	-0.211*** (0.023)	–	–	–	-0.177*** (0.022)
Computer-use at work	–	–	0.085*** (0.006)	–	0.064*** (0.006)	0.063*** (0.006)
Schooling	–	–	–	0.056*** (0.002)	0.052*** (0.002)	0.051*** (0.002)
R^2	0.205	0.163	0.174	0.209	0.219	0.224
Adjusted R^2	0.204	0.163	0.173	0.209	0.218	0.223

Source: Author's calculations in R based on data obtained from OECD

Notes:

1. Meanings of abbreviations in Table 2.
2. Each column represents a different ordinary least squares regression.
3. Standard errors in parenthesis.
4. Observations in no added controls regression 11,775, all other regressions 9,282.

Appendix 7. Regression analysis of digital skills impact on wages by country with and without the variable experience squared

Country	C_i	E_i	E_i^2	G_i	S_i	CW_i	N_i	Observations
Belgium	0.054*** (0.013)	0.013 (0.010)	-0.025 (0.219)	-0.076*** (0.021)	0.044*** (0.005)	0.038*** (0.013)	-0.032 (0.049)	722
	0.054*** (0.013)	0.012*** (0.002)	–	-0.076*** (0.021)	0.044*** (0.005)	0.038*** (0.013)	-0.032 (0.048)	722
Czech Republic	0.002 (0.018)	0.006 (0.012)	-0.160 (0.278)	-0.202*** (0.030)	0.056*** (0.006)	0.061*** (0.017)	-0.024 (0.096)	546
	0.002 (0.018)	-0.001 (0.002)	–	-0.203*** (0.030)	0.056*** (0.006)	0.061*** (0.017)	-0.023 (0.096)	546
Denmark	0.037*** (0.008)	0.005 (0.005)	-0.050 (0.106)	-0.100*** (0.013)	0.031*** (0.003)	0.062*** (0.007)	0.059*** (0.019)	1,133
	0.037*** (0.008)	0.003*** (0.001)	–	-0.100*** (0.013)	0.031*** (0.003)	0.062*** (0.007)	0.060*** (0.019)	1,133
Estonia	0.069*** (0.018)	0.008 (0.013)	-0.181 (0.302)	-0.427*** (0.031)	0.052*** (0.007)	0.102*** (0.015)	0.061 (0.218)	880
	0.069*** (0.018)	0.000 (0.002)	–	-0.428*** (0.031)	0.052*** (0.007)	0.103*** (0.015)	0.061 (0.218)	880
Finland	0.037*** (0.009)	0.010* (0.006)	-0.051 (0.140)	-0.198*** (0.015)	0.051*** (0.003)	0.073*** (0.010)	0.052 (0.046)	1,088
	0.038*** (0.009)	0.008*** (0.001)	–	-0.197*** (0.015)	0.051*** (0.003)	0.073*** (0.010)	0.052 (0.046)	1,088
Greece	0.034 (0.025)	0.022 (0.018)	-0.005 (0.472)	0.018 (0.047)	0.085*** (0.011)	0.007 (0.027)	-0.158 (0.207)	314
	0.034 (0.025)	0.022*** (0.004)	–	0.018 (0.047)	0.085*** (0.011)	0.007 (0.027)	-0.158 (0.206)	314
Ireland	0.038** (0.016)	-0.001 (0.011)	0.303 (0.253)	-0.035 (0.028)	0.050*** (0.007)	0.020 (0.013)	0.046 (0.049)	609
	0.038** (0.016)	0.012*** (0.002)	–	-0.035 (0.028)	0.050*** (0.007)	0.019 (0.013)	0.041 (0.049)	609
Lithuania	0.055*** (0.021)	0.012 (0.017)	-0.276 (0.374)	-0.238*** (0.039)	0.067*** (0.008)	0.087*** (0.019)	0.038 (0.067)	576
	0.054*** (0.021)	-0.000 (0.003)	–	-0.241*** (0.038)	0.067*** (0.008)	0.086*** (0.019)	0.035 (0.067)	576
Netherlands	0.065*** (0.014)	-0.002 (0.010)	0.200 (0.216)	-0.033 (0.026)	0.054*** (0.006)	0.052*** (0.013)	0.049 (0.037)	633
	0.065*** (0.014)	0.008*** (0.002)	–	-0.034 (0.026)	0.055*** (0.006)	0.052*** (0.013)	0.051 (0.037)	633
Poland	0.059** (0.025)	0.041** (0.019)	-0.736* (0.440)	-0.186*** (0.047)	0.095*** (0.011)	0.047* (0.026)	-0.394*** (0.147)	305
	0.058** (0.025)	0.011*** (0.003)	–	-0.188*** (0.047)	0.095*** (0.011)	0.046* (0.027)	-0.369** (0.147)	305
Slovak Republic	0.069*** (0.018)	0.011 (0.015)	-0.190 (0.343)	-0.301*** (0.033)	0.064*** (0.007)	0.060*** (0.017)	0.009 (0.089)	514
	0.070*** (0.018)	0.003 (0.003)	–	-0.301*** (0.033)	0.064*** (0.007)	0.061*** (0.017)	0.012 (0.089)	514

Country	*Ci*	*Ei*	*Ei2*	*Gi*	*Si*	*CWi*	*Ni*	Obser-vations
Slovenia	0.049*** (0.015)	0.021** (0.009)	-0.324 (0.207)	-0.125*** (0.024)	0.123*** (0.009)	0.034** (0.015)	0.069 (0.052)	773
	0.048*** (0.015)	0.007*** (0.002)	–	-0.127*** (0.024)	0.123*** (0.009)	0.035** (0.015)	0.074 (0.052)	773
United Kingdom	0.082*** (0.012)	-0.008 (0.009)	0.322* (0.184)	-0.111*** (0.019)	0.055*** (0.005)	0.079*** (0.010)	-0.060 (0.040)	1,189
	0.081*** (0.012)	0.007*** (0.002)	–	-0.111*** (0.019)	0.055*** (0.005)	0.079*** (0.010)	-0.074* (0.039)	1,189

Source: Author's calculations using R based on the data from OECD PIAAC
Notes:

1. Meanings of abbreviations in Tables 1 and 2.
2. Statistical significance levels: *<0.10, **<0.05, ***<0.01.
3. Skills score standardized within each country.
4. Each row displays the results of a different regression analysis.
5. Standard errors in parenthesis.

Appendix 8. Returns to problem-solving skills by gender across countries

Country	Problem-solving	Returns for Females	Experience	Experience Squared	Female
Belgium	0.046*** (0.015)	0.021 (0.023)	0.014 (0.010)	-0.038 (0.220)	-0.080*** (0.022)
Czech Republic	0.012 (0.023)	-0.022 (0.032)	0.006 (0.012)	-0.169 (0.279)	-0.197*** (0.031)
Denmark	0.026** (0.010)	0.023 (0.014)	0.005 (0.005)	-0.062 (0.107)	-0.103*** (0.013)
Estonia	0.056** (0.026)	0.021 (0.032)	0.007 (0.013)	-0.172 (0.302)	-0.433*** (0.032)
Finland	0.023* (0.013)	0.026 (0.016)	0.011* (0.006)	-0.074 (0.140)	-0.201*** (0.015)
Greece	0.030 (0.033)	0.010 (0.048)	0.023 (0.019)	-0.011 (0.473)	0.017 (0.048)
Ireland	0.034 (0.021)	0.009 (0.028)	-0.001 (0.011)	0.307 (0.254)	-0.036 (0.028)
Lithuania	0.074** (0.033)	-0.029 (0.039)	0.011 (0.017)	-0.257 (0.375)	-0.226*** (0.042)
Netherlands	0.065*** (0.015)	0.004 (0.029)	-0.002 (0.010)	0.200 (0.216)	-0.033 (0.026)
Poland	0.058 (0.037)	0.001 (0.048)	0.041** (0.019)	-0.736* (0.441)	-0.186*** (0.048)
Slovak Republic	0.068*** (0.026)	0.003 (0.034)	0.011 (0.015)	-0.188 (0.344)	-0.301*** (0.034)
Slovenia	0.060*** (0.021)	-0.020 (0.026)	0.021** (0.009)	-0.311 (0.207)	-0.120*** (0.025)
United Kingdom	0.091*** (0.016)	-0.017 (0.021)	-0.008 (0.009)	0.324* (0.184)	-0.107*** (0.020)

Source: Author's calculations in R based on data obtained from OECD

Notes:

1. No special R packages used for calculations.
2. Dependent variable is the logarithm of hourly wages adjusted to PPP in $US, problem-solving score standardized by country to standard deviation of 1 and experience squared divided by 1,000.
3. Estimations control for actual work experience variables, schooling, computer-use at work and native speaker indicator.
4. Standard errors in parenthesis.
5. Significance Level values *, **, *** indicate significance at 10%, 5% and 1% level.

Appendix 9. Non-exclusive licence

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

I Miina Hõbenael

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

“Returns to Digital Skills on Wages: A Micro-Level Analysis across European Countries Using PIAAC Data”

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.*