Who is at risk of automation?

Estimating the effects of automation technologies on employment

Irem Naz Temizel
Delft University of Technology
Faculty Technology, Policy and Management

TU Delft

accenture

Who is at risk of automation? Estimating the effects of automation technologies on employment

Master thesis submitted to Delft University of Technology

in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in **Engineering and Policy Analysis**

Faculty of Technology, Policy and Management

by

Irem Naz Temizel

Student number: 4917065

To be defended in public on August 17th 2020

Graduation committee

Chairperson	: Dr. ir. B. Enserink, Section Policy Analysis
First Supervisor	: Dr. E. Schröder, Section Economics of Technology and Innovation
External Supervisor	: R. Postma, Accenture
External Supervisor	: Y. Sprockel, Accenture

Preface

This thesis is my final product as a student and indicates that my journey at the Delft University of Technology comes to an end. These two amazing years passed in the blink of an eye. I had so much fun, met with great people and learned a lot. In the middle of the pandemic crisis, I conducted the most important research of mine. The thesis process was much more exciting and instructive than I could ever imagine thanks to certain people. Therefore, I would like to express my gratitude to everyone who immensely supported me and made this process valuable.

My thesis committee was full of enthusiastic people. They were all willing to exchange knowledge and discuss with me the findings. Enno Schöder, my first supervisor, always encouraged me to find my own way through the thesis process but also he supported me in every critical point. He provided me with constructive feedback and I learned a lot from him. Also, he introduced me to Daniel Samaan, from the International Labor Organization, and let me expand my knowledge. Bert Enserink, my chair, provided me with his detailed feedback and comforting suggestions. It was always very encouraging to discuss any topic with him. Roberto Postma, my first advisor at Accenture, has a unique way of thinking and I learned a lot from him. It was super motivating to see his enthusiasm and discuss my findings. Yuri Sprockel, my second advisor at Accenture, always put a tremendous amount of time to think with me and improve my work. Talking to him both about my project and my concerns about life was always inspiring.

I feel very fortunate to choose this master program as I gained really nice friends whose support I always felt. Also, Accenture Talent Factory interns and organizers became good colleagues of mine and always provided me with fresh ideas.

My biggest supporters are my parents Hülya and Güven and my brother Can. It would not be possible without their support, love and vision. I adore each of them and learned a lot from them in life. Can always says "If you have a problem solve it", a secret for happiness. Finally, Alp, *ma chance de Paris*, inspired me every day to do things better and encouraged me with his ideas, love and cleaning skills. I am excited to bring the second TU Delft diploma at home. Cheers!

Irem Naz Temizel
The Hague, August 2020

Executive Summary

In recent years, concerns about the negative impacts of automation technologies (i.e. AI) on employment increased. Yet, how technologies will impact the labor force is still uncertain as it depends on the diffusion process of new technologies. Nevertheless, technological progress usually increases the total output generated and, the new output is redistributed among society. Eventually, net gain creates winners with more benefits and losers who are vulnerable to technological changes. Policies are powerful tools to help those who are expected to be at high risk of automation and hence, at risk of losing their jobs. Therefore, this study investigates who is at risk of automation and offers policy recommendations to reduce inequality and ensure the vulnerable groups are seen. The research question is formulated as follows:

Who is at risk of automation?

This study has four main outcomes. First, to estimate the share of employment at risk of automation across OECD countries. Second, to estimate the jobs and industries at risk of automation. Third, to define the socio-demographic characteristics of workers at risk of automation. Finally, to provide policy recommendations to mitigate the negative impacts of automation on the vulnerable groups.

Estimating the share of employment at risk of automation

The share of employment at risk of automation has been calculated before in different studies. While famous research of Frey and Osborne (2013) (FO) predicts that 47% of the US employment at significantly high risk of automation, Arntz et al. (2016) estimate this share as 9%. The biggest difference between the two approaches is that the latter study takes into account that the task structure and skills used at work may differ across individuals within the same occupation. Nedelkoska and Quintini (2018) (NQ) apply the method of FO by considering that tasks and skills vary for individuals within the same occupation.

This study is based on NQ's study and aims to improve their model. The employment share at the risk of automation is calculated with individual-level data across 33 countries (30 OECD and 3 non-OECD countries). Overall, about 14% of the total employment of 33 countries was found to be at significantly high risk of automation which is the same as NQ's estimation. As opposed to FO's 47% prediction, we calculated that 10% of the US employment is at significantly high risk. Therefore, this study suggests that automation risk is less threatening and more manageable when the risk is considered at the individual level.

Jobs and industries at risk

Automation risk is assigned to each individual in our dataset. Then, the individual risks are aggregated into occupations and industries. Overall, we found that automation risk declines as the frequency of social intelligence-related skills and the level of education increases. Planning for others, presentation and influencing are skills that have the largest negative impact on the risk of automation. While most of the elementary jobs (cleaning, food preparation, laborers etc.) and agriculture-related jobs are at the highest risk, managerial jobs and professionals (teachers, scientists etc.) are at the lowest risk. Low-risk jobs require high-level of education (i.e. college degree) and also include high-level of teaching, advising and influencing skills. On contrary, personal care workers which belong to the elementary job category has a lower risk compared to manufacturing jobs as it requires high social interaction even though the level of education is low. In terms of industries, most of the primary (fishing and agriculture) and secondary (manufacturing) industry jobs are at higher risk than the jobs in the service sector.

Socio-demographic characteristics of workers at risk

The vulnerable groups who are expected to be affected by the technological advancements the most across 33 countries are more likely to be less educated, low-income earners. In general, younger and older populations are at the highest risk of automation. While younger people mostly perform unskilled jobs due to having only basic education, older people tend to be more technologically outdated. Females are always at higher risk compared to their male colleagues in an occupation. However, automation risk should not be interpreted independently from the impact factor ($risk \times size$). So, even though females are always at higher risk, the impact of risk also depends on the male/female ratio in an occupation. Besides, workers with no contract or a temporary contract have a higher risk of losing their jobs compared to ones with an indefinite contract.

Policy recommendations

We discussed policies about how to reduce the adverse effects of technological advancements on the high-risk groups. While human capital incentives such as training focus on the adaptation of workers to technological changes, governments can also steer the technological progress with financial interventions such as introducing tax or subsidy policies. We highlighted the importance of training workers. However, even though people who are at the highest risk need training the most, they are less likely to receive it. Therefore, government-sponsored programs need to provide affordable training and encourage workers at risk to join. Also, we discussed the importance of reducing the monopoly power of tech giants to reduce inequality.

The main conclusion of our study is that automation technologies are unlikely to substitute the large numbers of workers. Yet, lower-skilled (lower-educated) workers are at higher risk compared to the high-skilled. Therefore, to cope with the possible negative effects of technological progress, the policy attention should be put on the most vulnerable groups through the means of training and various financial incentives.

Contents

1 Introduction **1**
1.1 Outline 3

2 Research definition **5**
2.1 Research gap 5
2.2 Research questions 6
2.3 Research method 7

3 Technology, employment and skills **9**
3.1 History of technological advancements, employment and skills . . 9
3.2 Technological changes 12
3.2.1 Skill-biased technological changes 12
3.2.2 Routine-biased technological changes 14
3.3 Effects of the current technological advancements on employment and skills 16
3.4 Approaches 18
3.4.1 Occupation-based approach: Understanding the study of Frey and Osborne 18
3.4.2 Task-based approach 23
3.5 Conclusion 25

4 Estimating the risk of automation **27**
4.1 Datasets 27
4.1.1 O*NET 27
4.1.2 PIAAC 29
4.2 Methodology 30
4.2.1 Data preparation 30
4.2.2 Logistic regression 34
4.2.3 Linear regression 40
4.3 Conceptual model 41

5 Who is at risk of automation? **43**
5.1 Skills used at work and automation risk 43
5.2 Socio-demographic characteristics and automation risk 45
5.3 Country-level risk 47
5.3.1 Distribution of automation risk 47
5.3.2 Employment shares at risk of automation 51
5.4 Occupations and industries at risk 52

5.4.1 Occupations . . . 52
5.4.2 Industries . . . 54
5.5 Distribution of employee characteristics . . . 57
5.5.1 Education level . . . 57
5.5.2 Gender . . . 58
5.5.3 Age . . . 60
5.5.4 Earnings . . . 62
5.5.5 Other characteristics . . . 64
5.6 Conclusion . . . 66

6 Conclusion **67**
6.1 Research summary . . . 67
6.2 Policy recommendations . . . 70
6.2.1 Human capital incentives . . . 71
6.2.2 Tax and subsidies . . . 72
6.2.3 Reducing the monopoly power . . . 73
6.3 Reflection on societal and ethical relevance . . . 74
6.4 Limitations and future work . . . 75

Bibliography **77**

List of Figures **85**

List of Tables **87**

A SOC-ISCO correspondence table for the labelled occupations **89**

B Occupations and predicted probabilities **93**

C Education level per country vs automation risk **107**

D Employment shares per risk levels by country **109**

Chapter 1

Introduction

Since the last two decades, the evidence of robots and machines being the substitute of human-beings have become more pronounced. A recent study by Borjas and Freeman (2019) found that an additional robot is the substitute of two to three workers overall and three to four workers for particular groups. With the current advancements in digital technologies such as AI, machine learning and other forms of smart automation, machines keep reproducing or outperforming human performance in various work activities including rational human-like decision-making (Agrawal et al., 2019; Thaler, 2015). According to the McKinsey (2017)'s report, 375 million people may need to switch to an entirely new occupation by 2030. Thus, similar questions arise in everyone's mind over future employment: What kind of jobs are at risk of automation? How to keep automation from taking our jobs? Do we have the right capabilities to keep up with the pace of technological advancements?

The effects of current technological advancements are not expected to be similar to those of previous waves of automation. Prior technologies increased the labor substitution by replacing machines in unproductive, arduous and routine tasks (Autor, 2015; Ernst et al., 2019). During the first two waves of automation, this replacement caused a shift of low/unskilled labor to less automatable sectors (Bessen, 2015; Handel, 2012). Then, with the increasing usage of computers in the 80s, many middle-skilled jobs disappeared (Jorgenson and Vu, 2016) which caused job polarization, the wage gap between the high-skilled and low-skilled workers, and consequently caused income inequality. However, people kept their vital importance for flexible, creative, cognitive and non-routine complex tasks (Autor, Levy, et al., 2003).

Nonetheless, new advancements in robotics and increasing implementation of automation technologies reduce the human necessity to accomplish some of these tasks and threaten the high-skilled workers (Autor, 2015; Frey and Osborne, 2013). Today, robots are progressively able to do cognitive tasks besides performing a wide variety of routine physical work activities better and cheaper compared to human-beings. Smart automation technologies (AI technologies, machine learning algorithms, big data techniques etc.) improve the decision-making process by delivering more reliable results in a shorter time (McKinsey, 2017).

Technological advancements can be complementary or substitutionary depending on the type of technology, tasks in an occupation and industry structure (Arntz et al., 2016; United Nations, 2003). Complementary technologies create new tasks such as maintenance and operational activities and increase the overall performance of the occupation without replacing workers. For instance, using new technologies for medical diagnosis reduce time to analyze the symptoms and gives time to doctors to put more attention to patients well-being (Ernst et al., 2019). On the other hand, substitutionary technologies allow companies to directly replace their labor due to the economic advantages of machines (Frey and Osborne, 2013). Workers may be displaced to new jobs or replaced entirely which lead to negative consequences for the society (Bessen, 2015). More specifically, these consequences would be higher technology unemployment, declining wages, higher inequality and job polarization (Freeman, 2015).

Yet, the overall output generated by the technological progress increases and is redistributed among society. Eventually, the net gain creates its own winners, who adapt to the changes and receive more benefits, and losers, a considerable proportion of workers at risk of automation, who are replaced by the new technology (Korinek, 2019; Nedelkoska and Quintini, 2018). So, the impact is not the same for every group in society. Leaving the vulnerable group behind without any policy attention can harm equality (Korinek, 2019).

One of the main research areas of economists has always been how new technologies influence employment. In the literature, various explanatory models have been developed to estimate the impacts of technology and translate the findings into future trends (Acemoglu and Restrepo, 2017; Autor, 2015; Autor, Levy, et al., 2003; Frey and Osborne, 2013). As each of the researches has a unique way to scope down and approach to the topic, results differ even when working with the same datasets.

One of the methodologies to assess the impacts of technology is to combine expert opinions with an econometric model (Arntz et al., 2016; Frey and Osborne, 2017; Nedelkoska and Quintini, 2018). The most well-known and influential work that uses this method is conducted by Frey and Osborne (2013). This study is based on predicting the jobs at risk of automation and concludes that 47% of employment in the United States are at high risk of automation in the next decade or two. However, results are criticized due to not considering differences in tasks, activities and required capabilities within an occupation (Arntz et al., 2016; Arntz et al., 2017; Nedelkoska and Quintini, 2018). Studies that consider each individual within the same occupation can perform different tasks predict remarkably lower results compared to those of Frey and Osborne.

Besides estimating the employment shares at risk of automation, these studies shortly discuss workers' characteristics such as wages and education levels and put only a little attention to policy recommendations based on the future predictions and characteristics of the workers at risk (Arntz et al., 2016). Nonetheless, Nedelkoska and Quintini (2018) provide broad coverage to the socio-demographic characteristics of workers, who currently hold the risky occupations, and highlight the necessity to translate the empirical results into policies. Because various groups of people are more vulnerable to technological changes, these people need to be adapted to the fast-changing environment. Arntz et al. (2016) specify that "*the likely challenge for the future lies in coping with rising inequality and ensuring sufficient (re-)training especially for low qualified workers.*"

With this study, we aim to put more emphasis on the characteristics of workers at risk and provide policy recommendations to mitigate the adverse effects of the automation technologies.

1.1 Outline

This research can be divided into three main sections. Also, the research structure is graphically depicted in Figure 1.1.

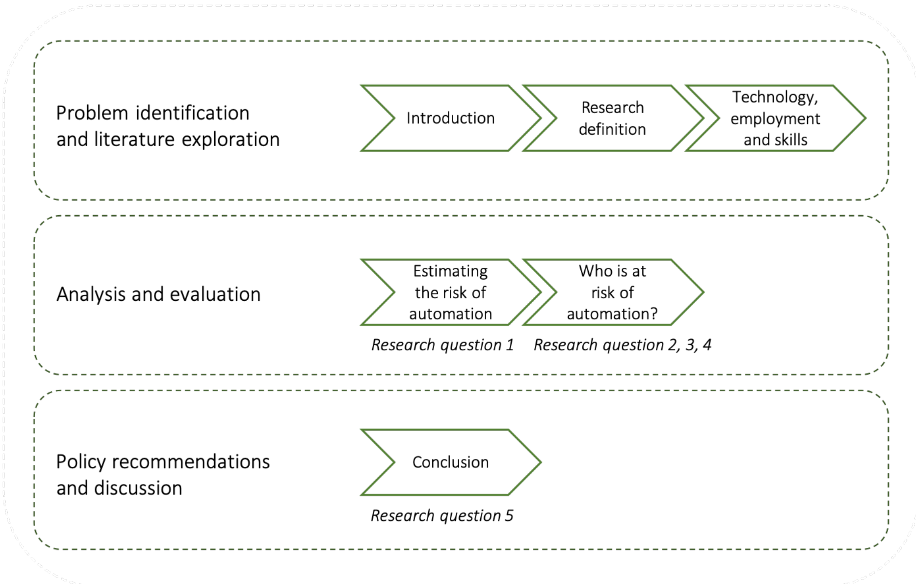

Figure 1.1: The research structure

1. **Problem identification and literature exploration:** This section is about identifying the problems and reviewing the literature. Initially, the societal and scientific importance of the research is introduced. This part continues with the definition of research: research gap, research questions and method. Chapter 3 explores the literature and discusses how automation technologies are expected to impact employment and what is different compared to the previous waves of automation.

2. **Analysis and evaluation:** In the second part of the study, we perform an analysis to predict the shares of employment, jobs and industries at risk for 33 countries (30 OECD and 3 non-OECD countries). Then, the characteristics of workers are investigated to prepare a baseline for policy recommendations. As the study of Nedelkoska and Quintini (2018)

contains a detailed analysis from estimating the risk of jobs to providing policy recommendations, this study is taken as the role model. In chapter 4, Estimating the risk of automation, we explain the data used and method applied. Chapter 5, Who is at risk of automation?, is the presentation of the results, also, reflects and discusses the findings derived from the analysis. The other aim is to highlight whether the results are similar to those of Nedelkoska and Quintini.

3. **Policy recommendations and discussion:** The final section includes chapter 6. The conclusions of the study are provided and finally, the analysis results and the literature review are combined to provide policy recommendations. Also, the recommendations are discussed with technology experts.

Chapter 2

Research definition

The starting point of this study is to explore the impacts of technology on employment and provide policy recommendations based on the findings. More specifically, we investigate the risk of automation on employment and the socio-demographic characteristics of employees at risk. In this section, first, the research gap and research questions are defined and the problem is scoped. Then, the applied method is presented.

2.1 Research gap

As mentioned earlier in the introduction, policy attention needs to be put on the vulnerable groups to prevent the rising inequality and high ratio of unemployment. To build robust policies, comprehensive knowledge is required about the people at risk (and their socio-demographic characteristics) and the share of these people in the total population (Arntz et al., 2016; Nedelkoska and Quintini, 2018). A variety of studies on predicting the share of employment at risk of automation is available in the literature. Results differ depending on the methodology, dataset and characteristics of the chosen country. However, these studies include a limited elaboration on the characteristics of workers at risk and policy recommendations for the relevant groups (Arntz et al., 2016; Nedelkoska and Quintini, 2018).

Therefore, we aim to conduct a study that first predicts the jobs at risk, then defines the workers at risk and offers policy recommendations to reduce inequality and ensure the vulnerable groups are seen.

The study of Nedelkoska and Quintini (2018) (NQ hereafter) is taken as the role model due to its comprehensive analysis where more than 30 countries are included in the analysis. NQ predict the employment shares at risk of automation based on the skills used at work, then focus on the workers' characteristics and policy recommendations. The method used to calculate the shares of employment at risk requires data to be first trained and then tested. Data is tested to understand if the model performance is high enough. NQ train the model with Canadian dataset due to its largest sample size. However, they neither discuss the performance of the model nor show whether the Canadian dataset fits the model the best, meaning that the model performance is the highest.[1] Therefore, compared to this study, we conduct an analysis with different country datasets and choose the one that fits the model the best. If a different dataset is found to be a better fit for the model, the results can change accordingly. Hence, the following findings and recommendations may vary. On the other hand, if the results do not change substantially, then this would lead us to similar conclusions.

2.2 Research questions

The main research question is formulated as follows:

Who is at risk of automation?

With this short but effective research question, we aim to define the jobs, industries and workers at risk by looking into skills used at work. The main reason why the analysis focuses on the skill use is that skills are the features that can be improved and influenced by the polices. In addition, automation refers to the applications of AI, machine learning and robotics.

The research question is divided into manageable sub-research questions. These sub-questions will eventually lead to the answers for the main research question.

1. **Which country's dataset is the best to predict the share of jobs at risk of automation by using the task-based approach?**
2. **What are the shares of employment at risk of automation in each OECD country?**
3. **Which jobs and industries are at high risk of automation in OECD countries?**
4. **What are the socio-demographic characteristics of workers at risk of losing their jobs?**
5. **What kind of policies could reduce the social cost of rapid technological progress?**

The main improvement to NQ's study is that this study chooses a representative country dataset that gives more accurate results. Therefore, the first step would be running the model with several different datasets and compare the performance of the model. The chosen dataset is used to predict the proportions of employment at risk for other OECD countries. The second question aims to highlight how different OECD countries are expected to be affected by automation. Then, the results are aggregated into jobs and industries across these countries to analyze the total impact of the automation technologies. The

[1] There are performance metrics to measure the model fit. The metrics are applied to the test data. So, to decide which country's dataset should be used to train the model, performance metrics are compared and dataset with the highest result becomes the best fit for the model. Detailed information is provided in chapter 4.

third question picks out the jobs and industries that are expected to have structural changes in the next decades. The fourth question identifies the workers holding these jobs and their socio-demographic characteristics. Finally, with the fifth question, the analysis results (job, industries and workers at risk) and the literature review are combined to offer policy recommendations to help workers at risk. As a result, a synthesis of the findings will indicate workers at risk and the possible policies to prevent the negative effects of technological progress on the high-risk groups.

2.3 Research method

This section provides an overview of the methods used in the research and defines possible outcomes for each sub-question. Throughout the thesis, the method used for each sub-questions is discussed in detail. The research has two main parts: analysis and policy recommendations. The first three sections are devised for the analysis while the last section is for policy recommendations. Figure 2.1 indicates the outcomes of the sub-research questions.

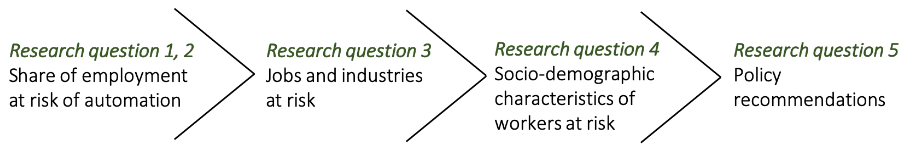

Figure 2.1: The outcomes of the sub-research questions

1. **The share of employment at risk of automation**

   Two approaches are available to predict the employment at risk of automation: occupation-based and task-based. While the occupation-based assumes that tasks of an occupation do not vary across the company, industry or country; the task-based considers the heterogeneity of workers' tasks and skills within an occupation (Arntz et al., 2016; Dengler and Matthes, 2018; Nedelkoska and Quintini, 2018). In this research, we choose to apply the task-based approach as it is more realistic to accept that tasks and skills vary across individuals within the same occupation.

   Both approaches combine expert opinions with a statistical model. The expert opinions used in this research are about how occupations will be affected by the automation technologies and what kind of skills are not easy to automate given the current state of knowledge. This information comes from Frey and Osborne's study (2013)(FO hereafter), one of the most influential researches about predicting the risk of automation on employment. NQ build their statistical model on the expert opinions provided in FO's study. We apply the same methodology as provided by NQ.

   As the task-based approach requires, NQ use individual-level data. The data comes from the Program for the International Assessment of Adult Competencies (PIAAC), a program that assesses and analyses adults' skills in over 40 countries (OECD, 2017). This program provides us with the individual-level data about the adult skills (Survey of Adult Skills). The survey measures the proficiency level of certain skills and also consists of skills used at work, at home and in general. In this study, we also use PIAAC dataset because our interest is to learn how adults use their skills at work and relate this information with automation risk.

NQ select Canadian data from PIAAC as the representative country due to its large sample size. However, the sample size is not an adequate reason to choose a dataset to train the model. Therefore, we improve their work by applying the same method to other country datasets and compare the model performances (Research question 1). Also, we provide a country comparison on the employment shares at risk (Research question 2).

2. **Jobs and industries at risk**

   The risk of automation for each individual in different countries is available at the end of the first part of the analysis. Thus, the level of the impact of automation risk can be assessed with the aggregation of individuals that perform the same occupation (job level), that operate in the same industry (industry level), and finally, that work in the same country (country level). For the cross-country comparison in the first part, individual level automation risk is on the country level. The interest of the second part is job and industry level (Research question 3) which will help us to provide OECD wide recommendations.

3. **Socio-demographic characteristics of workers at risk**

   PIAAC provides socio-demographic information for all individuals that attended to the survey. Therefore, the aggregation will be possible for the socio-demographic characteristics of workers at risk (Research question 4). Also, we investigate to what extent the characteristics of workers are related to automation risk. Thus, an analysis is conducted to explain the probability of automation as a function of the socio-demographic characteristics of workers. The explanatory power of the relationship indicates the level of importance of policy attention towards the vulnerable groups.

4. **Policy recommendations**

   When the characteristics of the workers at risk are known, policy recommendations will be more accurate. The analysis results and the literature review are combined for policy advice (Research question 5).

Chapter 3

Technology, employment and skills

This chapter provides background information for the analysis of automation risk. First, it discusses the effects of the previous technological advancements on the employment shares and skill levels of employees. After that, it gives information about how technological changes are related to the employment structure and occupational wages. Following that, looking at the previous changes, possible impacts of current technologies on employment are reviewed. Finally, it discusses the important details of the different approaches to calculate the expected risk of automation for jobs and provides reasons to choose a certain approach.

3.1 History of technological advancements, employment and skills

Over the years, employees with different levels of skills have been affected by the technological advancements in different ways. The breaking points are industrial revolutions. Figure 3.1 illustrates the waves of technological advancements since the 18^{th} century. Even though the fourth wave is shown in the figure, its effects have not been evident yet. Therefore, only the expected effects are emphasized for the last wave. The first three waves are discussed in this section and the fourth wave is discussed in the third section.

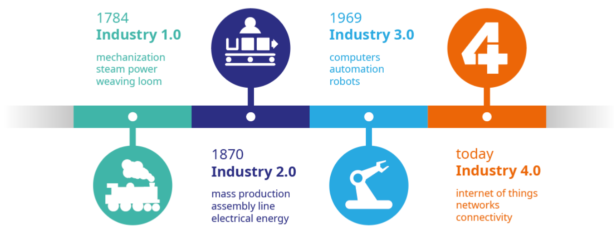

Figure 3.1: Waves of technological advancements (Busch Group, 2020)

The first wave, at the end of the 18^{th} century, started with the mechanization, which led to the automation of the agriculture sector. The self-made tools were replaced by the sophisticated machines for the industrial production and unskilled labor were specialized in the non-routine manual tasks (i.e. tractor driving) (Autor, Levy, et al., 2003). Yet, as productivity increased, the need for unskilled agriculture labor decreased steadily (Ernst et al., 2019). Figure 3.2 illustrates the changes in the US employment shares by industries between 1800 and 2000 (Acemoglu, 2008). Even though the shares for the agriculture sector declined, it remained the backbone of the economy for a long time and provided job opportunities. In the early 1950s, the US and also other OECD countries experienced a sharp decline in the agriculture employment shares (Handel, 2012). Consequently, the unskilled agriculture workers had to find new job opportunities which they found in the manufacturing jobs.

At the end of the 19^{th} century, the importance of the manufacturing sector boosted. The new energy sources (electricity, gas and oil) emerged thanks to the new technological advances in the industry. The steel demand increased, new methods of communication (telegraph and telephone) were developed. At the beginning of the 20^{th} century, the automobile and plane were invented (Pouspourika, 2020). These advancements created new job opportunities. So, unskilled labor in the agriculture sector shifted to new jobs in the manufacturing sector. There has been an increase in the manufacturing employment shares between the 1950s and 1970s (Acemoglu, 2008; Handel, 2012). Overall, during the first two waves the demand for unskilled labor increased (Ernst et al., 2019).

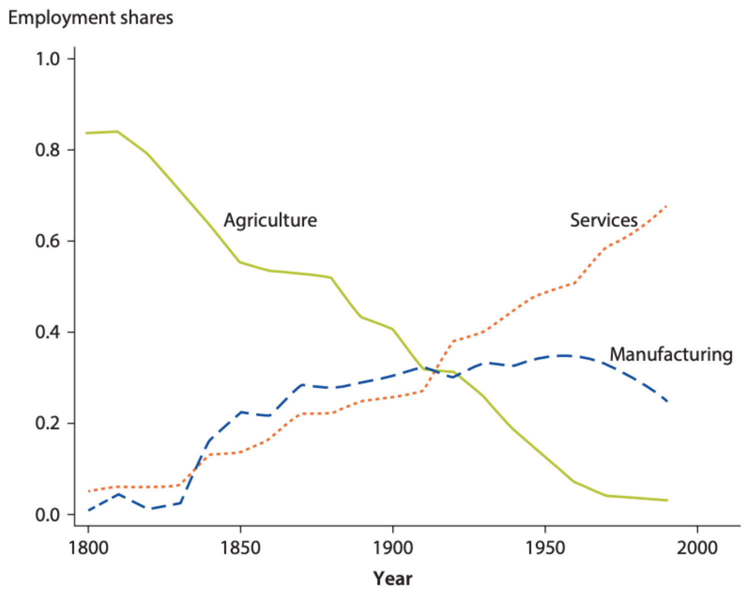

Figure 3.2: US Employment in Agriculture, Manufacturing and Services (Acemoglu, 2008)

Note: The figure is the reprinted by Halland et al. (2015)

The third wave began with the advancements in electronics and telecommunications, mainly the introduction of computers in the 1970s (Pouspourika, 2020). Starting from the 1980s, the number of robots used in the industry increased steeply which was one of the reasons for the decline in the manufacturing employment shares (Carbonero et al., 2018).[1] Some of the manufacturing laborers were replaced by robots as robots are capable of replacing the routine and repetitive tasks. Manufacturing labor was highly specialized in manual routine tasks with the adoption of the scientific management approach which was offered by Frederick Taylor (Ernst et al., 2019). This approach aims to maximize the labor productivity where the workers focus only on the manual repetitive tasks; a small part of the total production (for instance, tightening screws of the car door only). As the manufacturing jobs are required some level of training and skills, labor in this sector are considered as middle-skilled (Nedelkoska and Quintini, 2018). So, the third wave harmed the middle-skilled labor the most.

While the employment shares in the agriculture and manufacturing sector declined after the 50s and 70s, respectively, the shares in professionals, managerial and service jobs increased steadily (Handel, 2012). Figure 3.3 shows the changes in the occupational groups in the US. [2] Both low-skilled (elementary occupations and sales jobs) and high-skilled workers (managers, technicians and professionals) are required to perform the service jobs. So, middle-skilled workers in the manufacturing sector that were replaced by machines shifted to the (mostly low-skilled) service jobs due to its increasing job opportunities (Goos and Manning, 2007). With the replacement of middle-skilled, job polarization was observed which refers to an employment gap between low- and high-skilled workers in developed countries. The employment gap led to wage inequality. (Acemoglu and Autor, 2011; Nedelkoska and Quintini, 2018)

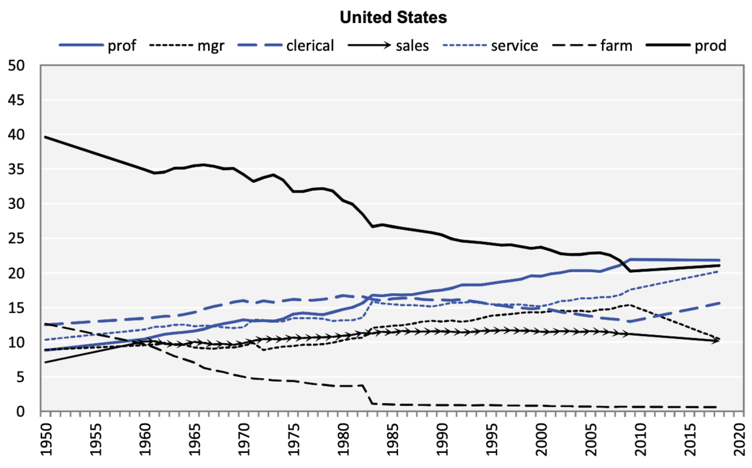

Figure 3.3: US Employment by occupation (Handel, 2012)

[1] There are a couple of reasons for the decline in the manufacturing employment shares: offshoring, decline in mobility and change in the skill requirements (Hernandez, 2018). As we are interested in the skill use at work, our focus is on the implementation of robots and accordingly changes in the skill requirements of the manufacturing workers.

[2] For detailed figures for other OECD countries see Handel (2012).

3.2 Technological changes

Technological change is the invention, innovation and diffusion of the technology. The technology is first invented, then commercialized (or released as open-source) and continuously improved (and get cheaper) and finally spread over the industry and society (Rousseau, 2010). The relationship between the employment structure and occupational wage can be explained by two different technological changes: skill-biased and routine-biased. The outcomes of the technological changes provide a baseline to discuss the possible impacts of the fourth wave of technological advancements in the next section.

3.2.1 Skill-biased technological changes

The skill-biased technological changes (SBTC) are the changes in the technology that favors the skilled over unskilled workers. It suggests that the relative productivity of skilled workers and also their relative demands got higher with the introduction of computers because the high-skilled were mostly performing non-routine tasks and computers started to perform routine tasks. When the relative demand is higher for skilled workers, higher wages are expected for them which causes wage inequality. Here, the skilled stands for more educated, more experienced or more able (Violante, 2008). High-educated workers are preferred as the level of skills increases with the educational attainment. When further technological advancements are common, the education enhances individuals' ability to learn, implement and spread new technology. Also, considering that non-routine tasks are intense in the occupations that require high-skilled workers, technology is complementary for skilled workers (Acemoglu, 1998; Card and Dinardo, 2002; Card and Lemieux, 1994; Goldin and Katz, 2009; Katz and Murphy, 1992; Krueger, 1993; Phelps and Nelson, 1966; Violante, 2008).

Table 3.1: Decomposition of task shifts into between and within industry components in the US (Autor, Levy, et al., 2003)

	1. Nonroutine analytic			2. Nonroutine interactive			3. Routine cognitive			4. Routine manual			5. Nonroutine manual		
	Total	Btwn	Wthn	Total	Btwn	Wthn	Total	Bwtn	Wthn	Total	Btwn	Wthn	Total	Btwn	Wthn
1960–1970	2.57	1.74	0.83	1.15	−0.34	1.49	2.20	1.14	1.06	4.01	2.39	1.62	−3.03	−2.28	−0.74
1970–1980	3.02	1.54	1.48	4.68	0.26	4.42	−0.14	0.33	−0.47	1.63	0.79	0.84	−2.25	−1.00	−1.25
1980–1990	2.97	0.92	2.05	5.31	0.52	4.79	−3.48	−1.42	−2.07	−1.47	−0.16	−1.31	−2.58	−1.27	−1.31
1990–1998	3.12	0.67	2.45	4.48	0.54	3.94	−4.88	−1.31	−3.57	−3.88	−0.38	−3.50	−0.63	−0.31	−0.31

Note: The US industries adapted the technological changes earlier than EU countries so the demand for the non-routine tasks increased earlier for the US. On the other hand, EU countries experienced late but more rapid change in the input (Handel, 2012; Nedelkoska and Quintini, 2018)

From the introduction of computers to the beginning of the millennium, the price of the equipment for information technology fell by 10% per year in the US (Gordon, 1990; Greenwood and Yorukoglu, 1997). The price decline induced the widespread adoption of computers in various industries. Following that wide usage of information technologies reduced the cost of communication, data storage and monitoring activities within firms (Milgrom and Roberts, 1990). So, workers started to compete with computers for specialized, routine, repetitive tasks and eventually the performance of computers were preferred (Nedelkoska and Quintini, 2018). However, routine and non-routine tasks are engaged in an occupation. According to Autor, Levy, et al. (2003), routine tasks are substitutable by the computers while non-routine tasks are complimentary. So, depending on the frequency of non-routine tasks, computers become complementary or substitutionary. Table 3.1 shows that starting from the 70s, within the same industry in the US, the labor input for the non-routine analytic and interactive tasks has increased while the input for the routine tasks and non-routine manual tasks has declined. Handel (2012) reaches the conclusion that these trends continue: (repetitive) physical and craft skills are declining while cognitive and high social interaction-required skills are increasing.

In parallel to an increase in the demand of employees working mostly on the non-routine tasks, the demand for tertiary education increased (Acemoglu and Autor, 2011). An indication to understand whether the technological changes are skill-biased is to look at the *return to skill*: the relative wages of college graduate workers to high school graduate workers. Michaels et al. (2014) report for 11 OECD countries that the relative wage for college graduates increased by 10 points between 1980 and 2004 as well as their supply. On the other hand, the wage for lower-educated labor declined by 18.7 points. So, the demand for low-educated labor was lower relative to the high-educated. Nedelkoska and Quintini (2018) suggest that the technological changes that influence the employment structure and wages are strongly dependent on the adaptation pace of the educational institutions. However, the response of these institutions (the supply of labor with the necessary skill levels) to demand shifts comes with a lag.

Critiques of SBTC

Various economic models support the SBTC structure where the central point is workers' level of education (Acemoglu, 2002; Aghion, 2002; Hornstein et al., 2005). However, historical evidence does not always support the outcomes of these models. Specifically, the recent phenomenon of job polarization, the high- and low-skilled employment shares increased relative to the middle-skilled in the last two decades and this was documented for the United States (Autor, Katz, et al., 2006; Autor and Dorn, 2013; Autor, Katz, et al., 2008), for the United Kingdom (Goos and Manning, 2007), for Germany (Dustmann et al., 2009; Spitz-Oener, 2006) and European countries (Goos, Manning, and Salomons, 2009; Michaels et al., 2014).

Figure 3.4 shows the changes in the US employment shares based on the skill levels (which is ranked by the mean wage) in three different decades (Acemoglu and Autor, 2011). During the first decade (1979-1989), the high-skilled high wage employment shares grew and the low-skilled low wage employment shares declined. This supports SBTC as the demand for the high-skilled increased. However, this creates wage polarization as the demand for low-skilled employment decreases. In the following decade (1989-1999), job polarization became pronounced where the high-skilled high wage and low-skilled low wage employment shares increased while the growth of middle-skilled shares was negative. Moreover, wage polarization was deepened during the 1990s as the employment was populated to either high wage or low wage occupations. As an example of wage polarization, Acemoglu and Autor (2011) report for the US that the workers with 18 years of education in 2009 earned 22% more in comparison to similar workers in the 1980s and workers with 7 years of education earned 18% less. Also, rising income inequality since the 1980s was documented for other OECD countries: The United Kingdom (Gosling et al., 2000), Canada (Boudarbat et al., 2003) and Germany (Antonczyk et al., 2010; Dustmann et al., 2009).

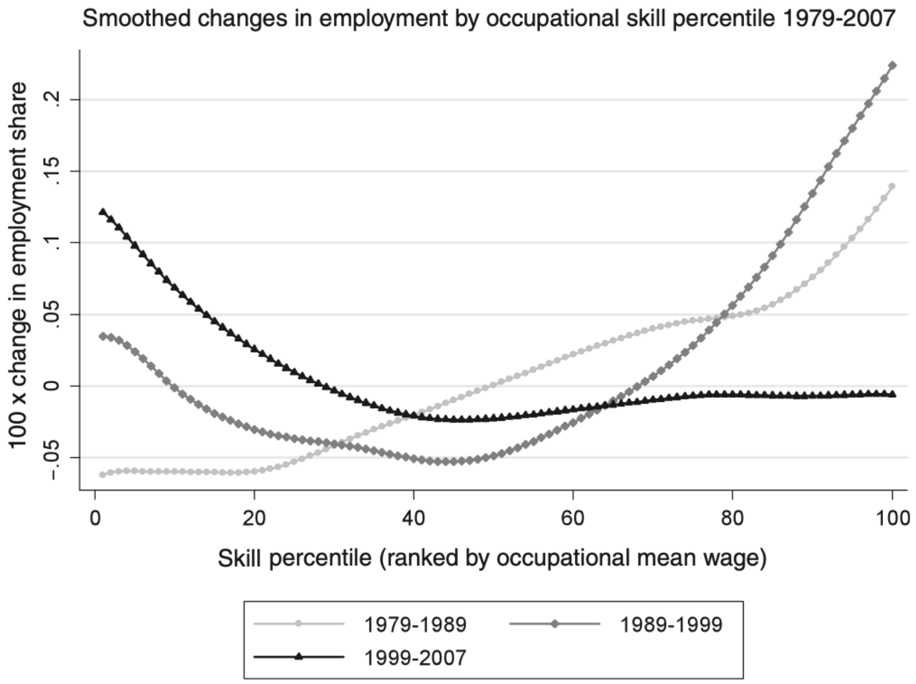

Figure 3.4: Changes in US employment by occupational skill percentile 1979-2007 (Acemoglu and Autor, 2011)

3.2.2 Routine-biased technological changes

When technical changes are accepted as the skill-biased, it is assumed that technology is complementary for high-skilled workers since they mostly focus on non-routine jobs. So, the demand for high-skilled is high and the demand decreases when the educational attainment is lower. However, job polarization that has been observed over the years can be explained when technological changes are routine-biased (Goos, Manning, and Salomons, 2014). RBTC centers the degree to which a job is routinisable instead of the level of education. For instance, there can be occupations that are intense with non-routine tasks but do not require high-level of education such as personal care work. Occupations are defined as the bundle of tasks and when the non-routine tasks are intense in an occupation, the advantages on employment and wages are high (Acemoglu and Autor, 2011; Caines et al., 2017). Historically, workers with different skills were specialized in different tasks: low-skilled in routine cognitive (record-keeping, calculation etc.) and non-routine manual tasks (truck driving etc.); high-skilled in non-routine cognitive tasks (medical diagnosis, legal writing etc.); middle-skilled in routine cognitive and manual tasks (picking, sorting etc.) (Nedelkoska and Quintini, 2018). Since the middle-skilled jobs include more routine tasks, middle-skilled workers are expected to be at more risk of being replaced by machines.

Critiques of RBTC

According to RBTC, higher wage growth is expected for occupations that are intense with non-routine tasks (Acemoglu and Autor, 2011; Caines et al., 2017). However, non-supporting evidence from the US (Autor and Dorn, 2013) shows that both routine and non-routine occupations include fairly high shares of low and high wage growth occupations (Figure 3.5). This means that a routine occupation can have a similar wage growth with a non-routine occupation. For instance, a machine operator who belongs to the routine occupation, and truck driver, in the non-routine occupation, have similar wage growths. Caines et al. (2017) suggest that the level of task complexity is an important aspect of wage growth. The task complexity is assessed for each O*NET occupations by looking at skills, abilities and work activities of the occupations. Tasks with the lowest level of complexity such as carrying, driving, archiving, cleaning or over-the-counter interaction involve raw physical, cognitive and interactive skills and abilities. On the other hand, high-complex tasks can be performed with some level of education. Examples to the high-complex tasks are problem-solving, decision-making, effective communication and tasks that require social interaction skills.

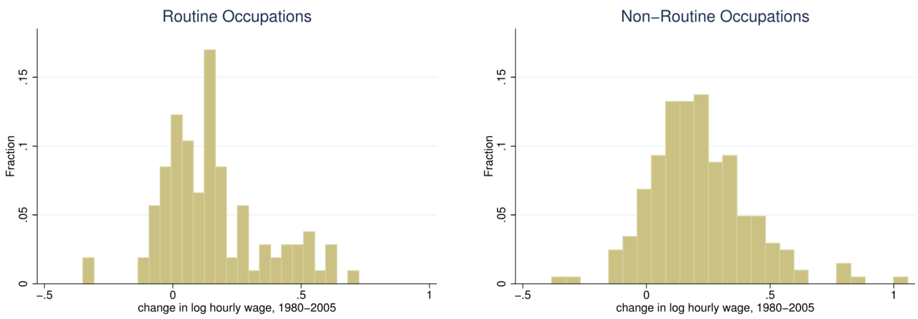

Figure 3.5: Distribution of hourly wage growth for routine and non-routine occupations in the US (Autor and Dorn, 2013)

Overall, RBTC centers the density of non-routine tasks in an occupation while the education and level of skills are the key aspects for the SBTC. Education is necessary to perform complex tasks (both routine and non-routine). Yet, there is a lag between the education (supply of labor) and demand shifts in the industry as education takes time. Still, after graduation, some skills can be learned during the work. Therefore, it is important to grasp the value of re-qualification and retraining activities to reduce the negative impacts of technological changes. According to Nedelkoska, Neffke, et al. (2015), between 1975 and 2010, the displaced workers in Germany that moved to occupations that require skill upgrades did not have a long-term earning losses while workers that moved to less-skill required occupations experienced large long-term earning losses.

3.3 Effects of the current technological advancements on employment and skills

The fourth wave of technological advancements are happening now and its impacts are getting more pronounced. This section compares the current wave with the previous ones and explains the possible impacts of current technologies on employment and skills from different perspectives.

With the current advancements such as in artificial intelligence (AI), it is possible to automate some of the complex non-routine tasks (Acemoglu and Restrepo, 2018). Machine learning (ML) algorithms provide opportunities to automate non-routine cognitive tasks by eliminating the human biases and by outperforming human in scalability (Frey and Osborne, 2013). In 2017, a study, aiming to show the effects of human bias in court cases, used a large dataset excluding demographic features like ethnicity, race and gender and found that the algorithm reduces the jail population by 42% and crime by 25% (Thompson, 2019). Moreover, the fraud detection which is now almost fully automated is a good example for the scalability feature of ML as the system can more accurately define the false pretences in big data (Phua et al., 2010). On the other hand, autonomous mobile robots (AMR) are used to automatize the non-cognitive manual tasks. Industrial robots were introduced in the previous wave and now AMR is incorporated with AI for logistics jobs at plants (as forklifts and cargo handling vehicles), agriculture sector (as agriculture vehicles), hospitals (as robots to transport samples and food) and homes (as vacuum cleaner and mop) (Federal Ministry of Labour and Social Affairs, 2017; Frey and Osborne, 2013).

As the AI-based innovations focus more on the automation of the complex non-routine tasks compared to the previous technological waves, it is expected to be less skill-complementary and more substitutionary (Ernst et al., 2019). Previously, with SBTC, machines were complimentary to skilled labor and, machines usually automatized the repetitive routine tasks. Consequently, the demand for skilled labor rose further with the introduction of computers in the 70s. The high-skilled labor was always at the lowest risk as they perform complex tasks such as human judgment, problem-solving, analytical skills, or various soft skills (Acemoglu and Restrepo, 2018).

Now, AI threatens high-skilled labor in complex non-routine tasks (Acemoglu and Restrepo, 2018). According to the McKinsey's report in 2016, even the highest-paid occupations like financial planners and surgeons are at risk of being (partially) automated (Chui et al., 2016). Acemoglu and Restrepo (2018) declare that the high-skilled will be impacted this time. Also, Ernst et al. (2019) suggest that the demand for high- and middle-skilled workers will decrease as opposed to the previous waves and the productivity of low-skilled will increase.

The reason is that AI takes the job of a high-skilled worker and provides expert knowledge to non-specialists. For instance, in the agriculture sector in which low-skilled labor is intense, AI increases productivity by guiding the agriculture labor to select the right seeds and track the growth of the plants. On the other hand, Frey and Osborne (2013) claim that while the high-skilled move to down-skilled jobs, the low-skilled will be pushed further with the risk of getting unemployed. Therefore, AI may cause deskilling of the workers.[3]

On contrary, according to McKinsey (2017)'s report, it declares that most of the job growths in the US and other advanced economies will be in the occupations that require a college degree or more while the occupations requiring secondary degree education or lower will decrease. For wage growth, occupations at the high end of the wage distribution and some low-wage occupations such as teaching assistants or nursing assistants are expected to grow. On the other hand, middle-income occupations will experience the largest occupation decline. Therefore, income inequality is expected to continue. However, if investment in the energy transition, building and infrastructure increases as a choice of policy in advanced economies, the demand for the middle-skilled could increase. Besides, as opposed to the above-mentioned studies that claim that complex tasks will be automated as well, McKinsey (2017)'s report highlights that the time spent for managing, communicating and advanced cognitive capabilities like creativity and reasoning will increase, while the time declines for the activities that machines can perform such as physical activities, collecting and processing data. Also, Grundke et al. (2018) suggest that workers with high numeracy skills are rewarded with a higher wage if they also have higher communication, managing and self-organizing skills. Therefore, even though automation risk is partial for high-skilled labor, still, the high-skilled are at the lowest risk. In 2014, a decision-making algorithm called VITAL was nominated to a Hong Kong-based venture capital firm's board of directors to show AI can be used for the decision-making processes (Barfield, 2018). Yet, in 2019, it is announced that VITAL is no longer used (Kahn, 2020).

In short, different studies anticipate different possible effects of technological advancements on high-skilled workers. The technological advancements have two competing effects on employment: destruction effect, arises from technology replacing the labor and the workers shifting to other jobs, and capitalization effect, occurs when more companies keep entering to productive industries and as a result, new job opportunities emerge (Aghion and Howitt, 1994). Historically, the destruction effect was dominant in the short run and the latter effect was observed in the long run (The Economist, 2016). However, according to McKinsey's report in 2015 as different from the previous waves, AI is expected to have 3000 time more impact as the innovations are happening 10 times faster at 300 times larger scale than the industrial revolutions in the 18^{th} and 19^{th} centuries (Dobbs et al., 2015). The direct substitution has already been observed. In June 2020, around 30 journalists employed by Microsoft were replaced by an AI software to select stories and edit the content (Marks, 2020). Or, self-driving cars threaten the driving jobs ranging from the taxi or Uber drivers to construc-

[3]Deskilling is the ability of the technology to substitute human skills and consequently human forgets such skills to use (Parasuraman et al., 2000; Zuboff, 1985)

tion machinery, truck and bus drivers which translates into 3.8 million workers in the US in 2017 (Winick, 2020). Yet, there are also many complementary examples such as medical diagnosis or anomaly detection in production (Frey and Osborne, 2013).

The overall productivity is expected to increase in the long run (Acemoglu and Restrepo, 2018). However, how the destruction and capitalization effect will be dominant over time will change with the improvement and diffusion processes of the current technological changes (Ernst et al., 2019). According to a study conducted in the 80s, the average time between the first commercial and the time of sale for 46 inventions in the 19^{th} and 20^{th} centuries was 14.4 years (Gort and Klepper, 1982). Therefore, the uncertainty is high about the future technologies and their impacts. Yet, technological changes heavily depend on the country policies, tax incentives and investments in the technological researches (Mazzucato, 2015; Neufeind et al., 2019). Policies influence the job design and task structure of jobs so, tasks in an occupation can vary with the training and supervision (Chentouf and Ernst, 2014). As the task structure of jobs changes, the skill requirements differ as well. Skills are features that can be improved and influenced by policies. Since this study aims to define jobs and industries at risk but also to define the groups of workers at risk, the substitution risk is calculated for occupations as a function of skill use at work. The reasons why only the substitution risk is calculated are that (1) the substitution risk is calculated with the current knowledge (current tasks and skills) and (2) the complementary effect is not certain since the changes in the diffusion process of the technological changes is indefinite (Brynjolfsson, Rock, et al., 2017).

3.4 Approaches

There are two approaches to calculate automation risk for occupations, namely occupation-based and task-based (Arntz et al., 2016). The results differ substantially. Therefore, it is wise to look from the perspectives of both and decide which one to choose for the analysis.

3.4.1 Occupation-based approach: Understanding the study of Frey and Osborne

As discussed earlier, FO predict 47% of the US employment is at high risk of being automated in the next decade or two. This analysis frightened most of the scientists and policymakers worldwide due to the fact that almost half of the US workers have the possibility to become unemployed. FO's study sparked a widespread discussion about the possible threats of current and future technological developments. Their study not only created a baseline for the occupation-based approach but also inspired studies that adopt the task-based approach. Thus, their perspective on the topic is explained in detail.

FO focus on the technological advancements in machine learning (ML) and mobile robotics (MR) related fields. ML algorithms including data mining, artificial intelligence, computational statistics and machine vision provide increasingly reliable opportunities to automate non-routine cognitive tasks like simultaneous language translation and legal writing. On the other hand, MR allows the automation of manual tasks like autonomous robots for logistic jobs and pattern recognition. So, unlike the previous technological developments, FO assume in their study that the current technological advancements are capable of performing tasks that have not been considered to be performed by machines but humans until now.

As discussed in the previous section, technological advances have either destruction effect or capitalization effect. FO argue that even though the latter effect has been historically dominant and human-beings have been highly capable of adopting new skills by education, with the current technological developments, cognitive tasks are expected to be automated and consequently cause high ratios of technological unemployment (Brynjolfsson and McAfee, 2012; Goldin and Katz, 2009).

With the focus area of advancements in ML (and MR) and expectations for the future employment in mind, FO build their research on a task model (Autor, Levy, et al., 2003). The task model suggests that computers are substitutes for routine tasks and complimentary for cognitive non-routine tasks. The model is designed by considering the routine tasks can only be substituted by machines if profitability is high for companies. So, substitution depends on not only the technological capabilities but also the comparative price differences between humans and machines. On the other hand, FO narrow down the scope and focus only on the substitution effect arising from technological capabilities. Considering that Autor et al. suggested this model in 2003, FO claim that, given the current technological advancements, the model can be expanded by redefining tasks that are substitutable by machines (such as accepting some of the cognitive non routine tasks can be automated). Yet, computerization of the non-routine tasks still has boundaries. FO call these boundaries *engineering bottlenecks*. Engineering bottlenecks refer to tasks that cannot be substituted easily by machines in the near future. Even though most of the tasks defined as bottlenecks can be overcome partially, still innovative approaches are required for better improvements. FO argue that the pace to solve these bottlenecks will have a high impact on the speed of computerization. Three task categories are defined.

1. **Perception and manipulation tasks**: Some of the simple tasks like identifying the geometry of the product are already automated. Yet, machines can perform only a limited number of tasks with reliable results. Humans still outperform machines in complex perception and manipulation tasks from handling objects in an unstructured environment to planning and selecting actions. So, there is room for improvements in the capability of machines in perception, learning, and planning, especially in an unstructured environment.

2. **Creative intelligence tasks**: Creativity is defined as the ability to create unique and valuable ideas (Boden, 2005). Ideas vary from writing a poem to scientific theories. Even though it is possible to see examples of machines creating drawings (Belzer and Kent, 2000) or composing music

(Cope, 1989), creativity changes by time depending on the social perception in different contexts. Therefore, humans are expected to dominate this area in the near future.

3. **Social intelligence tasks**: Tasks that require social interactions and the ability to respond by recognizing the emotions and psychology of others are difficult for machines to take over from human with the given knowledge. Therefore, human seems to be preferred for tasks requiring skills like negotiation, advising or communication even in the long run.

FO use 2010 version of O*NET data to estimate the substitution risk arising from technological capabilities. O*NET includes detailed information about 903 occupations in the US, task contents and variables (abilities, skills and knowledge) to perform these occupations. To relate this data with the employment and wage information, 903 O*NET occupations are aggregated into 702 occupations of the Labor Department's Standard Occupational Classification (SOC).

FO aim to estimate future impacts. However, the impacts of most of the technological developments on employment have not been observed yet. Thus, they consult experts about future expectations. ML experts from the University of Oxford are asked to label 70 occupations selected from 702 SOC descriptions, whose labels are believed to reach a consensus at the end of the meeting. By looking at the occupations and their task descriptions, ML experts assign 1 (if the occupation is expected to be *fully* automated) or 0 (if the expectation is *partially* automation) to each of 70 occupations. Then, FO predict the probability of the 702 occupations to be *fully* automated in the near future by relating occupational labels with the bottleneck-related O*NET variables.

However, the drawback of the expert opinions is that the labels may contain subjective bias, meaning that the labels are defined subjectively by the limited number of experts and labels may be misleading. Therefore, first, to minimize this bias, FO test if *subjective* labels are related to the *objective* O*NET attributes that are related to the engineering bottlenecks. To do that, they examine the power of the O*NET attributes in estimating the probability of an occupation to be automated: Both occupational labels and the bottleneck-related attributes of the labelled occupations are implemented into the four different probabilistic models, namely logistic regression and variances of Gaussian process classifiers; exponentiated quadratic, rational quadratic and linear covariances. While logistic regression predicts the probability of automation risk for 70 labelled occupations given the O*NET attributes with 82.7% success, exponentiated quadratic model fits the data the best with 89.4% success. The high success ratios confirm that the subjective judgement is consistent with the objective 9 O*NET variables and thus estimating the risk of automation for 702 occupations is fairly reliable. So, FO do not only estimate the automation risk probabilities for the remaining 632 occupations but also labelled 70 ones. One of the occupations may be labelled as 0, but the probability of risk can be higher than expected for that occupation. Table 3.2 shows the bottleneck-related attributes as defined in the study of FO.

FO divide occupations into three categories as low risk (less than 30%), medium risk (30-70%) and high-risk (>70%) occupations. Then, they merge the risk levels with the employment numbers for the US from the Bureau of Labor Statistics. And the famous 47% of the US employment that are at high risk of automation is calculated. FO do not specify any time for the predictions but mention "*occupations are potentially automatable over some unspecified number of years, maybe a decade or two*". FO reach the conclusion that the risk is higher for the low-wage occupations that require low-skilled workers. Also, as can be seen in Figure 3.6, most of the jobs that are at high-risk are related to service, sales and administration.

The findings of FO are interpreted as potential threats arising from technological advances (Arntz et al., 2016). And, FO are criticised for not considering that automation can be applied to tasks rather than the entire occupation. This critic leads us to explore the task-based approach.

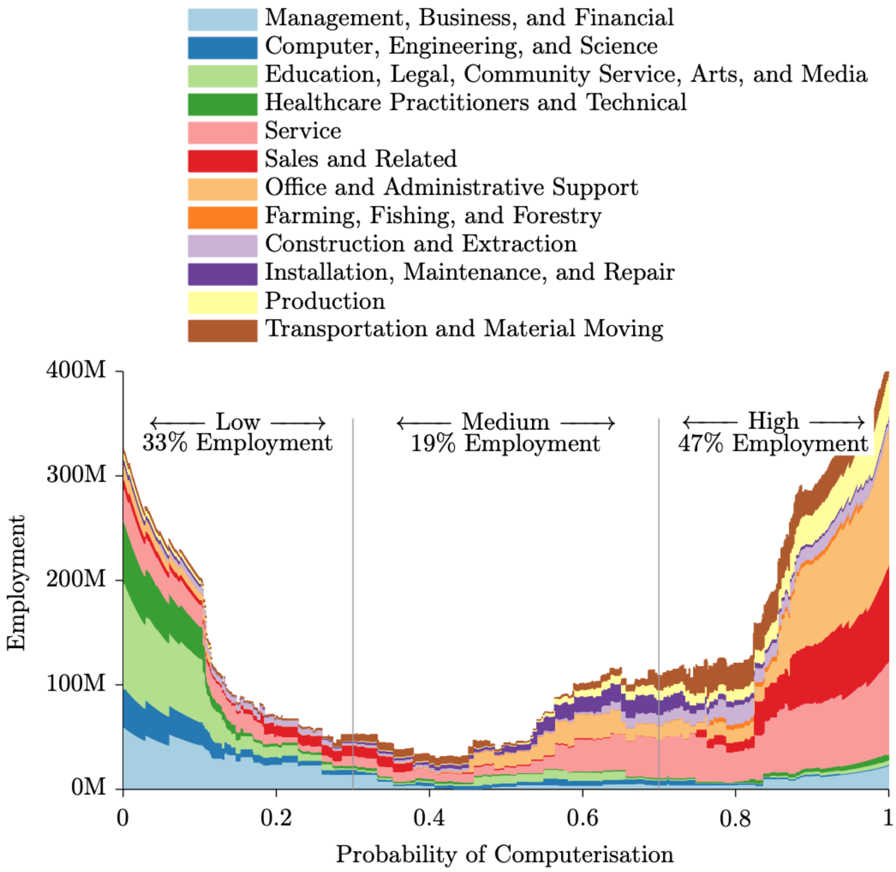

Figure 3.6: The distribution of US employment by risk levels (Frey and Osborne, 2013)

Table 3.2: Engineering bottleneck related O*NET variables (Frey and Osborne, 2013)

Computerisation bottleneck	O*NET Variable	O*NET Description
Perception and Manipulation	Finger Dexterity	The ability to make precisely coordinated movements of the fingers of one or both hands to grasp, manipulate, or assemble very small objects.
	Manual Dexterity	The ability to quickly move your hand, your hand together with your arm, or your two hands to grasp, manipulate, or assemble objects.
	Cramped Work Space, Awkward Positions	How often does this job require working in cramped work spaces that requires getting into awkward positions?
Creative Intelligence	Originality	The ability to come up with unusual or clever ideas about a given topic or situation, or to develop creative ways to solve a problem.
	Fine Arts	Knowledge of theory and techniques required to compose, produce, and perform works of music, dance, visual arts, drama, and sculpture.
Social Intelligence	Social Perceptiveness	Being aware of others' reactions and understanding why they react as they do.
	Negotiation	Bringing others together and trying to reconcile differences.
	Persuasion	Persuading others to change their minds or behavior.
	Assisting and Caring for Others	Providing personal assistance, medical attention, emotional support, or other personal care to others such as coworkers, customers, or patients.

3.4.2 Task-based approach

FO's results are directly transferred to different countries at the occupational level: for all European countries (Bowles, 2014), Germany (Brzeski and Burk, 2015) and Finland (Pajarinen and Rouvinen, 2014). The employment numbers at risk of automation are high, similar to those of FO. The reason why automation risk is high with the occupation-based approach is that FO build their research on the fact that task descriptions in an occupation do not vary across individuals. So, they do not take into account that workers can perform different tasks within the same occupation. However, task structure changes to a great extent across individuals within the same occupation (Autor and Handel, 2013).

As an alternative, Arntz et al. (2016) (AGZ hereafter) take into account that an occupation does not always include the same tasks so tasks may differ for each individual within the same occupation. This approach is called the task-based. AGZ use PIAAC data, which is an adult skill survey conducted by OECD (OECD, 2017). As individual-level data on the job content is accessible in this survey, it allows us to consider that tasks performed vary for individuals within the same occupation. The risk is calculated for not only the US but also other OECD countries. The study is concluded that 9% of the individuals in the US are at high risk as opposed to 47% by FO while the share of employment alters between 6% to 12% for different OECD countries.

On the other hand, when tasks differ within the same occupation, attributes associated with the engineering bottlenecks remain the same. These bottlenecks are universal and when solved, it will influence the pace of automation for all countries. Yet, the importance and intensity of using these attributes for each individual within the same occupation may change depending on the industry and country structure (Nedelkoska and Quintini, 2018). NQ base their analysis on the FO's study but adopt the task-based approach as AGZ do. So, NQ also use PIAAC data for their analysis. Different from the study of AGZ, NQ focus on how the intensity of the skills used at work that are related to the engineering bottlenecks influences the risk of automation. NQ conclude that 10% of employment in the US are at high risk while this ratio is 14% across all OECD countries.

To replicate the FO's study, NQ need to relate the labelled occupations and engineering bottlenecks to the variables in PIAAC data. First, 70 labelled occupations in O*NET are manually matched with the occupational codes in PIAAC. Since there is no perfect match between the occupation descriptions, some of the O*NET occupations have more than one equivalents in PIAAC and sometimes the same ISCO-08 code is used for two of the 70 occupations (see Appendix A in the study of NQ). Secondly, engineering bottleneck-related variables in O*NET are translated into PIAAC. Even though PIAAC variables sufficiently match with those in O*NET, no perfect match exists. For instance, PIAAC does not involve any question related to *assisting and caring for others*. Therefore, this may affect the people working especially in the healthcare and service sectors. Also, PIAAC does not have any specific variable that highlights *cramped work space,awkward positions*, which is about working in an unstructured environment. Table 3.3 indicates the PIAAC variables that represent one of the types of engineering bottlenecks.

Table 3.3: Engineering bottleneck related PIAAC variables (Nedelkoska and Quintini, 2018)

Engineering bottlenekcs	Variable in PIAAC	Varible code	Variable description
Perception manipulation	Fingers, (dexterity)	F_Q06C	How often - using skill or accuracy with your hands or fingers?
Creative intelligence	Problem-solving, simple	F_Q05A	How often - relatively simple problems that take no more than 5 minutes to find a good solution?
	Problem-solving, complex	F_Q05B	Problem solving - complex problems that take at least 30 minutes thinking time to find a good solution?
Social intelligence	Teaching	F_Q02B	How often - instructing, training or teaching people, individually or in groups?
	Advise	F_Q02E	How often - advising people?
	Plan for others	F_Q03B	How often - planning the activities of others?
	Communication	F_Q02A	How often - sharing work-related information with co-workers?
	Negotiate	F_Q04B	How often - negotiating with people either inside or outside your firm or organisation?
	Influence	F_Q04A	How often - persuading or influencing people?
	Sell	F_Q02D	How often - selling a product or selling a service?

NQ use logistic regression, which is one of the methods that FO use to predict the risk of automation of occupations in different countries. Since NQ conduct the analysis for all OECD countries, they choose a representative country to select the 70 labelled occupations from. Canada dataset is chosen because its sample size is the largest with more than 26,000 observations while the second largest dataset belongs to Poland with almost 9,500 rows. The larger dataset is better for the analysis since the more people that hold the same labelled occupation, the higher representatives of the different individuals within the same occupation. On the other hand, the drawback to rely on only one country's observations is that other countries are assumed to have a similar industry structure and global position in the predictions. So, while replicating NQ's study, we take into account this drawback and choose the representative country by looking at the performance of the analysis. After defining the datasets used, the ways of improving the NQ's method will be elaborated in the following section.

3.5 Conclusion

This chapter had four main steps to provide background information about skills and employment. First, to explore how the importance of skill demands and shares of employment changed with the previous technological advancements. Second, to identify how the technological changes explain the market outcomes, specifically the employment structure and wages. Third, to explain different views about the future impacts of the current technological developments on skills and employment. The final fourth goal was to define two approaches that were used to calculate the risk of automation for different occupations.

The demand for low-skilled labor was high during the first two waves of automation. Due to the automation in the agriculture sector (1^{st} wave), the low-skilled labor shifted to the manufacturing sector and the manufacturing employment shares increased between 1950 and 1970 (2^{nd} wave). As the interaction was high between the manufacturing laborers and machines, some level of skills is required. So, manufacturing workers are educated to be middle-skilled. Then, with the introduction of computers and robots, repetitive routine tasks were automated and middle-skilled people at plants were affected negatively, which caused job polarization. In parallel to that, the demand of high-skilled labor increased (3^{rd} wave).

The technological changes are interpreted as skill-biased, meaning that the more educated workers are, the less susceptible they are to the negative impacts of automation. Computers and robots were only capable of performing routine tasks and computer-performance was preferred in such tasks. So, the importance of non-routine tasks increased which was mostly done by high-skilled labor. Therefore, technology became complementary for high-skilled workers. Increase in the demand of the high-skilled, as well as an increase in their wages, also support this view. However, this trend changed a decade later in a way that the share of middle-skilled jobs decreased and job polarization was observed. This trend was explained with the routine-biased technological changes. If the intensity of routine tasks is higher, then the risk is higher and wages are lower irrespective of the level of education. In short, both views agree that high-skilled workers were at the lowest risk and wage inequality was high. However, when looked at the wage distributions between 1980 and 2005, wages were similar for routine- and non-routine-intense jobs (Autor and Dorn, 2013). So, in the literature, it was found that the level of task complexity is more important than tasks being only non-routine in wage growth.

AI, on the other hand, now provides an opportunity to automate complex non-routine tasks and its impacts are bigger (and faster) than the previous technologies. Overall, it is more skill-substitutionary rather being complementary. So, AI threatens the high-skilled labor which may result in deskilling of them. It is also possible that high-skilled workers may take low-skilled labors' jobs since the importance of low-skilled jobs is expected to increase. On the other hand, the results of other studies show that high-skilled will be still at the lowest risk and the middle-skilled jobs will be affected the most. However, how technologies will impact the labor force is still uncertain as it depends on the diffusion process of the new technologies. Still, policies could change the task structure and provide opportunities to vulnerable groups to learn new skills.

To define who is at the highest risk, the substitution risk of automation is calculated in various studies. There are two main approaches: the occupation-based and task-based. The main difference is that the occupation based approach uses the standard features (tasks, skills, abilities and knowledge) of an occupation to calculate the risk while the task-based uses features that are specific to each individual. So, the latter approach considers that each individual within the same occupation may have different tasks to perform and skills to use. Therefore, the risk calculation is very different in these two approaches. FO who calculated the risk of automation with the occupation-based approach, identify three task categories that are not easy to automate in the coming two decades and nine related features that consist of skills, abilities and knowledge. NQ use only the skills that are associated with FO's nine features to calculate the risk of automation and also consider that the tasks and skills of each individual differ. The reason why only skills are included in the analysis is that skills can be learned, practised and improved by interventions. So, policies have an impact on the skills. In this study, we use NQ's method and conduct the analysis to calculate the risk of automation. In the next chapter, the data used and method followed are explained in detail.

Chapter 4

Estimating the risk of automation

This chapter aims to present the methods applied and data used for predicting the jobs and industries at risk of automation as well as defining the socio-demographic characteristics of workers at risk. First, it gives an overview of the data sources used. Then, it concludes with explaining the steps of the analysis and the ways of improving the NQ's model. Finally, in the end, an answer to the first research question is provided.

4.1 Datasets

In this section two datasets are discussed: The Occupational Information Network (O*NET) and Program for the International Assessment of Adult Competencies (PIAAC). FO use O*NET data to identify occupation descriptions while Arntz et al. and NQ combine information from O*NET and PIAAC. Both datasets can be considered for multi-purposes. This section explains which parts of these datasets are selected for the analysis.

4.1.1 O*NET

The Occupational Information Network (O*NET) is a project, which is sponsored by the US Department of Labor/Employment and Training Administration, provides information about occupation-specific descriptions such as knowledge, skills, tasks and work context on around 1000 occupations that cover the entire US economy (O*NET Resource Center, 2020). The project aims to understand how the characteristics of the US labor force change over time. O*NET Data Collection Program collects data from selected workers and occupation experts through questionnaires and keeps data up to date by annually incorporating new information into the O*NET database. Each occupation includes various tasks and activities and requires a different mix of skills, abilities, knowledge. So, this dataset includes occupation names with their standardized characteristics.

The content model of O*NET, in Figure 4.1, provides an overview about the types of information about work (O*NET Resource Center, 2020). Occupation-Specific Information category provides information about the occupations and their task descriptions. When machine learning experts label the 70 selected occupations, they consider the tasks of the occupations from this section. Currently, there are 1016 occupations and only 867 of them have detailed tasks descriptions while the remaining has only the titles. In 2013, there were 903 occupations in total. Frey and Osborne (2013) (FO) consider only 702 since the Bureau of Labor Statistics has only these occupations' wage and employment data. The unit of observations depends on the variables: occupation, task, skill etc.

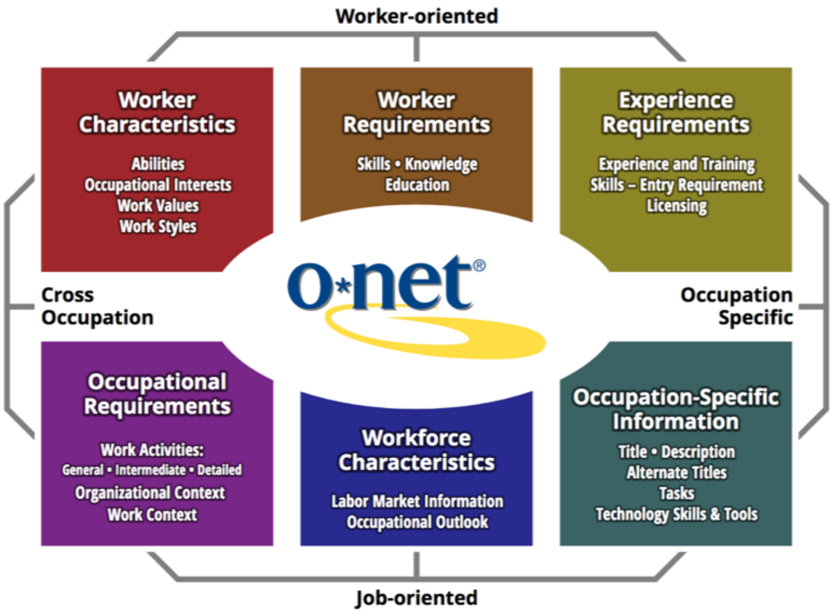

Figure 4.1: The content model of O*NET (O*NET Resource Center, 2020)

On the other hand, engineering bottleneck-related attributes defined in FO's study are from different categories. The number of unique values is 33 for knowledge, 35 for skill and 52 for ability. O*NET defines *fine arts* as knowledge, *social perceptiveness, negotiation* and *persuasion* as skills and *finger dexterity, manual dexterity* and *originality* as abilities. So, bottleneck-variables have not been selected from a single O*NET category. However, the definitions of skills, abilities and knowledge are very different and putting them together can be misleading. This concern will be discussed in the next section.

4.1.2 PIAAC

OECD has conducted an international adult skill survey as part of the Program for the International Assessment of Adult Competencies (PIAAC) for 40 different countries. The survey has two cycles so far and it is planned to be realised in every 10 years. In the first cycle, data was collected in three rounds, between 2011 and 2018 (see Figure 4.2 for more details). The second cycle started in 2018 and the new results are planned to be launched in 2023. The survey measures the skills of adults in three different areas: literacy, numeracy and problem-solving in technology-rich environments. Also, it investigates the skills used at work and in personal life. The unit of observation is person. The biggest advantage of PIAAC over O*NET is that PIAAC includes individual-level information which can change within the same occupation.

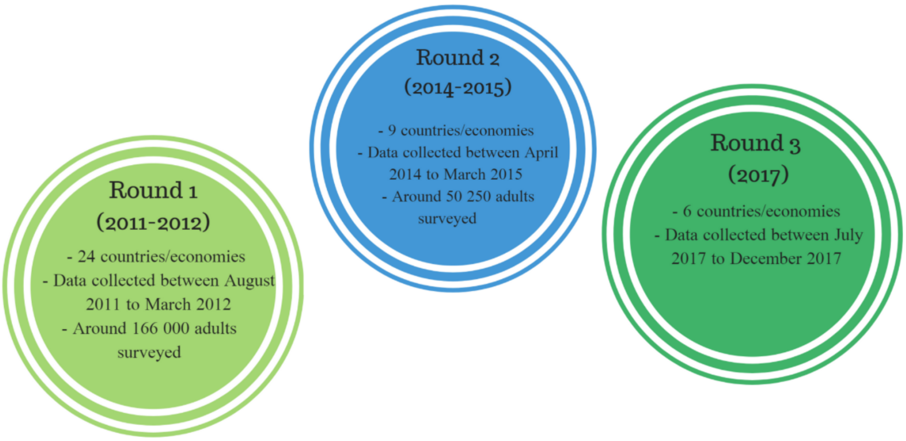

Figure 4.2: Rounds for the first cycle of PIAAC (OECD, 2017)

Nedelkoska and Quintini (2018) (NQ) are interested in the skill use at work and numeracy scores along with the occupation and industry codes. Therefore, we also consider those parts of the survey. The survey results help countries to understand how their education and training systems are and how they can improve the weak parts of their systems (OECD, 2017). The target groups are the adults aged between 15 and 65 who provide personal information such as age, gender, education levels and earnings. Demographic information of the individuals is especially useful while defining the socio-demographic characteristics of workers at risk. The standardized questions in the survey and approximately 5000 sample for each country facilitate the creation of a cross country analysis.

NQ include 32 different countries into the analysis. Some of the OECD countries are not involved in the PIAAC survey like Iceland, Latvia, Luxembourg and Switzerland. Australia joined but the data is not publicly available. NQ have data for Ireland and Northern Ireland. However, only data for Ireland is available online. Portugal joined too but later was dropped out of the analysis due to an error in the background questionnaire (OECD, 2013). Hungary was also dropped out of the analysis, however, the dataset is still available online.

NQ do not analyze Mexican data but it is also available. Therefore, we included Hungary and Mexico into our analysis. Colombia is not included in the survey as they joined the OECD in 2020. In addition, a few non-OECD countries are added into the NQ's analysis: Cyprus, Russia and Singapore. We include these countries as well.

As PIAAC survey assess the adult skills, NQ select the bottleneck-related attributes of PIAAC only from the skill category and investigate the impacts of skill use at work on automation risk of occupations. However, the attributes that are related to the engineering bottlenecks defined in FO's study are from different categories: ability, skills and knowledge. Yet, not all the categories can be improved by training or influenced by policy implementations. Knowledge is the theoretical information and does not have to be practical. Skills are inferred capabilities and can be developed with certain training while abilities are inbuilt characteristics and not easy to influence. So, skills are the capabilities that can be impacted by policies. This is the reason why we also choose only the skills to estimate the risk of automation.

Skill-use questions in PIAAC are in the form of *How often - do you use your skill x at work?* and the answer is at the ordinal scale, varying between 1 to 5. While 5 represents that skills are used highly, 1 represents non to little usage. So, this survey takes into account that the level of importance and intensity of these skills can vary for different individuals within the same occupation. Depending on the answers to the skill questions, automation risk for individuals will change.

4.2 Methodology

This section aims to explain the methods used during the analysis and highlight what is different from NQ's study.

4.2.1 Data preparation

As discussed earlier, the first step of NQ while replicating FO's study with the task-based approach is that finding the corresponding 70 labelled occupations and engineering bottlenecks in the PIAAC data.

Labelled occupations: Corresponding O*NET and PIAAC

The occupation classification systems are different in O*NET and PIAAC datasets. O*NET adopts the Standard Occupational Classification (SOC) system, which is a statistical standard in the US (U.S. Bureau of Labor Statistics, 2018). The last version of the SOC code is from 2018. However, FO use 2010 version of SOC for the labelled occupations, so we look into the 2010 version as well. There are four hierarchical levels: 23 major, 97 minor, 461 broad groups and finally 840 detailed occupations (U.S. Bureau of Labor Statistics, 2018). O*NET has an additional 63 detailed occupations which do not contain any SOC code. However, each of the 70 O*NET occupations has a SOC code. Figure 4.3 shows an example from SOC system.

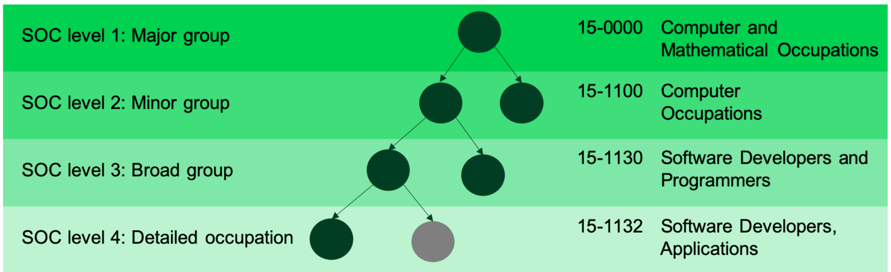

Figure 4.3: Hierarchical levels in the SOC system

On the other hand, the PIAAC survey includes another occupational code system called the International Standard Classification of Occupations (ISCO). ISCO-08 classification was created in 2008 which is the revision of the 1988 version, ISCO-88. ISCO-08 also has four hierarchical levels: 10 major groups, broken into 43 sub-major groups. Each sub-major group is broken into minor groups, of which there are 130. At the most detailed level, 436 unit groups exist (ILO, 2012). One of the notable differences between these two systems is that occupations are classified based on the similarity of skill levels and skill specializations in ISCO-08 while similarity is on the task content for SOC (ILO, 2012; U.S. Bureau of Labor Statistics, 2018).

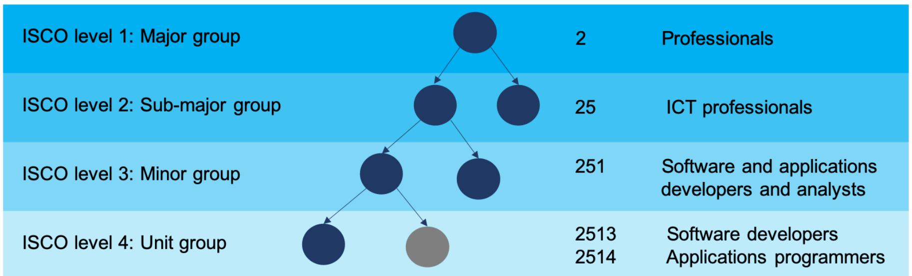

Figure 4.4: Hierarchical levels in the ISCO-08 system

The number of occupations at the most detailed level in SOC (840) is higher than in ISCO-08 (436). Figure 4.4 shows an example of the ISCO hierarchical levels. This example is comparable with the above SOC example. The highest level of ISCO is broader and actually covers level 1 in the SOC system. At the lowest level, on the other hand, more than one ISCO-08 code is needed to explain the SOC code. Reversely, an ISCO-08 coded occupation may cover more than one SOC-coded. For instance, "car, taxi and van drivers" is in the ISCO-08 system and the equivalents in SOC are "light truck or delivery services drivers" and "taxi drivers and chauffeurs". So, while transferring the 70 occupations from SOC to ISCO-08 system, it is possible that one occupation in O*NET has multiple equivalents in PIAAC or multiple O*NET occupations have only one representative in PIAAC.

NQ manually match equivalents of 70 labelled occupations with PIAAC data. However, a correspondence table from 2010 version of SOC to ISCO-08 is available online (U.S. Bureau of Labor Statistics, 2018). Interestingly, the PIAAC equivalents of 70 occupations in these two tables are fairly different.

NQ's match table

NQ subjectively select 79 ISCO-08 coded occupations as the equivalents of 70 occupations. When an occupation has more than one equivalents, all of them are labelled the same. According to NQ, some of the occupations do not have any ISCO-08 match: dishwashers; parking lot attendants; technical writers; paralegals and legal assistants; gaming dealers; farm labor contractors; claims adjusters, examiners, and investigators. Furthermore, we checked if all the ISCO-08 coded occupations in NQ's match table are in the current PIAAC dataset that we are using for the analysis. We found 11 occupations that are indicated to be in PIAAC dataset by NQ but could not be found with the same names in the current PIAAC dataset (see Table 4.1). Considering that NQ published their research in 2018, we conducted this study in 2020 and ISCO codes are lastly revised in 2008, we would not expect to have different names in NQ's match table. As NQ do not provide the related SOC and ISCO codes in their table, it is not possible to trace back how they come up with these names.

BLS's correspondence table

US Bureau of Labor Statistics (BLS) provides not a perfect but a complete match between SOC codes and ISCO-08 codes. It is not perfect since the classification is done based on the similarities of different aspects (tasks in SOC and skills in PIAAC). There are 80 corresponding ISCO-08 coded occupations and all of them can be found in the PIAAC data. Thus, compared to NQ's match, BLS offers a more reliable correspondence and we choose to continue with the PIAAC equivalents from the BLS's correspondence table (see Appendix A.1).

Labelled occupations selected by Frey and Osborne (2013)	**Occupations that are in NQ's table but not it PIAAC with this name**	**PIAAC equivalents in BLS table**
Athletes and Sports Competitors	Athletes, sportspersons and related associate professionals	- Athletes and sports players
Compliance officers	Process control technicians, other	- Customs and border inspectors - Government social benefits officials - Government licensing officials
Computer-Controlled Machine Tool Operators, Metal and Plastic	Stationary plant and machine operators, other	- Metal working machine tool setters and operators
Chief Executives	Directors and chief executives	- Senior government officials - Traditional chiefs and heads of villages - Managing directors and chief executives
Zoologists and Wildlife Biologists	Biologists	- Biologists, botanists, zoologists and related professionals
Meter Readers, Utilities	Meter readers	- Meter readers and vending-machine collectors
Healthcare Practitioners and Technical Workers, All Other	Health professionals, other Health associate professionals, other	- Midwifery associate professionals - Traditional and complementary medicine associate professionals - Health associate professionals not elsewhere classified
Electrical and Electronics Drafters	Electronics and telecommunications engineering technicians	- Draughtspersons
Credit Authorizers, Checkers, and Clerks	Credit and loan officers	- Statistical, finance and insurance clerks

Table 4.1: Occupations that are in NQ's match table but not in the actual PIAAC with this ISCO-08 occupational name

Engineering bottlenecks

FO conducted their study in 2013. So, the bottleneck-related O*NET attributes were defined and occupations were labelled in 2013. NQ identify the corresponding PIAAC attributes in 2018. In five years, impacts of technological advancements, especially impacts of AI, has become more visible and widespread. So, it is well possible that some of the advancements after the FO's study can provide solutions to automate some of the bottleneck-related tasks sooner than expected. Nonetheless, the time between the survey is conducted and bottlenecks are defined are not far from each other. All of the countries that are involved in the NQ's analysis are participated in the PIAAC survey either in the first round or the second round of the first cycle, between 2011 and 2015. So, we expect the answers to the skill questions to be consistent with the defined bottlenecks. All bottleneck-related skill variables defined by NQ are included in the analysis.

When looked at all the defined *skill use at work* questions in PIAAC, we identify two more skills that are not easy to be taken over from human and be replaced by the machines. These skills are *presentation* and *cooperation.* Participants are asked to answer how often they present and how much time they spend to cooperate with their co-workers. Presentation is not only about the visuals and information but also the influence of the presenter on the listeners. Cooperation, on the other hand, requires social interaction and capability to communicate with people. Even though FO do not include presentation or cooperation as bottleneck-related O*NET attributes, these skills are by definition fit the *Social intelligence* category. To illustrate that presentation and cooperation are appropriate skills as bottlenecks, we conduct the analysis with and without including them and compare the results.

4.2.2 Logistic regression

The next step is to predict the probability of an individual being at risk of automation. Probabilities are calculated using 80 labelled ISCO-08 coded occupations and bottleneck-related skills. We apply logistic regression as NQ do. Logistic regression is a statistical model that predicts the probability of an object being in a certain *class* given the values of the categorical or numerical independent variables (*features of the object*). In our case, the classes are the labels: 1, if automation risk of an occupation is expected to be *high*; 0, if the risk is expected to be *partial* in the near future. The occupational labels are denoted as $y \in \{1, 0\}$. It is important to highlight that we are not interested in labelling an occupation as 0 or 1 but estimating the probability of an occupation being at risk. We investigate the probability of $y = 1$ for occupations described by the bottleneck-related skills. NQ relate 10 skills with the bottlenecks and we add 2 more (presentation and cooperation). These skills form the feature vector, denoted as $x \in R^{12}$.

To learn the patterns between the occupational labels and skills used at work, only labelled data can be used in the logistic regression. The labelled data is selected from the representative country. We also run the analysis with the labelled data from all countries but the explanatory power of the model was low compared to using a single country dataset or using a combination of country datasets. The country selection process is explained in the following section. Labelled data refers to the answers of individuals who hold one of the 80 labelled occupations. Data is divided into train and test sets to check the consistency between the labels and answers to skill use questions. In the literature, there is not a single rule about the split ratio. The main point is that the ratio of the training data should be higher so that the model can learn the patterns, then test data is used to measure the performance of the model and to check if the pattern is consistent. We split as 75% to train and 25% to test.

The number of unique occupations with a label is 80, but multiple individuals exist for each labelled occupations. So, the training data will be $D = (x_{train}, y_{train})$ where $x_{train} \,\epsilon\, R^{12\times 0.75\times N}$ and $y_{train} \,\epsilon\, \{1, 0\}^{0.75\times N}$. N represents the number of people in PIAAC who hold one of the 80 labelled occupations within the representative country. N is defined after deciding the representative country from which the training dataset is selected. N is multiplied by 0.75 as training set is 75% of the data. The training dataset D provides information about how y varies as a function of x. We use D with two purposes. First, to decide which country to select by looking at the performance of the logistic regression and calculate the coefficients accordingly. Secondly, to calculate the probability of an unlabelled occupation being at risk of automation, $p(y_{unlabelled} = 1 \mid x_{unlabelled},\ y_{train},\ x_{train})$ by using the coefficients.

Country selection

The probability of automation risk of occupations is modelled as a sigmoid (logistic) function. We select the representative country based on the highest performance of the model, so by applying the test data into the model.

$$p(y_{test} = 1|f_{test}) = \frac{1}{1 + e^{f_{test}}}$$

f is the discriminant function for the probabilistic classification that is modelled as logistic

$$f(x) = \beta \times x$$

where β represents the coefficients for each skill feature $x \,\epsilon\, R^{12}$. βs are calculated in light of the training data: the logistic relationship between the features (x_{train}) and labels (y_{train}). So, depending on the country, data change, thus the βs vary. To choose which βs fit the model best, we rely on the results of the performance metrics, more specifically AUC.

Performance metric: Area Under the ROC curve (AUC)

Two terms need to be explained to understand what AUC is: the confusion matrix and receiver operating characteristics (ROC). The confusion matrix is used to calculate most of the metrics (See 4.2). It shows the actual and predicted labels for the test data. ‘Positive’ symbolizes the category that we are looking into. Since our interest is to calculate automation risk, positive represents $y = 1$ and negative shows $y = 0$. True or false depends on the match between the actual and predicted labels.

Table 4.2: Confusion matrix

		Predicted category	
		Positive (1)	Negative (0)
Actual category	Positive (1)	True positive (TP)	False negative (FN)
	Negative (0)	False positive (FP)	True negative (TN)

ROC curve is a graph that shows the performance of a classification model at all possible thresholds. Threshold values can vary but the default value is 50%. The model predicts a probability for an object being in a certain class. Based on the probability being higher or lower than the threshold, the object is assigned to a class. For instance, if 60% is set as the threshold and the probability of the object is lower than 60%, the class would be 0. Different threshold values result in the changes in true positive ratio (TP ratio) and false positive ratio (FP ratio) so the ROC curve and AUC change as well. For the analysis, we use the default threshold (50%). The ratios are calculated as follows.

$$TP\,ratio = \frac{TP}{TP + FN}$$

$$FP\,ratio = \frac{FP}{TP + FN}$$

AUC is the area under the ROC curve. AUC represents the probability that the model positions the object with a positive category ($y = 1$) higher than the object with a negative category ($y = 0$). This is also obvious when looked at the curve as it is closer to the true positive axis (TP). If estimating the probability of an object being in a certain category is preferable rather than assigning the object to the correct category, AUC would be the right metric to look at. AUC varies between 0 and 1. The more correctly the model predicts, the closer AUC is to 1.

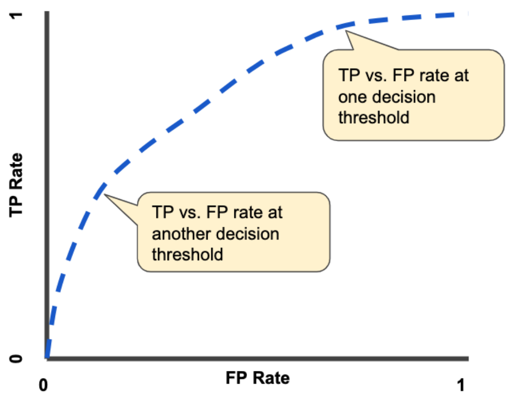

Figure 4.5: ROC curve (Google, 2020)

Countries included into the selection process

33 countries are included in the analysis. Different from the NQ's analysis, data from Mexico and Hungary are included while data from Northern Ireland is not available online. In total there are 585 unique occupation names with an ISCO-08 code. All digit levels are included but all the labelled data are at the 4-digit level. We calculate the probability of risk for each individual even though their occupation code is less than 4 digit level. Due to privacy reasons, some of the countries do not provide any detailed ISCO codes. Instead, all occupations are coded either *Valid skip* or *Not stated or inferred*. There are 9 countries without any valid code: Austria, Canada, Estonia, Finland, Germany, Ireland, Singapore, Sweden and the United States. It is not possible to select one of these countries as the representative country because labelled occupations cannot be matched. So, 24 countries remain for country selection. We apply logistic regression to each country separately and also to the combination of various countries.

Table 4.3: Performance of logistic regression with various countries; best performance in bold

Country	AUC	Number of labelled rows	Number of labelled occupations
United Kingdom	0.80	880	34
New Zealand	0.78	893	69
France	0.77	1050	59
Poland	0.69	1045	68
New Zealand + United Kingdom	0.75	1773	73
United Kingdom + Poland	0.74	1925	74
New Zealand + France	0.68	1943	75
New Zealand + United Kingdom + Poland	**0.80**	**2818**	**78**
All countries	0.66	17735	80
Canada	0.74	4656	79

Table 4.3 shows the results for AUC with the number of labelled occupations in the country datasets. Only the first three countries with the highest AUC scores (United Kingdom, New Zealand, France) and their combinations are presented in the table. In addition, Poland is included as it is the second country with the largest dataset (after Canada). Since Canada does not provide detailed occupational codes due to privacy reasons, information for Canada is taken from NQ's study. We also select labelled rows from all countries (countries with ISCO-08 code available), but the AUC is notably low compared to the other results. Even though the UK has the highest AUC, most of the labelled occupations are missing. Countries may not include samples for each occupational code. So, reliability is low for the UK. The highest AUC in the presence of 78 labelled occupations belongs to the combination of New Zealand, the UK and Poland. So, representative countries are New Zealand, the UK and Poland. In this case, N is 2818. The advantage of using a combination of countries is that the level of representation is higher compared to using only one country.

To provide more details about how AUC is calculated for the combination of the United Kingdom, New Zealand and Poland, confusion matrix and the ROC curve are presented below. Test data from the representative countries, which is 25% of the total labelled occupations in the representative countries, are classified as 0 or 1 depending on the labels in the training data. According to the confusion matrix in Figure 4.6, 517 $(303 + 214)$ of the occupations are predicted correctly while 188 $(100 + 88)$ of them are mislabelled.

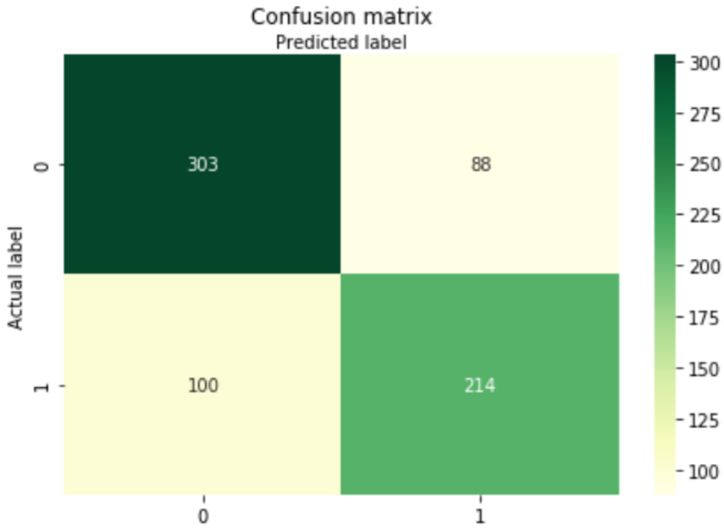

Figure 4.6: Confusion matrix for the selected representative countries

With values in the confusion matrix, true positive (TP) and false positive (FP) rates are calculated which lead us to the ROC curve in Figure 4.7. AUC presented under the ROC curve is 0.80. The important remark is that 0.80 AUC value does not only include 10 bottleneck-related skills defined by NQ but also two other skills: cooperation and presentation as we suggest these skills involve human interaction. Therefore, we conduct the analysis with the same labelled occupations from the representative countries with and without the additional skills. When two skills are added to the analysis, the AUC value increases by exactly 1 point. Even though the difference is not remarkable, we include cooperation and presentation into the analysis.

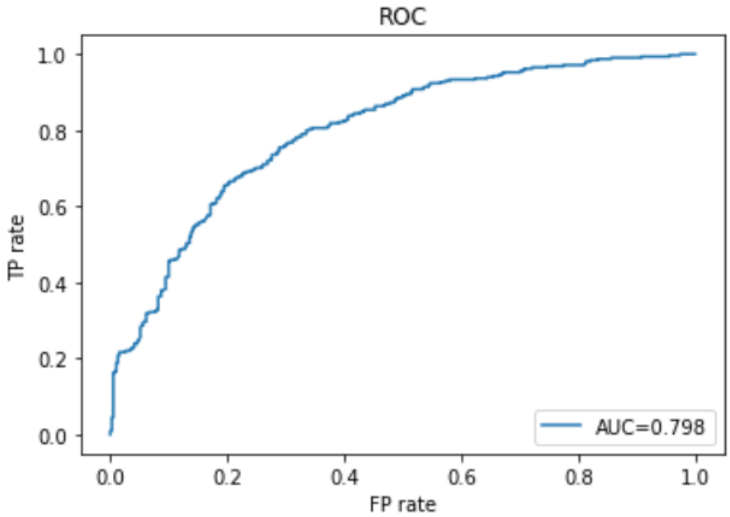

Figure 4.7: ROC for the selected representative countries

Coefficients and probabilities

Coefficients (βs) are estimated on the representative countries' data, more, specifically, with the training data from these countries. In the discriminant function ($f = \beta \times x$), βs are known and this function can be applied to all other individuals across 33 countries in PIAAC.

After estimating coefficients with a model that works with 80% success, the probability of unlabelled occupations at risk across countries can be calculated. However, we also include individuals with the labelled occupations and perform the analysis for all individuals regardless of having a labelled occupation or not. The aim is to fix the subjective bias that may occur when machine learning experts label 70 occupations. So, we also assign labelled occupations a probability as a function of skills vector. In this case, the logistic function is introduced again.

$$p(y_{individual} = 1 | f_{individual}) = \frac{1}{1 + e^{f_{individual}}}$$

For the discriminant function $f_{individual}$, the *feature vector* is denoted as $x \ \epsilon \ R^{12 \times 142,258}$, matrix of 12 skills for each of the 142,258 individuals. Each individual receives a probability of being substituted by the automation technologies based on their answers to skill questions. Because the answers at the ordinal scale, the risk differs for individuals within the same occupation. Including individual level probabilities with more than 100,000 rows into the report is impossible. Thus, in the following chapter, the results are first presented at the employment level and then at the country, industry and job level, which are the answers to the 2^{nd} and 3^{rd} research questions.

Survey weights

Each individual has a *final survey weight* that indicates the number of people in the employed population who hold the same occupation within the same country and would possibly give similar answers to the survey questions. Therefore, all observations need to be weighted using the final survey weights to be able to generalize the survey findings. Only then, the probability of risk of automation calculated for each individual becomes valid for each represented group of people in the employed population in a country. The column name is "SPFWT0".

4.2.3 Linear regression

Linear regression is the method used to explain automation risk as a function of the socio-demographic characteristics of workers. Three models are built. Model 1 includes age, gender, the level of education, PIAAC numeracy scores and country fixed effects as the independent variables. In model 2, occupation and industry dummies are added on top of the model 1. Sample size decreases between two models due to missing values in occupation and industry columns. Finally, model 3 has the variables from the first model and uses the sample size of the second model. The third model aims to see if the results are robust. If the coefficients do not differ to a great extent, it means that results are robust. The independent variables are the same as those in NQ's study. The variables in PIAAC are presented in Table 4.4. Dummy variables are created for the variables except for age and numeracy scores. Education has 9 levels including *primary or less education*. Numeracy score represents the mean of 10 possible scores for an individual. There is no explanation in NQ's study about why only numeracy scores are included in the analysis.

Table 4.4: Independent variables in PIAAC

Name in PIAAC	Label	Level
AGE_R	Age	Ratio
GENDER_R	Gender	Nominal
EDCAT8	Level of education (8 categories)	Ordinal
PVNUM1 - PVNUM10	Numeracy scale scores	Scale
CNTRYID	Country ID	Nominal
ISCO08_C	Current Job (ISCO 2008)	Nominal
ISIC4_C	Current Industry (ISIC rev 4)	Nominal

The age variable is not accessible for 7 countries: Canada, New Zealand, Hungary, Singapore, Germany, Austria, The United States. Also, publicly available files (PUFs) in PIAAC does not include any information about the level of education of individuals for Canada, Germany and Estonia. Therefore, we do not include these 8 countries into the regression analysis. The results are valid only for the remaining 25 countries in Model 1. In addition, New Zealand does not provide industry names in its PUF, thus we remove it in Model 2 and Model 3.

4.3 Conceptual model

Figure 4.8 is the summary of this chapter along with the outcomes for each research question. Data-related boxes are colored in gray. Methods and processes are presented in shades of blue. Finally, outcomes of the analysis (the answers to the sub-research questions) are in green.

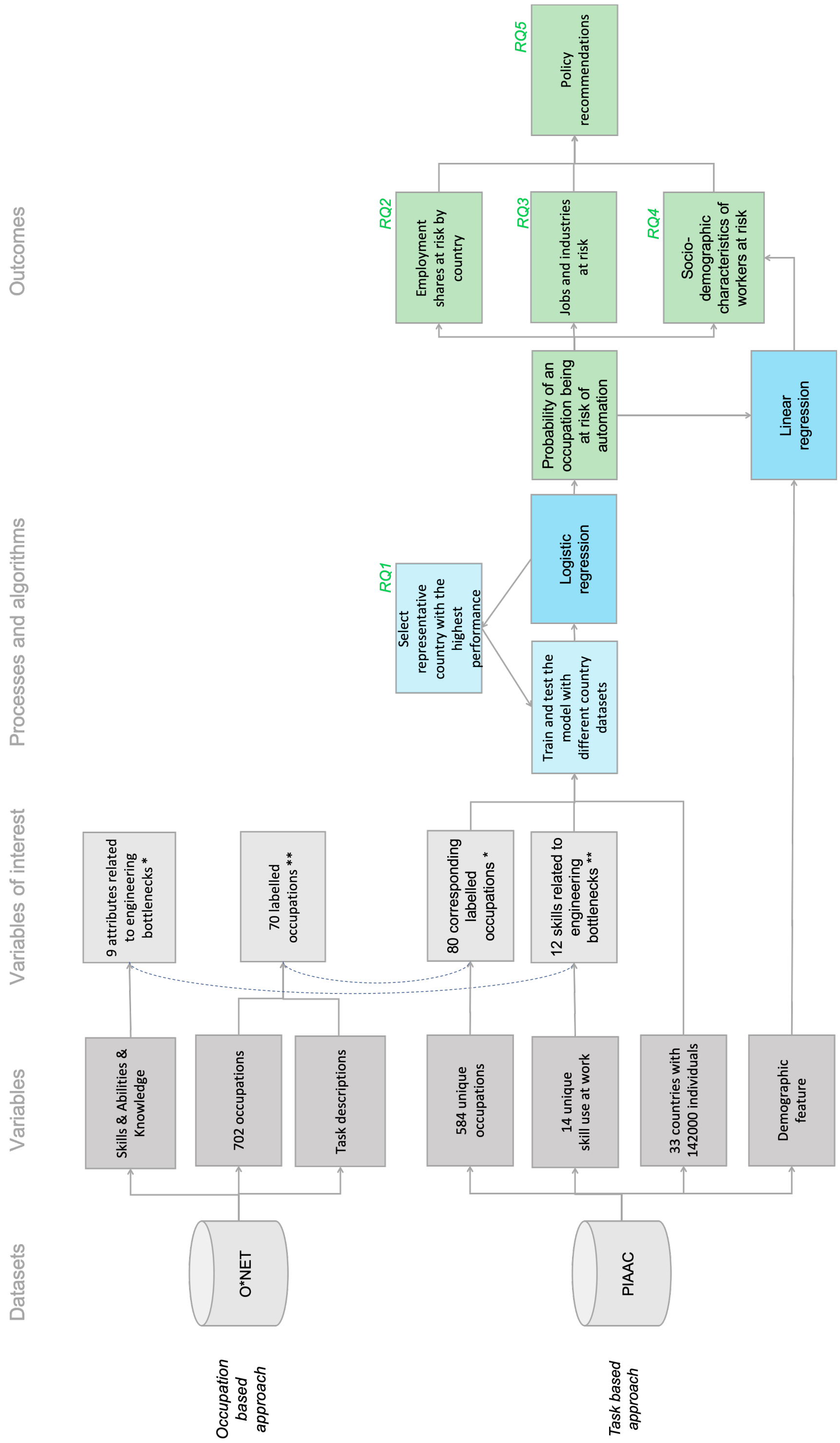

Figure 4.8: Detailed flow of the research

Chapter 5

Who is at risk of automation?

This chapter aims to present the results of the analysis. The starting point was to investigate how skills used at work are related to automation risk. The analysis continues with the exploration of the socio-demographic characteristics of workers. First, it discusses the coefficients from the regression models. Then, it presents automation risk at three different levels: country, industry, and occupation. Finally, it explains the characteristics of workers. This chapter seeks to answer the 2^{nd}, 3^{rd} and 4^{th} research questions. Also, during the discussion, the results are compared with those of NQ.

5.1 Skills used at work and automation risk

According to FO, bottleneck tasks can not be substituted easily by the current technologies in the near future. To perform these tasks, FO define certain bottleneck-related abilities, knowledge and skills. However, we only investigate the relationship between bottleneck-related skills and automation risk. To understand the relationship, first, we need to understand what logit coefficients tell us as these coefficients are used to predict automation risk for each of the 142,258 individuals from different countries. Table 5.1 summarizes the logistic regression results.

The initial expectation is to have a negative relationship between the bottleneck-related skills and automation risk, meaning that when these skills are used, the risk should reduce. Therefore, the signs of logit coefficients are expected to be negative. *Planning for others*, *presentation* and *influencing* are the top three skills that reduce the risk of automation the most. On the other hand, using *finger dexterity*, *complex problem solving*, *communication* and *selling* skills increase the risk. The most unexpected skill that is not negatively associated with automation risk is *communication*. The survey question related to communication is about the frequency of sharing work-related information with colleagues. So, it is one of the key skills for most of the workers and it may be over/undervalued by the workers. On the other hand, finger dexterity, which is under the perception and manipulation bottleneck category, is about using hands and fingers.

However, the description is not as precise as the O*NET descriptions and does not include the meanings like manipulation of objects or working in an unstructured environment. When only using fingers is considered, a positive sign is not that surprising. Regarding the coefficient signs, the only difference with NQ's study is that NQ find that complex problem solving has a negative impact on automation risk. The different may arise from using different country datasets so that answer differs. Besides, coefficients for simple problem solving and negotiation are not significant ($p\ value > 0.05$) in our study so it is appropriate to treat them as 0.

Table 5.1: Logistic regression results on automation risk as a function of bottleneck-related skills in PIAAC

	Logistic regression results		
	Logit coefficients	Standard errors	p values
Finger dexterity	0.0866	0.030	0.004
Simple problem solving	0.0078	0.039	0.844
Complex problem solving	0.1850	0.041	0.000
Teaching	-0.1507	0.037	0.000
Advising	-0.0833	0.036	0.019
Planning for others	-0.3147	0.032	0.000
Communication	0.1623	0.042	0.000
Negotiation	-0.0154	0.034	0.650
Influencing	-0.1696	0.036	0.000
Selling	0.1446	0.028	0.000
Presentation	-0.2462	0.052	0.000
Cooperation	-0.1319	0.034	0.000
Contant	0.7271	0.209	0.000
Observations	2818		
Pseudo R-squared	0.146		
AUC	0.798		

For the technical comparison with NQ's study, the dataset size (number of observations), pseudo-R-squared and AUC are discussed. In general, larger datasets are preferred for predictive analysis and machine learning approaches because the model can learn more details about the dataset and provide more reliable insights. Therefore, in our case, larger dataset is better. To calculate the logit coefficients, NQ use 4656 observations (the number of individuals with a labelled occupation from the Canadian data) while the number of observations is 2818 for our representative countries. Nonetheless, the dataset size is not the best indication to understand model performance. Therefore, we look into the performance metrics. Pseudo R-squared is used to measure how well independent variables (skills) explain the dependent variable (automation risk). So, R-squared allows us to compare which dataset explains automation risk the best. R-squared is 0.137 for NQ and 0.146 for our analysis. Since our model has a higher pseudo-R-squared, the explanatory power of our model is higher. Finally, AUC measures how accurate the probabilities are for automation risk. In our analysis, AUC is higher by 5.5 points which is preferable. Overall, our model performance is higher than NQ's. So, higher AUC and pseudo-R-squared values for our model indicate that NQ's claim about choosing Canadian dataset due to its large sample size does not hold up. NQ's model is improved by using different datasets (the combination of the UK, New Zealand and Poland) as the representative countries even though the sample size is smaller.

5.2 Socio-demographic characteristics and automation risk

Next, we explore the relationship between the socio-demographic characteristics of workers and automation risk. Automation risk is the predicted probability for each individual that comes from the logistic regression whereas worker's characteristics come from the PIAAC survey and include information about individuals' education level, gender, age and numeracy scores. Table 5.2 illustrates the linear regression coefficients for three different models. Model 1 and 3 have additional country dummies while model 2 includes country, occupation and industry dummies.[1] Even though model 2 has a higher explanatory power (See adjusted R-squares in Table 5.2), we take into account only the coefficients of model 1.[2]

Age and high numeracy scores are negatively associated with automation risk which does not change much when additional occupation and industry dummies are added in the second model. Besides, our model shows that women are more vulnerable to automation technologies than men. With additional dummies, the risk for females increases further. Since we calculated automation risk in relation with the skills used at work, high automation risk for females indicates that females use less bottleneck-related skills such as presentation, advising and teaching compared to their male colleagues within the same occupation.

[1]Dummy variables take only the value 1 or 0 to indicate whether the categorical effect exists or not. If a dummy independent variable takes the value 0, it means that the coefficient does not influence the dependent variable. On the other hand, if the value is 1, then the coefficient influences the value of the dependent variable. In our model, country, occupation and industry are the dummy independent variables. For instance, if an individual lives in the Netherlands, the country dummy will be 1 for the Netherlands and 0 for other countries. The coefficient for the Netherlands will change the value of the dependent variable.

[2]Control variables like occupation dummies may cause a "bad control problem" in the presence of the education variable (Angrist and Pischke, 2009). Bad control variables are correlated with both the explanatory (x) and outcome (y) variables (Söderbom, 2011). In our case, it is expected that people with a higher level of education are more likely to get a high-skilled job since the type of occupation that someone holds depends on the level of education. So, occupation control is the outcome of the education variables while it should be explanatory of automation risk. Still, results can be used for the robustness check. We compare the coefficients in model 1 and model 3. Both models include the same variables but have different sample sizes. We find that coefficients are very similar, meaning that results are robust (Table 5.2). The reason why sample sizes are different is that model 3 uses the sample size of model 2 and when industry and occupation dummies are included in model 2, sample size decreases due to the missing industry and occupation values.

Automation risk is likely to be lower with the increase in the level of education. Figure 5.1 illustrates the monotone relationship between automation risk and the education level. Thus, we can deduce that the risk of automation can be reduced by educating people and teaching them new skills. When occupation and industry dummies are included in model 2, the effects of education decreases, however, still the relations remain negative. NQ find slightly stronger negative impacts between education coefficients and automation risk. Overall, the results highlight the need to put extra attention to female workers and also the importance of education that has long-term societal impacts.

Table 5.2: Linear regression results on automation risk as a function of socio-demographic characteristics of workers

	OLS regression results					
	Model 1		Model 2		Model 3	
	OLS coefficient	Robust standard errors	OLS coefficient	Robust standard errors	OLS coefficient	Robust standard errors
Numeracy	-0.0003	0.000	-0.0002	0.000	-0.0003	0.000
Female	0.0204	0.001	0.0333	0.001	0.0233	0.001
Age	-0.0081	0.000	-0.0073	0.000	-0.0085	0.000
Age squared	0.000	0.000	0.000	0.000	0.000	0.000
Lower secondary education (ISCED 2, 3c)	-0.0199	0.003	-0.0069	0.003	-0.0186	0.004
Upper secondary (ISCED 3A-B, C long)	-0.0582	0.003	-0.0294	0.003	-0.0591	0.003
Post-secondary, non-tertiary (ISCED 4A-B-C)	-0.0705	0.004	-0.0254	0.004	-0.0694	0.005
Tertiary/professional degree (ISCED 5B)	-0.1204	0.003	-0.0521	0.004	-0.1209	0.004
Tertiary/bachelor degree (ISCED 5A)	-0.1617	0.003	-0.0669	0.004	-0.1629	0.004
Tertiary/master degree (ISCED 5A)	-0.1861	0.004	-0.0805	0.004	-0.1839	0.004
Tertiary/research degree (ISCED 6)	-0.2221	0.007	-0.0996	0.007	-0.2250	0.008
Tertiary-bachelor/master/research degree (ISCED 5A, 6)	-0.1713	0.006	-0.0715	0.006	-0.1703	0.006
Constant	0.7497	0.008	0.6553	0.014	0.7563	0.009
Country effects (33 countries)	Yes		Yes		Yes	
Occupation dummies (ISCO 08; 2-digits)	No		Yes		No	
Industry dummies (ISIC rev 3, 2-digits)	No		Yes		No	
Observations	89510		74929		74929	
Adjusted R-square	0.174		0.314		0.169	

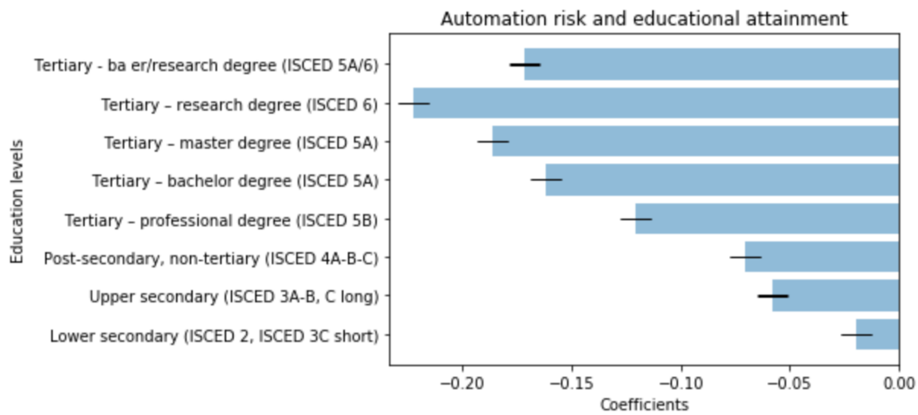

Figure 5.1: Partial correlation between the level of education and automation risk

The independent variables in model 1 explain 17.4% of the variation in automation risk. The ratio increases to 31.4% when occupation and country effects are included. Even though the number of observations in NQ's model is higher by around 55,000[3], the explanatory powers of our models are roughly 3 points higher than NQ's models. NQ indicate that R-square is "*astonishingly low given the attention that has been given to the role of skills in technology-labor relationships*". Even though our R-squares are higher, it is not ideal. Clearly, there are other variables that have higher impacts on automation risk but not in our scope. For instance, we do not consider task structures of occupations or companies' strategies.

5.3 Country-level risk

In this section, first, the distribution of automation risk is presented at the country level. Then, the shares of employment at different levels of risk is discussed and compared with the previous findings in the literature.

5.3.1 Distribution of automation risk

Table 5.3 shows mean and median automation risks by country.[4] Countries are ranked by increasing median risk. The median risk of automation for 33 countries is 48% and the mean is 45%. While Finland, New Zealand and the United States have the lowest median risk with 38%, the ratio increases to 58% at the other extreme for the Slovak Republic and Lithuania. The differences between

[3]The difference between the observation sizes is due to not including the same countries with NQ. As mentioned in the previous chapter, there are 25 countries included to the linear regression due to missing values for some of the variables. However, all the coefficients have high levels of significance ($p < 0.01$). Also, coefficients have the same signs and fairly similar magnitudes with NQ's finding except for the numeracy variable.

[4]While the median is the value in the middle, mean is the average value. Median values are preferred to compare countries if extreme values exist in the dataset or the distribution is skewed. Figure 5.2 indicates that the distributions are skewed, therefore, we discuss the median values.

the average median and the extreme values are half standard variation. So, the variation in automation risk is fairly large. As automation risk is calculated as a function of skills used at work, the high variation shows that skills vary to a great extent across countries. NQ calculate the median risk as 48% and standard deviation as 20% which are the same or very close to our findings. However, median risks show only the values in the middle and do not give us much detail. To gain a better idea of how the risk changes within countries, the distribution of risk is an important aspect to investigate.

Table 5.3: Country-level automation risk

Country	Median	Mean	SD
Finland	0.38	0.39	0.19
New Zealand	0.38	0.39	0.19
United States	0.38	0.39	0.20
Norway	0.39	0.39	0.18
Sweden	0.40	0.40	0.19
Singapore	0.41	0.41	0.20
United Kingdom	0.41	0.41	0.20
Denmark	0.42	0.41	0.19
Ireland	0.43	0.42	0.21
Canada	0.44	0.42	0.20
Netherlands	0.45	0.44	0.18
Israel	0.45	0.44	0.20
Korea	0.46	0.44	0.19
Estonia	0.46	0.44	0.19
Cyprus	0.48	0.46	0.20
Belgium	0.48	0.45	0.19
Japan	0.48	0.45	0.18
Austria	0.48	0.45	0.19
Chile	0.49	0.46	0.20
Slovenia	0.50	0.46	0.21
France	0.51	0.46	0.19
Czech Republic	0.51	0.48	0.19
Poland	0.52	0.47	0.20
Russia	0.52	0.48	0.19
Spain	0.53	0.48	0.20
Italy	0.53	0.48	0.19
Germany	0.54	0.50	0.18
Mexico	0.54	0.49	0.20
Turkey	0.54	0.48	0.19
Hungary	0.54	0.49	0.18
Greece	0.54	0.50	0.18
Slovak Republic	0.58	0.51	0.21
Lithuania	0.58	0.52	0.18
All countries	**0.48**	**0.45**	**0.19**

The distribution of automation risk per country, shown in Figure 5.2, is bimodal, has two peaks. This means that majority of individuals are either at high risk (around 60%) or low risk (around 30%). The density for South and Eastern European countries are concentrated around the right bimodal peak: The modes for Lithuania, Hungary and Poland are around 65%. Also, Germany and Japan have comparatively higher peaks on the right. One of the reasons is that these countries have large shares of manufacturing jobs (around 20%). However, the highest shares in manufacturing sector belong to Czech Republic (31%) and then Slovenia (29%) and we observe that their right peaks are still lower. So, the peaks do not only depend on the industry or occupation shares in a country. Since we take into account that skills differ across individual within the same occupation, the peaks also depend on the differences in the skills used at work across countries. On the other hand, Northern countries have moderate polarization with slightly higher left peaks at less than 25%. The highest shares for Norway and Sweden are at human health and social work activities while manufacturing has around 10% shares. Norway, the United Kingdom and New Zealand have more distinctive left peaks at around 20%. NQ argue that the reason of New Zealand having a low risk of automation compared to other OECD countries is that there has been a boost in cognitive jobs: professionals since 1990s and managerial jobs since 2010.

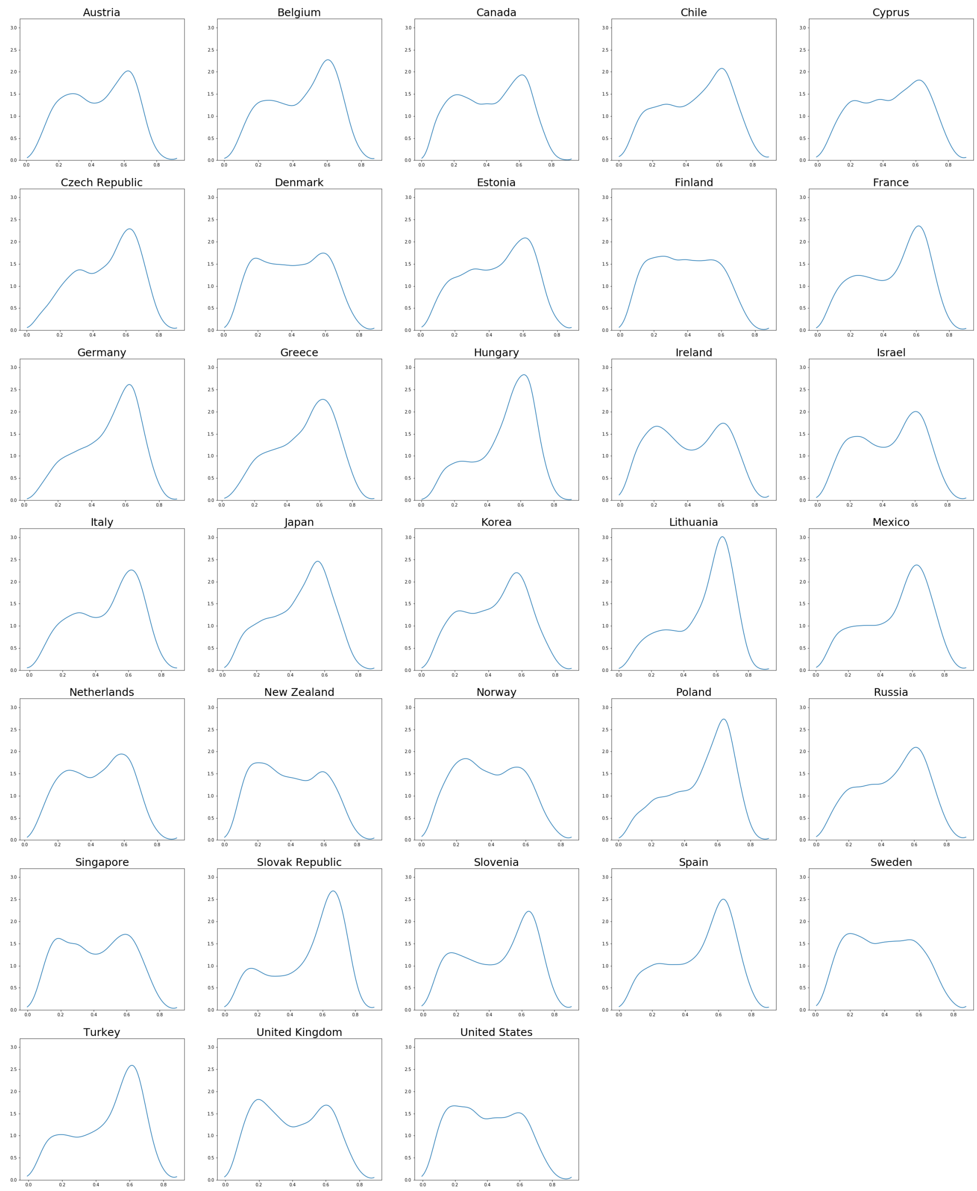

Figure 5.2: Distribution of automation risk by country

5.3.2 Employment shares at risk of automation

Knowing that the median risk of automation is 48%, the future seems to be scary. However, when the probabilities are translated into the employment shares, it turns out that the risk is more manageable than assumed. Figure 5.3 illustrates the employment shares at different levels of risk. Only 14% of the employment is at significantly high risk of automation (> 70%) across 33 countries. Yet, the variation of significant risk is rather high. While the shares of employment at significant risk vary between 6% to 11% for Northern countries and Northern America (the United States and Canada), these ratios jump to more than 20% for Mexico and the Slovak Republic. Besides, the largest employment ratios at small risk again belong to Northern countries (For more details see Appendix D.1).

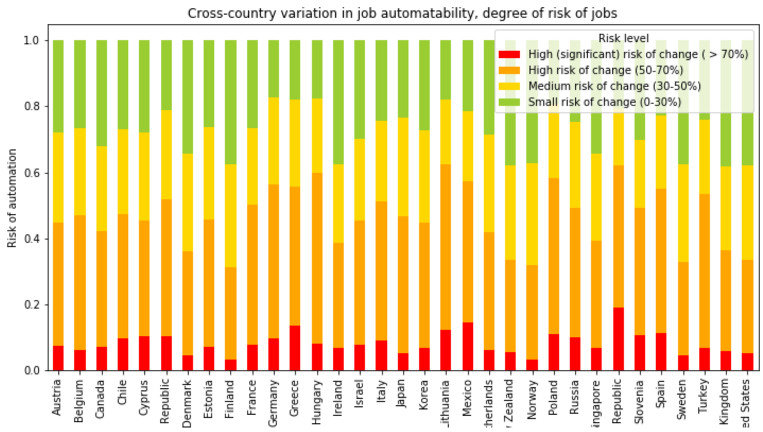

Figure 5.3: Employment shares by risk levels

NQ also predict that 14% of employment is at significantly high risk while this ratio is calculated 9% by AGZ for 21 OECD countries. AGZ found a lower ratio as they do not include countries with high automation risk such as Mexico, Turkey and Chile into their analysis. In addition, we find that 51% of the labor force is at high risk (between 50% and 70%) while 12% is at low risk of automation (>30%). NQ calculate different ratios: 32% for high risk and 26% for low risk. The reason can be that NQ do not include data of Hungary and Mexico, countries with a high concentration of risk at 60%.

Even though NQ and our study find different employment shares for different levels of risk due to including different countries, automation risk calculated with the task-based approach is consistent when several studies are compared. Let's take Germany as an example. 15% of individuals in the labor force is at significant risk according to our study while AGZ estimate 12%. Dengler and Matthes (2018) calculate also 15% risk for German employment by taking into account the variation of tasks and skills within the occupations. They use different datasets than PIAAC. The second example is the United States. While we calculate 10% for the US which is the same as NQ's, AGZ estimated 9% for the US. On the other hand, FO estimate that 47% of total US employment is at significantly high risk by adopting the occupation-based approach. The reason why the difference is extremely high between two approaches is that we take into account the variation of bottleneck-related attributes within the occupations. FO use 70 occupational labels to train the model while we use 2818 individuals with labels where some of the individuals hold the same occupation.

5.4 Occupations and industries at risk

In this section, automation risk is presented at the occupational and industry levels.[5] Automation risk for each individual is calculated with logistic regression as a function of skill use at work. Then, the risks are aggregated into the occupation and industry levels. This section answers which jobs and industries at high risk of automation.

5.4.1 Occupations

Table 5.4 illustrates how automation risk and skills vary across the ISCO occupational groups.[6] The risk declines when the frequency (values between 1 and 5) of using negatively (positively) correlated bottleneck-related skills increases (decreases). The columns are in shades of red if the risk is high and in shades of green if the risk is low. Simple problem solving and negotiation have no color because their coefficients are insignificant, so we do not discuss them. The labor force is fairly equally distributed among the occupational groups.

Managers and professionals (teaching, engineering, health, law or technology related jobs) have the lowest risk as they highly use bottleneck-related skills that are negatively associated with automation risk. Even though these groups have also red values, the variance of the skill questions influences the results. For instance, the coefficient magnitudes of communication and influencing are similar, approximately 0.16 (Signs are different). However, the difference between the maximum and minimum values among the occupation groups is 1.03 (4.53-3.50) for communication and 2.94 (4.08-1.14) for influencing. So, although communication has a positive sign, the values are quite similar across the occupational group. On the other hand, the mean automation risk is the highest for operators and assemblers at plants; skilled agricultural, fishery and forestry workers

[5] To differentiate the risk at the occupation and industry level, detailed occupation and industry codes are required. However, ISCO codes are not publicly available for 9 countries (Austria, Canada, Estonia, Finland, Germany, Ireland, Singapore, Sweden and the US) due to privacy reasons. Therefore, results are valid only for the remaining 24 countries. At the country level, on the other hand, the risk is presented for 33 countries. The reason is that the risk probabilities are calculated independently from the occupation codes: To calculate the logistic coefficients, we only use the 4-digit level labelled data (labels and labelled occupations' answers to skill question) from three representative countries. Then, we use these coefficients to calculate the probabilities for each individual in PIAAC. Therefore, a probability is assigned to even individuals with no occupation code (instead the answers are *valid skip, refuse to say etc.*)

[6] Most of the observation in PIAAC have 4-digit ISCO occupation codes. Yet, in this section, the risk is presented at the 1-digit and 2-digit levels for two reasons. First, the number of occupations in these levels are more manageable to visualize. Second, some occupations have only the second (or first) level ISCO codes. Let's start with discussing risk probabilities at the highest level (1-digit) as we mostly use this level to reach conclusions for the characteristics of workers.

Major groups	Finger dexterity	Simple problem solving	Complex problem solving	Teaching	Advising	Planning for others	Communication	Negotiation	Influencing	Selling	Presentation	Cooperation	Automation risk	Labor distribution
Plant and machine operators and assemblers	4.06	3.36	2.22	1.83	2.38	1.60	3.94	1.76	1.93	1.51	1.26	3.27	0.56	0.12
Skilled agricultural, fishery, and forestry workers	4.13	3.23	2.30	2.02	2.22	2.25	3.50	2.09	2.17	2.04	1.30	3.17	0.52	0.11
Craft and related trades workers	4.58	3.65	2.65	2.26	2.82	2.14	4.21	2.09	2.35	1.67	1.34	3.70	0.51	0.10
Clerks	3.51	3.83	2.75	2.18	3.21	2.18	4.28	2.70	2.79	2.06	1.48	3.31	0.49	0.09
Service and sales workers	3.72	3.64	2.34	2.22	3.48	2.13	4.07	2.58	3.20	3.10	1.48	3.53	0.49	0.10
Elementary occupations	3.96	3.74	2.64	2.56	3.26	2.32	4.15	2.53	2.94	2.08	1.76	3.46	0.45	0.10
Technicians and associate professionals	3.52	4.08	3.11	2.67	3.67	2.61	4.43	2.99	3.30	2.10	1.89	3.49	0.42	0.09
Armed Forces occupations	3.71	3.71	2.43	2.43	3.71	3.14	4.43	1.00	1.14	1.00	2.43	4.14	0.37	0.13
Professionals	3.48	4.21	3.28	3.30	3.88	2.83	4.34	3.00	3.66	1.71	2.63	3.21	0.33	0.08
Managers, senior officials and legislators	3.31	4.31	3.47	3.35	4.16	3.85	4.53	3.89	4.08	2.92	2.54	3.66	0.29	0.09

Table 5.4: Automation risk and skill averages by ISCO major groups for the labor force

(gardener, farmer worker, animal producer); and craft and related trade workers (house building, painting, electronics). These groups specifically use their hands, which is positively associated with the risk of automation.

Figure 5.4 shows the mean probability of automation by more detailed occupational categories (2-digit). We discuss this graph in three groups: occupations at the top, in the middle and at the bottom. The highest probability of automation risk belongs to two occupations from the elementary occupations group: agriculture, forestry and fishery laborers (59%); and cleaners and helpers (58%). Also, food preparation and refuse workers are in the elementary occupations category which is at risk higher than 50%. Skill requirements are expected to be little to none in this group. Another category with high to medium risk is machine operators and assemblers (drivers and mobile plant operators with 56% and assemblers with 55%, machine operators with 54%, metal and machine workers with 51% and electric-electronics workers 49%). This category includes most of the manufacturing jobs and usually, the interaction between workers and machines is high. Therefore, some training and technical skill use are required. Besides, personal care jobs (one of the elementary occupations) are in the middle with a 43% risk. While social interaction is high in caring, a high level of training or education is not necessary. On the other hand, most of the teaching, engineering and managerial jobs are populated at the bottom of the graph. These occupations are required a high level of education and training. Overall, the graph tells us that the risk of automation declines as the skill level increases. Here, the skills are not only the bottleneck-related skills in PIAAC but also occupation-oriented skills. Our findings are consistent with NQ's findings.

5.4.2 Industries

Industries at the descending risk order are shown in Figure 5.5[7]. The risk is the highest for activities of households as employers of domestic personnel such as maids, cooks, gardeners, caretakers and tutors (56%). Other service industries at high risk are land transport, postal and courier services, (undifferentiated) goods and service-producing activities, food and beverage services. However, other than these exceptions, industries with low risk belongs to the service sector (management, art, social work and human health-related activities). On the other hand, the majority of the industries at high to medium risk are primary (fishing, 53% and agriculture, 51%) and secondary (manufacturing with average 47%) sectors. The exceptions in the primary sector are mining activities and oil extraction with low risk at around 38%.

[7]Each individual in PIAAC is classified with the first two levels of ISIC code (International Standard Industrial Classification). There are 21 sections and 88 divisions. We prefer not to aggregate the risk into 21 sections as it is at a very high level. On the other hand, showing 88 industry groups would be too detailed. So, we select the top 20 and bottom 20 industries from the table that is sorted by the descending automation risk. Austria, Estonia and Finland do not have ISIC code in their PIAAC file so the results do not include these countries.

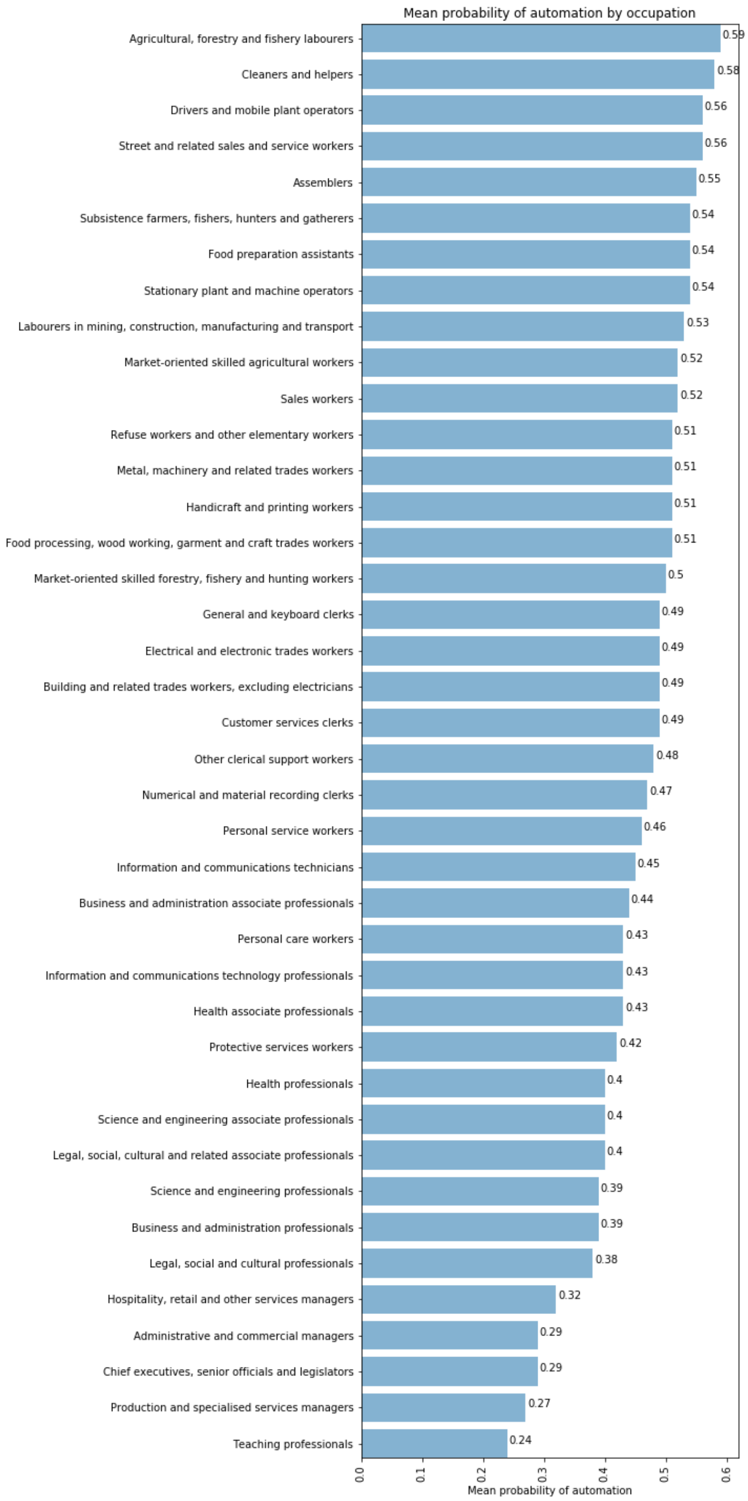

Figure 5.4: Automation risk by occupation

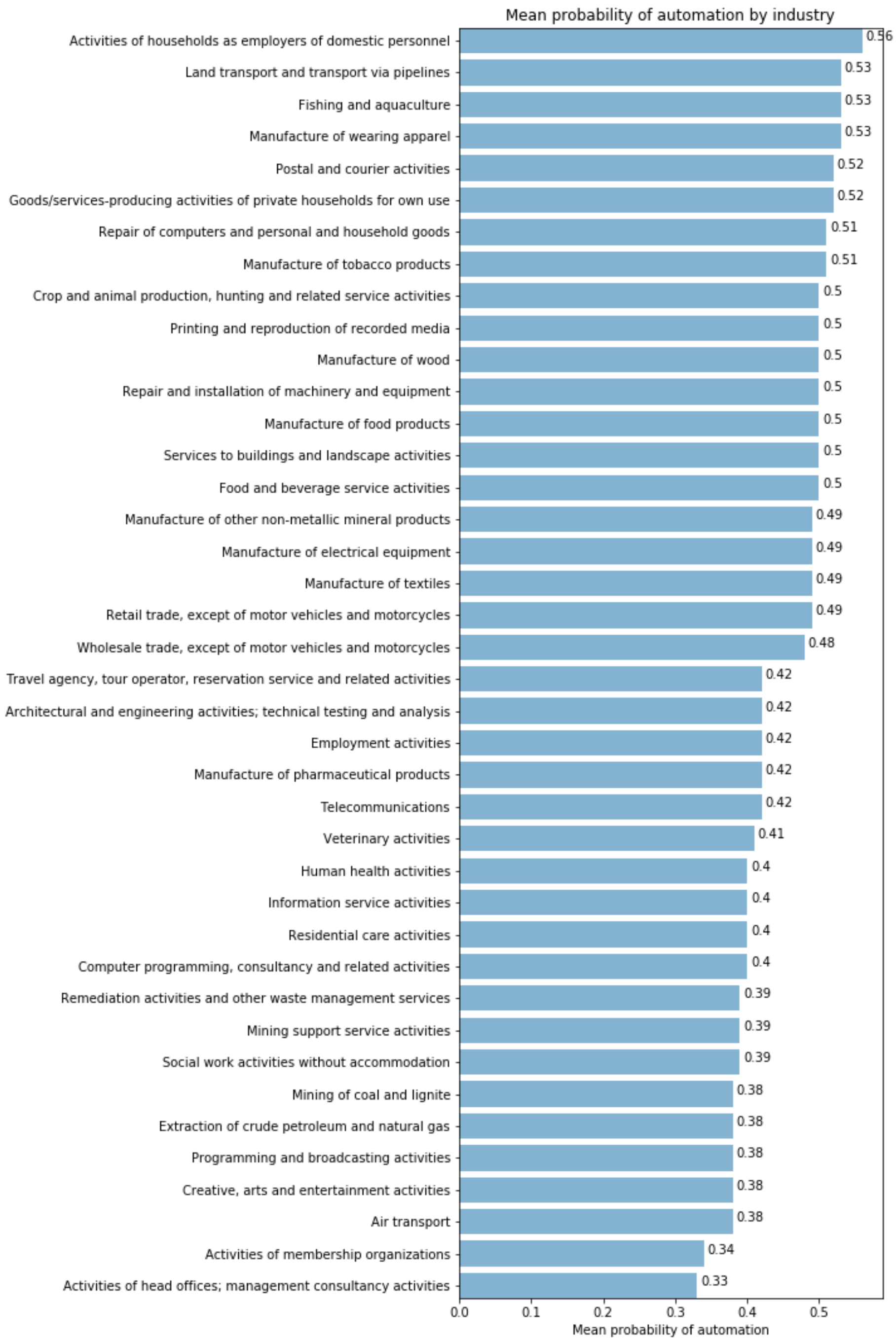

Figure 5.5: Automation risk by industry

5.5 Distribution of employee characteristics

5.5.1 Education level

The relationship between the educational attainment and automation risk is monotonic (Table 5.1) for all countries (Appendix C.1). The risk declines as the education degree increases. Among the nine levels, the first two levels of education are the introduction and completion of basic education (reading, writing, elementary understanding in mathematics) while the third level, upper secondary education, aims to teach work-related skills starting at the age of 15. The following levels are more advanced and can be academically based or occupation-specific (UNESCO Institute for Statistics, 2012). The first two levels are at the highest risk because no occupation-specific skills or knowledge are taught.

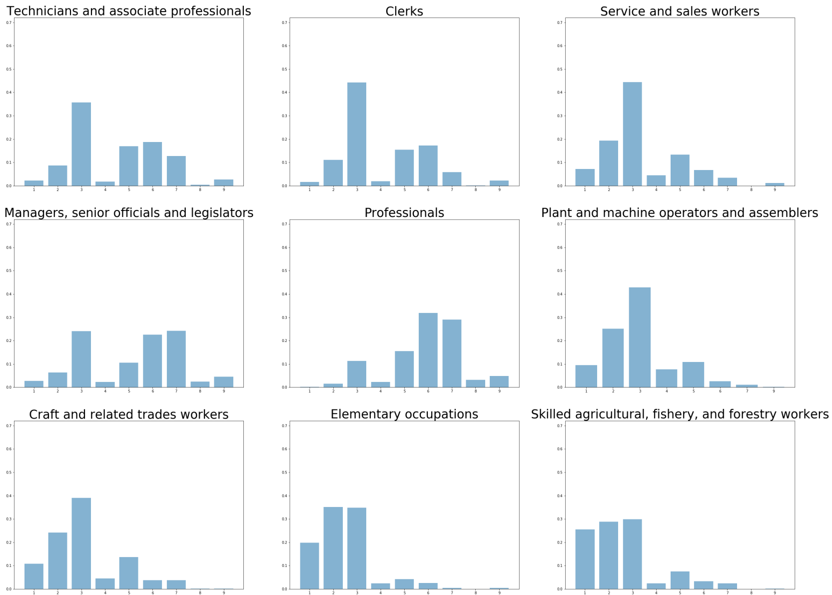

Figure 5.6: Ratio of people per education degree per major groups

In section 5.4.1, we divide occupations that are sorted in descending risk order into three groups: top (elementary occupations), middle (operators and assemblers at plants), bottom (managers and professionals). We argue that occupations on top are at the highest risk and do not require specific skills while the need for training and knowledge increase as going down. The ratio of people by education levels is illustrated in Figure 5.6. More than half of the workers operating either in the elementary occupations or skilled agriculture-related occupations completed only the basic education (until level 2) which does not provide job-specific knowledge. Following that, the highest ratio for upper secondary education (level 3) belongs to operators and assemblers at plants so, some training is required. Lastly, people with higher education levels are mostly working in the professionals or managerial related jobs which are at the lowest risk. Briefly, low-skilled workers are the highest risk and vulnerable to technological changes. Therefore, policies should focus on training and reskilling these people to adapt them to the changes in technology.

5.5.2 Gender

Female workers are more likely to be replaced by machines compared to their male colleagues in the majority of the countries. When we zoom into the occupations, we explore that the risk of automation is always higher for females regardless of the occupation category or skill level of the occupations. These findings are also consistent with the linear regression coefficients for females in Table 5.2. The coefficient signs for females are positive and when occupation and industry dummies are added in model 2, the risk increases further. Table 5.5 highlights the occupations that have higher than 70% and lower than 30% of the male population. We find that both genders can dominate occupations that require a different level of skills. For instance, female ratios are around 70% for high-skilled jobs like teaching and health-related jobs but also, for unskilled jobs like personal caring and cleaning jobs. The same is true for males: there are both high-skilled and unskilled occupations that male ratio is strikingly high. The male ratio is around 70 to 80% for the managerial and technology-related jobs while the ratio is high also for low-skilled jobs such as agricultural and manufacturing jobs. One of the main reasons for females being at higher risk is that females use less bottleneck-related skills at work compared to their male colleagues. This is a long-term societal issue. However, it can change with education or when females are more involved in the workforce.

On the other hand, when the impact factor ($risk \times size$) is considered, then the gender sizes make a difference. For instance, the manufacturing sector is male-dominant. Even though the risk is higher for females in this sector, automation of tasks will have a higher impact on male workers. Overall, policies should take into account the number of male/female workers at risk and also encourage female workers to take more initiative to use bottleneck-related skills at work.

Table 5.5: Occupations with higher than 70% and lower than 30% automation risk

Occupation	Risk for male	Risk for female	Male ratio
Teaching professionals	0.24	0.26	0.29
Production and specialised services managers	0.26	0.29	0.72
Chief executives, senior officials and legislators	0.3	0.32	0.73
Science and engineering professionals	0.37	0.46	0.71
Protective services workers	0.43	0.5	0.79
Personal care workers	0.37	0.43	0.13
Science and engineering associate professionals	0.37	0.44	0.82
Health professionals	0.38	0.39	0.26
Information and communications technology professionals	0.41	0.47	0.84
Health associate professionals	0.41	0.45	0.29
Information and communications technicians	0.42	0.48	0.78
Market-oriented skilled forestry, fishery and hunting workers	0.46	0.5	0.91
Electrical and electronic trades workers	0.51	0.63	0.94
Customer services clerks	0.45	0.52	0.3
Handicraft and printing workers	0.48	0.58	0.7
Building and related trades workers, excluding electricians	0.49	0.53	0.95
General and keyboard clerks	0.47	0.52	0.27
Metal, machinery and related trades workers	0.52	0.58	0.92
Market-oriented skilled agricultural workers	0.53	0.56	0.76
Assemblers	0.51	0.58	0.7
Labourers in mining, construction, manufacturing and transport	0.53	0.55	0.71
Drivers and mobile plant operators	0.58	0.58	0.94
Cleaners and helpers	0.54	0.58	0.2
Agricultural, forestry and fishery labourers	0.58	0.59	0.71

5.5.3 Age

At first glance, the relationship between age and automation risk appears to be U-shaped where younger and older age groups have higher risks (Figure 5.7). However, there are a few exception countries. Automation risks for Cyprus and Lithuania fluctuate but remain in a narrow interval. So, age does not affect the risk to a great extent compared to other countries. In Denmark and France, the risk does not change with age after a peak in younger ages. Other countries without a U-shaped distribution are Belgium, Czech Republic, Italy, Poland, Slovak Republic, Spain and Turkey where the risk decreases further with age.

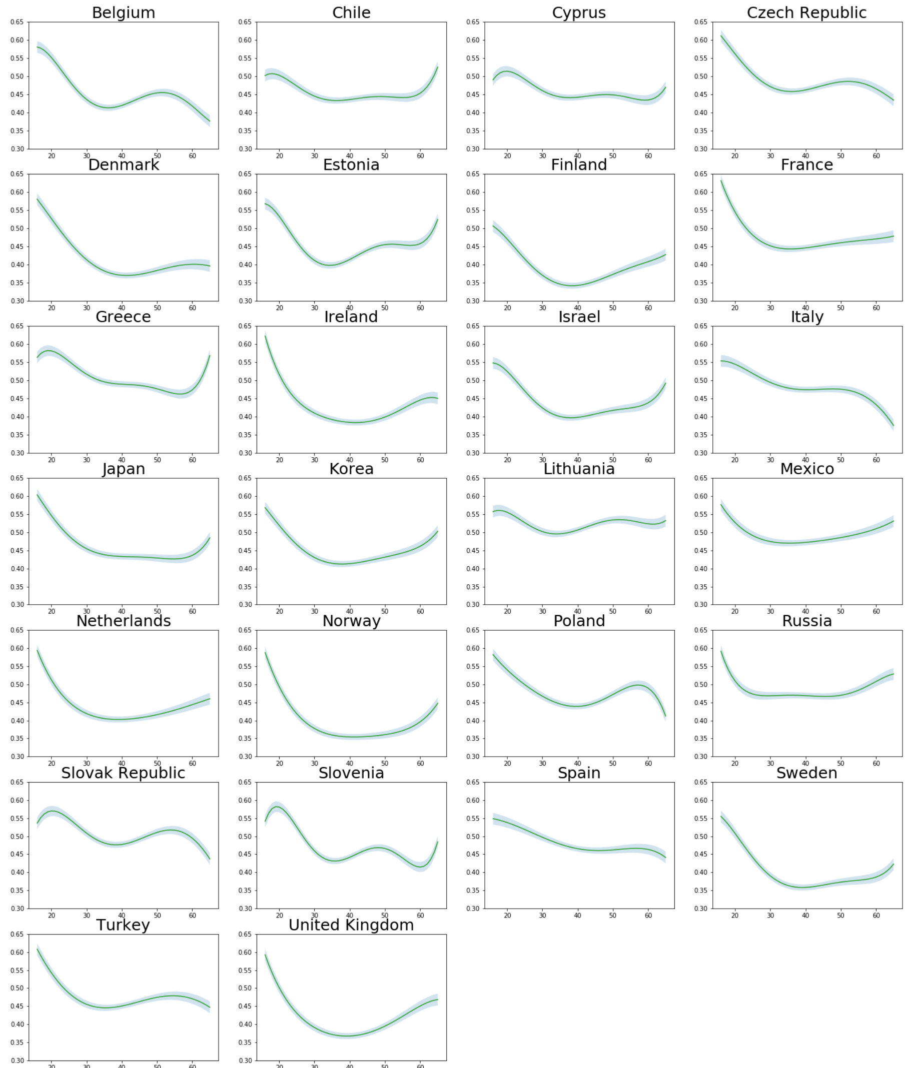

Figure 5.7: Age vs automation risk distribution per country

Studies suggest that the most productive years at work are between the ages of 30 and 40 (Lehman, 1953; Simonton, 1988). In parallel to that, when looked at Figure 5.7, the risk is the lowest between this age interval in most countries. On the other hand, younger populations, age 20 or younger, accounting for 7% of the total population, are at the highest risk. Table 5.6 shows how the young population is distributed across the occupational categories along with each categories' automation risks. Around 30% of the young population operates either as skilled agricultural workers or operators and assemblers at plants. The following occupation groups are craft workers and elementary occupations (cleaners, laborers and unskilled agricultural jobs) which together make 20% of the total young population. Around 60% of young workers operating in one of these four occupation groups have lower secondary education or just the basic education which means they have not earned any skills yet. Due to lack of skills, teenagers are at the highest risk.

According to NQ's findings, 34% of the teenage population is in sales and personal services while 20% is working in an elementary occupation. The difference occurs due to missing values in the age variable for 7 countries while NQ have a full dataset. We lose around 44,000 rows while calculating the population ratio and automation risk for the teenager and aged populations.

Table 5.6: Automation risk by ISCO major groups for the population younger than 20

Major groups	Automation risk	Young population %
Skilled agricultural, fishery, and forestry workers	0.60	0.16
Plant and machine operators and assemblers	0.59	0.14
Craft and related trades workers	0.58	0.10
Elementary occupations	0.57	0.09
Clerks	0.55	0.09
Service and sales workers	0.54	0.09
Technicians and associate professionals	0.45	0.11
Professionals	0.44	0.10
Managers, senior officials and legislators	0.37	0.11

In the majority of the countries, older populations, age 60 or older, are at the highest risk after the youngest populations. One of the reasons is that the average age of retirement [8] is calculated as 64.3 for men and 63.1 for women in 2013 (OECD, 2019). Older people who are close to retirement may be less willing to adapt to technological changes because they tend to use less technology-related skills and are less passionate about learning (Autor and Dorn, 2009). Therefore, training and reskilling strategies may be difficult for the older population. Table 5.7 illustrates automation risk and the population distribution for the groups that are older than 60. In most of the occupational categories, older people are at risk higher than 50%. On the other hand, the risk is the lowest in managerial and professional jobs (teaching, engineering, health, law or technology related

[8] The average age of retirement is the average age of exit from the labor force during a 5-year period.

works). While professional jobs mostly require a college degree, the red-colored occupations require an upper secondary degree or less. So, even though the older population is at high risk, the risk is not really applicable when a higher education degree exists.

Table 5.7: Automation risk by ISCO major groups for the population older than 60

Major groups	Automation risk	Older population %
Plant and machine operators and assemblers	0.55	0.15
Skilled agricultural, fishery, and forestry workers	0.55	0.12
Craft and related trades workers	0.52	0.14
Service and sales workers	0.50	0.15
Clerks	0.50	0.10
Elementary occupations	0.50	0.14
Technicians and associate professionals	0.41	0.07
Professionals	0.33	0.06
Managers, senior officials and legislators	0.31	0.09

5.5.4 Earnings

There is a monotonic relationship between earnings and automation risk like in the case of the level of education (Figure 5.8). So, the risk increases as the income declines. As the low-income and low-educated (low-skilled) workers are at the highest risk of automation, we can deduce that the low-skilled workers are at the low-income percentile (See also Table 3.4).

The variation in automation risk across countries does not change the monotonic relationship between the risk of automation and earning. Occupations with the lowest income are at the highest risk in a country irrespective from other countries' automation risk distributions. Let's take two countries with extreme mean risks: the Slovak Republic and Sweden. The risk varies between 40% to 60% for the Slovak Republic while the variation for Sweden is between 30% and 45%. While an individual with 45% of substitution risk in the Slovak Republic is at the highest income percentile, a Swedish person with the same risk is at the lowest income percentile. On the other hand, Russia has a flat distribution which is notably different than the other countries' distributions. Risk does not differ among Russian workers by the different levels of income. This is an interesting case because it means that earnings do not change with the changes in the skills used at work.

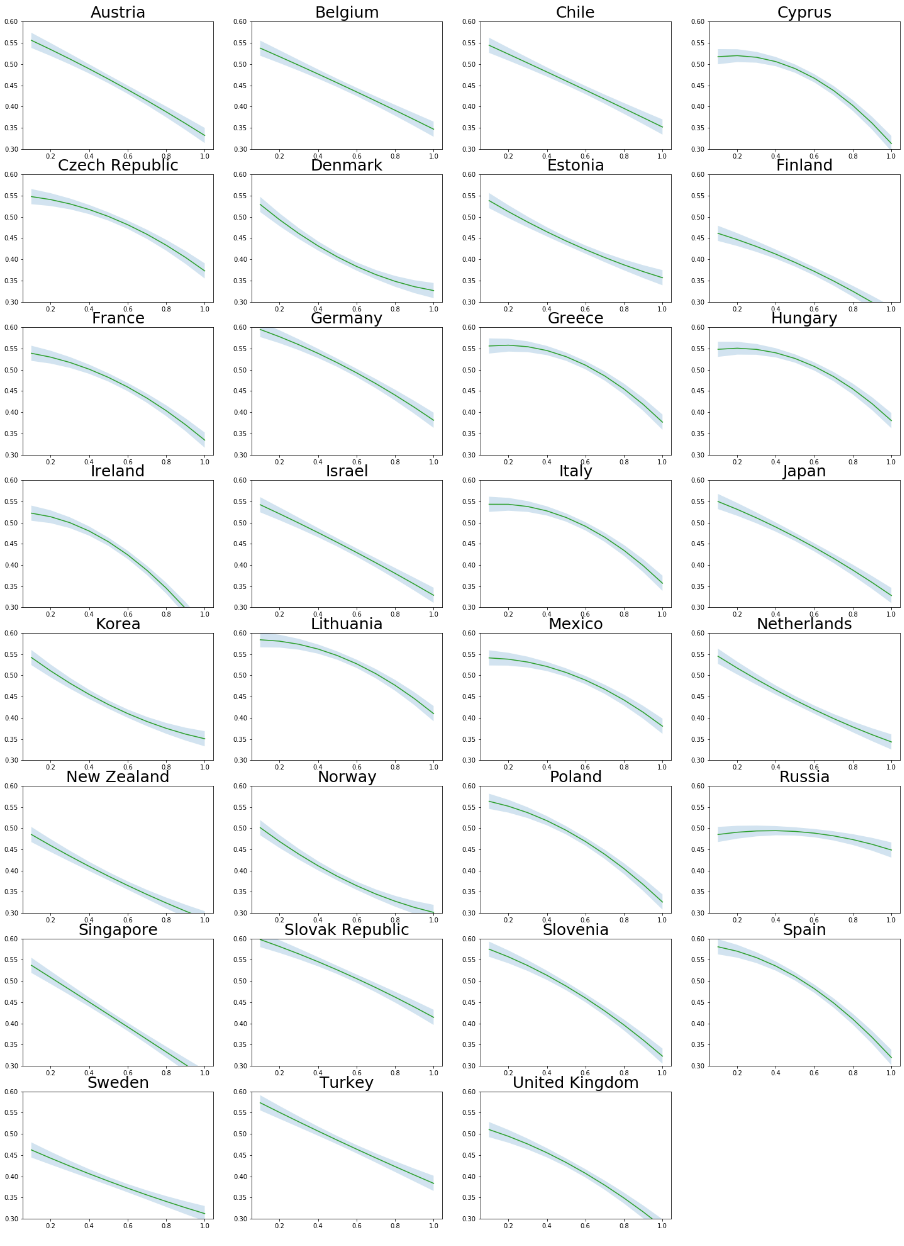

Figure 5.8: Earnings vs automation risk distribution per country

5.5.5 Other characteristics

Economic sector

Previously, we presented the risk of automation at the industry level. Now, we illustrate the risk of automation by the economic sector which is the higher level of aggregation (Figure 5.9). Still, some of the occupations can be both in the public and private sector. For instance, teachers can work either at a state school or a private school then, the economic sector will be different. The mean risk for non-profit organizations such as charities, professional associations or religious organisations is 35%, which is the lowest among other sectors. Usually, these organizations are based on human interaction, hence the low risk is expected. Yet, employees in this sector form only 3% of the population. On the other hand, the public sector including local governments and state schools, have a mean risk of 38% which is close to the risk of the non-profit organization. The highest mean risk belongs to the private sector (i.e companies) which accounts for 75% of the total population. Therefore, the mean risk for the private sector (46%) is very similar to the 33 counties average risk (45%). Also, the agricultural and manufacturing jobs belong to the private sector which also explains the high mean risk. So, the focus should be put on the employees in the private sector as the impact is the highest.

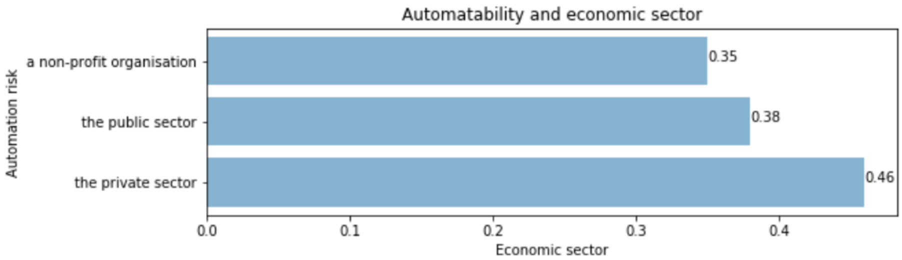

Figure 5.9: Economic sector vs automation risk

Contract type

Automation risk by the contract type is shown in Figure 5.10. Employees with indefinite contracts, which is 70% of the population, have the lowest mean risk (45%). Another 14% has a fixed-term contract with a comparatively low mean risk of 48%. On the other hand, 11% of the population that works without a contract has the highest mean risk (54%). Half of the employees without a contact falls into the service industry called "activities of households as employers". Child-care, cleaners and helpers, security guards are examples of these groups. These occupations do not require any job-related skills (also see Figure 5.6). Another 20% with no contract operates in the agriculture, forestry and fishing industry which again do not require more than upper secondary education.

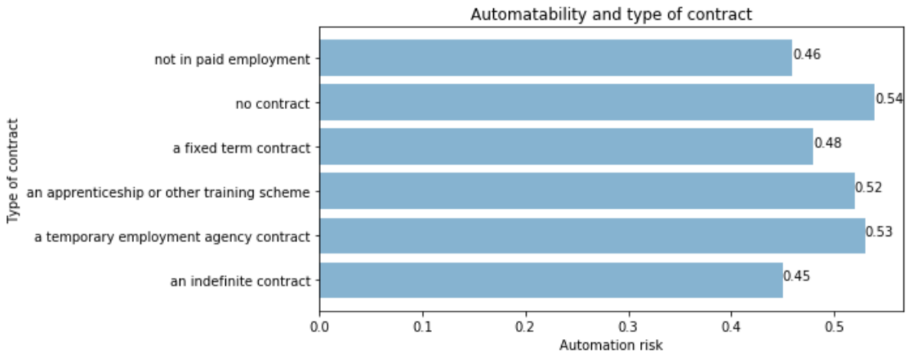

Figure 5.10: Contract type vs automation risk

Firm size

Finally, Figure 5.11 illustrates the firm sizes and corresponding mean automation risks. The risk decreases as the firm size increases. 60% of the population works in a place with 50 people or less and have the highest risk ratios (50% for 1 to 10 people and 44% for 11 to 50 people). Around 50% of the companies in agriculture, forestry and fishing; and accommodation and food service activities employ 10 or fewer people. On the other hand, 8% of the population works in a place with more than 1000 people. The top three sectors with large firm sizes are mining and quarrying by 30%, financial and insurance activities by 15% and information and communication by 14%. While the mean risk of these industries is around 42%, the risk declines to 39% for companies with more than 1000 employees. So, within an industry, larger firms are more likely to adapt to technological changes.

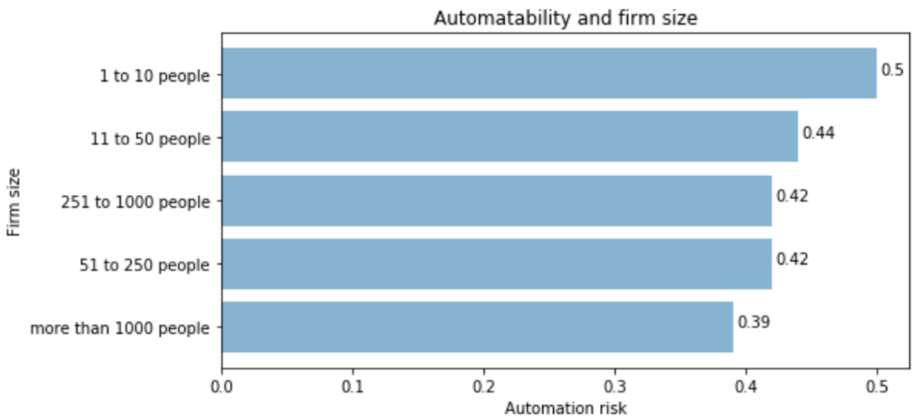

Figure 5.11: Firm size vs automation risk

5.6 Conclusion

The goals of this chapter were to explain the relationship between the skills used at work and automation risk; estimate the employment shares at high risk; assess the industries and occupations at risk; identify the socio-demographic characteristics of workers at risk.

The logistic regression results show that automation risk is low for jobs that require a high level of social interaction while the risk is high for jobs that require finger dexterity. Planning for others, presenting and influencing are the top three skills that reduce the risk the most. These skills are heavily used in professionals and managerial jobs which are at the lowest risk. Besides, the top three occupation groups that use their hands (machine operators; skill agriculture workers; craft workers) are highly expected to be replaced by machines. However, even though the risk is higher for certain occupations (and industries), the risk of automation varies across individuals within the same occupation (and industry). Besides, the mean share of employment at significantly high risk (> 70) is 14% for 33 countries. The findings are consistent with the other studies that also used the task-based approach. As opposed to FO's 47%, the risk was found to be more manageable.

The risk of automation decreases as the skill level (and the education level) of individuals increases. Jobs at the highest risk belong to the elementary and agriculture-related occupations where the skill requirements are little to none. The second group at risk is machine operators at plants where some skill is required as the interaction between human and machine is high. Finally, managerial-related jobs and professionals are at the lowest risk as these occupation groups require a high-level of education. For industries at risk, most of the jobs in the primary and secondary industries are at higher risk compared to the service sector jobs.

The vulnerable groups who are expected to be affected by the technological advancements the most across 33 countries are more likely to be less educated, low-income earners. In general, younger and older populations are at the highest risk of automation. While younger people mostly perform in unskilled jobs due to having only basic education, older people tend to be more technologically outdated. Females are always at higher risk compared to their male colleagues in an occupation. However, automation risk should not be interpreted independently from the impact factor ($risk \times size$). So, even though females are always at higher risk, the impact of risk also depends on the male/female ratio in an occupation. Besides, workers with no contract or a temporary contract have a higher risk of losing their jobs compared to ones with an indefinite contract. Finally, smaller firms within an industry are less likely to adapt to technological changes.

Chapter 6

Conclusion

6.1 Research summary

In recent years, concerns about the negative impacts of automation technologies (i.e. AI) on employment increased. Yet, how technologies will impact the labor force is still uncertain as it depends on the diffusion process of new technologies. Nevertheless, technological progress usually increases the total output generated and, the new output is redistributed among society. Eventually, net gain creates winners with more benefits and losers who are vulnerable to technological changes (Korinek, 2019). Policies are powerful tools to help those who are expected to be at high risk of automation and hence, at risk of losing their jobs. Therefore, we investigate who is at risk of automation and offer policy recommendations to reduce inequality and ensure the vulnerable groups are seen. This study has four main outcomes. First, to estimate the share of employment at risk of automation. Second, to estimate the jobs and industries at risk of automation. Third, to define the socio-demographic characteristics of workers at risk of automation. Finally, to provide policy recommendations to mitigate the negative impacts of automation on the vulnerable groups.

Regarding the first aim, the shares of employment at risk of automation has been calculated before in different studies. The occupation-based approach that is first proposed by Frey and Osborne (2013) (FO) is used to answer how susceptible jobs are to computerization, more specifically to find the share of the US employment at risk of automation. FO define three task groups called engineering bottlenecks that are not easy to automate in the coming decade or two. These task groups include nine (skills, abilities, knowledge) variables in total. Then, FO associate these nine variables with the expert assessment on the job automatability. This approach assumes that tasks do not vary within the same occupation so automation risk is the same for individuals who hold the same occupation. On the other hand, Arntz et al. (2016) (AGZ) argue that the risk is overestimated with the occupation-based approach since it does not take into account the fact that the tasks performed and skills used in an occupation may differ for each individual within the same occupation. So, AGZ propose the task-based approach to calculate the risk of automation for occupations.

Nedelkoska and Quintini (2018) (NQ) build their research on FO's study but consider that tasks and skills may vary across individuals within the same occupation. Besides, NQ analyse the socio-demographic characteristics of workers at risk which is also the second outcome. We, therefore, take the study of Nedelkoska and Quintini (2018) as the role model and then, improve their work. We consider only skills used at work while calculating automation risk for individuals because skills can be learned by a certain level of education and can be influenced by policies.

To summarize, our study aims to define jobs and workers at risk of automation and provide policy advises to reduce the possible adverse effects of the current technologies. To reach these aims, we associate engineering bottleneck-related skills with expert opinions on the jobs automatability. We formulate the main research question as follows:

Who is at risk of automation?

To answer the main question, we divide it into five sub-research questions. The first question is methodological and about how to improve NQ's work. The second, third and fourth questions are the results of the analysis and provide a baseline for the policy recommendations. The final fifth question is about the policy recommendations based on the analysis results and the literature review. Below the research questions are answered one by one.

Research question 1

Which country's dataset is the best to predict the share of jobs at risk of automation by using the task-based approach?

The fourth chapter of the thesis was dedicated to this question. The chapter included the data used and the method applied. We have already mentioned that NQ based their study on FO's study by adopting the task-based approach. While FO used O*NET dataset, NQ used PIAAC. These two datasets include different occupation code systems. Therefore, NQ had to find PIAAC equivalents of 70 O*NET labelled occupations and also 9 engineering bottleneck-related variables defined by FO. NQ did the matching manually. Then, they selected Canadian dataset to train the model and predict the risk of automation for jobs across all OECD countries. Before calculating the risk of automation for each individual across 33 countries, we did several improvements on NQ's method including the country selection:

1. Instead of matching 70 labelled occupations manually, we used the official correspondence table from the US Bureau of Labor Statistics which is available online. NQ's match was very different than this official table. Therefore, their estimated probabilities may not be valid. On the other hand, we had 80 corresponding PIAAC occupations as opposed to 79 occupations defined by NQ.

2. NQ defined 10 skills that are related to the engineering bottlenecks. There was no perfect match with the 9 O*NET variables as not all of them are skills. When the definition of social interaction, one of the engineering bottlenecks, is considered, *presentation* and *cooperation* are included in

the bottleneck-related skills. Adding them into the analysis improved the explanatory power of the analysis.

3. The only reason why NQ selected Canadian dataset to train the logistic regression model is that Canadian dataset has the greatest amount of observation in their data with more than 26,000 rows. We looked at the explanatory power of the model by training the model with different country datasets or the combination of country datasets. Finally, we concluded that the combination of New Zealand, Poland and the United States explain the model better compared to only the Canadian dataset.

Research question 2

What are the shares of employment at risk of automation in each OECD country?

The fifth chapter explains the results of the analysis in detail and provides answers to the second, third and fourth questions. After the risk of automation was calculated for each individual, the risk percentages were divided into four groups: the risk is significantly high if the percentage is higher than 70%, high if between 50% and 70%, medium if between 30% and 50% and small if the risk is smaller than 30%. The mean employment share for 33 countries was calculated as 14% (see also Appendix D.1 for a detailed table) which aligns with the results calculated with the task-based approach. The proportion was the same as NQ's finding. Besides, half of the jobs across countries have a high risk of automation while 23% of the occupations have medium risk. So, as opposed to the FO's calculation, 47% of jobs at high risk of automation, the risk was found to be less threatening and more manageable.

Research question 3

Which jobs and industries are at high risk of automation in OECD countries?

We found that the risk of automation declines as the frequency of social intelligence-related skills and the level of education increases. Most of the occupations that require low skills (or lower level of education) are at the highest risk: elementary jobs such as agriculture laborers (59%), cleaners and helpers (58%) and food preparation assistants (54%) and also skilled agriculture workers such as subsistence farmers, fishers, hunters and gatherers (54%). Occupations with medium-level risk are performed by the middle-skilled workers who are mostly employed at the manufacturing jobs or sales jobs. Finally, managerial jobs (around 30%) and professionals (around 40% for science-related professionals and 24% for teaching professionals) are at the lowest risk. These jobs require high-level of education (i.e. college degree) and also include high-level of bottleneck related skills such as teaching, advising and influencing. Besides, personal care workers which belong to the elementary job category has lower risk compared to manufacturing jobs as it requires high social interaction even though the level of education is low.

In terms of industries, most of the primary (fishing, 53% and agriculture, 51%) and secondary (manufacturing a with an average of 47%) industry jobs are at higher risk than the jobs in the service sector. There are exceptions: while some of the service sector jobs such as maids, cooks, gardeners, tutors are likely to be at higher risk compared to other service jobs, mining and oil extraction activities are at lower risk compared to other secondary jobs.

Research question 4

What are the socio-demographic characteristics of workers at risk of losing their jobs?

The education level has a positive impact on the level of skills and knowledge. Higher education leads to higher level of skills, therefore, reduces the risk of automation. So, the technological changes are skill-biased. Managers and professionals that are at the lowest risk have the highest proportion of tertiary education degree. On the other hand, most of the occupations at the highest risk fall into either elementary or agriculture-related jobs which require a low level of education. Also, within the same occupation, low-educated workers are more likely to be replaced by robots. As low-income earners are low-skilled workers, they are more susceptible to automation than the high-income earners (higher-skilled).

Females are always at higher risk compared to their male colleagues irrespective of the skill level of an occupation. So, females are not only dominant at the low-skilled occupations such as cleaning and personal care jobs but also in some high-skilled jobs such as health-related jobs. This is also true for males. The male ratio is much higher in managerial and technology-related jobs but also in manufacturing jobs. Besides, automation risk should not be interpreted independently from the impact factor ($risk \times size$). Even though females are always at higher risk, the impact of automation also depends on the male/female ratio in an occupation.

Usually, younger and older populations are at higher risk compared to workers age between 30 and 50. While the younger population mostly work at the low-skilled jobs due to having only basic education, older groups that are close to the retirement age are less likely to update themselves according to the technology trends.

Employees working in the private sector are at higher risk compared to those at public sector and non-profit organizations. One of the reasons is that most of the elementary and manufacturing jobs are in the private sector. Moreover, people without a contract or with a temporary contract are more likely to be replaced by the new technologies. Finally, smaller firms with less than 50 people are at higher risk compared to larger firms.

Overall, the groups that are at the highest risk of losing their jobs are more likely to be less educated, young, low-income earners working in a small firm in the private sector with no contract or a temporary contract.

6.2 Policy recommendations

This section aims to answer the fifth question. The analysis results and the literature review are combined to provide policy recommendations. Also, the recommendations are presented and discussed during the meetings with technology experts. The question is answered in three sub-sections: human capital incentives; tax and subsidies; reducing the monopoly power.

Research question 5

What kind of policies could reduce the social cost of rapid technological progress?

6.2.1 Human capital incentives

As technology is diffused among society and implemented into businesses; some tasks will newly emerge while some will be fully automated. So, workers will need to move between and within occupations according to tasks availability. McKinsey (2017) reports that around 75 to 375 million people in the labor force may shift to new occupational categories and learn new skills. In parallel, according to Bravo (2015), individuals are expected to be re-skilled rather than deskilled or up-skilled, meaning that knowledge will simultaneously increase in new tasks and decrease in other tasks rather than having only one-way change. As technological progress creates new demands for workers with different skill levels, investing in workers is required so that they can adapt to the changing environment. The human capital theory argues that investing in workers is a factor of production, just like investing in capital equipment. (Amadeo, 2020; Becker, 1994). So, teaching new skills and re-skilling workers are not only beneficial for workers but also firms that aim to maintain their profitability.

Bughin et al. (2018) report that the demand for technological skills is expected to increase the most by 2030 while time spent on social and emotional skills will be still the highest. Also, our analysis findings justify that people such as managers and teachers who highly use social and creativity-related skills are less likely to be replaced by machines. Yet, it is still crucial for workers to learn how to use occupation-related technologies as its need will increase the fastest. So, the education system should combine social skills and technological skills. However, adjusting the education system and harvesting the outcomes may take time (Nedelkoska and Quintini, 2018). Also, the education system targets the young population and the older population may not benefit as much (Agrawal et al., 2018). Therefore, training is a faster way to invest in human capital. Nedelkoska and Quintini (2018) highlight the importance of training to reduce the risk of automation: BIBB Employment Surveys for Germany ask respondents to report five consecutive training they followed and the tasks they performed afterwards. Even after the second training, it is observed that workers start to perform tasks that require more social intelligence and creativity, and as a result, the risk declines significantly.

Nedelkoska and Quintini (2018) define different actors involved in the training process: firms, workers and government. According to NQ's analysis findings, firms are less willing to provide on-the-job training to workers at higher risk of automation, especially if firms expect new technologies to *highly* automate the tasks of these workers and to be more profitable compared to them. Workers at the highest risk are three times less likely to receive on-the-job training compared to workers at the lowest risk. On the other hand, if tasks of a job can be *partially* automated in the future, then employers invest in their workers since technology is still labor complementary and also firms need workers to increase their profitability.

Overall, high-risk groups need retaining more than other workers, but they are less likely to get it by firms. While workers at the lowest risk spent on average of 59 hours in training in 2014, workers at the highest risk spent only 25 hours (Nedelkoska and Quintini, 2018). According to our findings, workers at high risk are usually less-educated, low-income earners. So, they may not have adequate savings to spend on training or enough motivation to start for a new re-skilling program. Also, considering that training programs are not as comprehensive as education, 25 hours of training is not sufficient to learn a new skill. While three to four years are required to get a college degree, training has limited time to provide the necessary knowledge. Then, government-sponsored programs should consider these limits. First, training should be sponsored by governments or the cost of training should be low enough that workers at risk can effort. Second, governments should encourage these workers to participate in training by increasing awareness about the risks of automation. Third, training programs should be comprehensive but also up-to-date enough to help high-risk workers in adapting to the technological changes or switching to a new job.

Besides providing training programs, another way to invest in workers is that governments conduct or sponsor researches that increase the distributive effect of technologies (Korinek, 2019). These technologies can be complementary for the low-skilled rather than a substitute option. For instance, virtual assistants or chatbots emulate human interaction to perform various tasks such as filling out forms and directing a request to the relevant department within governments or firms. These technologies complement humans and allow low-skilled workers to perform higher value-added tasks.

6.2.2 Tax and subsidies

While human capital incentives focus on workers and help them to keep up with the pace of technology, governments can also steer the technological progress from a more financial perspective. Depending on the effects of the distributive technologies, whether its negative or positive, different tax or subsidy schemes can be introduced (Korinek, 2019).

The ultimate goal of firms is to increase their profitability. If technology is more profitable than workers, firms prefer to invest in technology, even some of the implementations result in increasing unemployment. Therefore, innovations that increase unemployment can be discouraged through the means of tax policies. One of the much-discussed taxes is the robot tax. It is a concept indicating that firms should pay tax for replacing workers with robots, and then, this tax should be distributed among displaced workers (Silkin, 2019). While several robotics companies call this idea as an "innovation penalty" (Cousins, 2017), technology entrepreneurs such as Bill Gates and Elon Musk supports the robot tax concept (Clifford, 2016; French, 2017). According to Korinek and Stiglitz (2017), this tax could significantly increase the fiscal revenue without harming the investment incentives, if the supply of capital is inelastic enough.

On the other hand, subsidies that increase the desirability of labor over capital can be useful for the low-skilled. For instance, providing wage and hiring subsidies to lower-skilled workers decreases their costs to the firms and increases their demand. As a result, innovations become less likely to replace lower-skilled workers (Ernst et al., 2019; Korinek, 2019).

6.2.3 Reducing the monopoly power

The competition to get a job increasingly seems to be between human and artificial robots. As robots are created by the tech giants, this subsection discusses the necessity to reduce the monopoly power of firms as a way of steering technological progress and reducing inequality. Different from the previous policy recommendations, this one looks from a broader perspective rather than focusing on how to reduce the technological unemployment.

The driving factors of the current wave of technological progress are digitization and information goods. The information goods are intangible, non-rivalry products, meaning that it can be used without being consumed away (Ernst et al., 2019). For instance, a computer program as an information good can be developed, copied for many times at a small cost and used for many times. When private companies have ownership of these products, competitors cannot use it. Then, firms with information goods become natural monopolies or create high entry barriers. Consequently, they gain the power of charging consumers with high prices which leads consumers to demand less, or they provide free services in the exchange of intellectual property rights (Korinek, 2019).

Korinek (2019) indicates that the most efficient solution to reduce the monopoly power is publicly funding the innovation of information goods. So, the information goods could be distributed at a much lower price and everyone can easily reach the technologies. However, private companies are better at commercializing compared to publicly funded organizations. Therefore, private companies are still required to develop and commercialize new technologies. Then, governments need to let companies keep some of the monopoly powers by awarding them with intellectual property rights. One of the ways of reducing monopoly power is to weaken the intellectual property rights by extending consumers' rights on their data usage. The second way, on the other hand, is to introduce a tax on the technological rents of corporations, such as charging a licensing fee to cooperation for their publicly available technologies (Korinek, 2019).

To conclude, our study investigated who is at high risk of automation and provided some ideas to reduce the adverse effects of technology on the high-risk groups. While human capital incentives such as training focus on the adaptation of workers to technological changes, governments can also steer the technological progress with financial interventions such as introducing tax or subsidy policies. Also, we discussed the importance of reducing the monopoly power of tech giants to reduce inequality. Overall, the analysis results in combination with the policy recommendations give some valuable insights.

6.3 Reflection on societal and ethical relevance

In this study, we calculated the proportions of employment in OECD countries at significantly high risk of automation and the socio-demographic characteristics of these workers. Our study was based on the widespread debate started with Frey and Osborne (2013)'s question: How susceptible are jobs to computerization? We found that 14% of the total employment in 33 countries is at significantly high risk of automation. While this ratio can be accepted as *manageable*, we should keep in mind that the life and well-being of each worker in this 14% ratio are highly likely to be negatively affected. When we think about the economy or the development of a country, the most important factor is to grow as a society. Even though in the long-run, the employment levels have chance to balance with right policy actions, the effects will be more pronounced in the short-run, especially for the low-educated, low-income earners who are at the highest risk with the lowest chance of adapting to the technological changes. Carl Frey mentions in one of the interviews that "while cashiers are not yet demolishing self-service cash registers, technological progress is not safe for the anger of workers who cannot connect with the modern economy" (Witteman, 2019). The mentioned anger of workers at risk may cause unexpected political trends which harm society further and deepen inequality. Therefore, the crucial point is to *really* understand who is at risk and take *tangible* actions about how to deal with the risk. Then, the question arises: Who is responsible?

As we live in a complex environment where multiple actors with different objectives are involved in the policy-making process, there is not a simple answer about how to share the responsibility of steering the technological progress or helping the vulnerable groups during this process. However, as Korinek (2019) emphasises, the importance of steering technological progress is even higher when it comes to distributing the resources between humans and artificial robots. If robots are preferred over humans, technology companies may reap all the benefit. Then, how the society dynamics will change? Policy-makers have to consider to what extent they should include robots into the labor force in a way that robots complement workers. Also, new opportunities will emerge. So, defining the opportunities and how to use them for the vulnerable groups' benefit should be on the agenda of the decision-makers.

Besides the economic benefits of taking care of the vulnerable groups, it is also a moral responsibility. As a society, people should support each other instead of ignoring those left behind because it is the only way to develop. Therefore, governments should increase the awareness about the risks and opportunities of technologies on society. Firms, on the other hand, should take some additional responsibility to help workers at risk. Of course, the moral side of automation cannot be summarized in one paragraph as it is a different area of research. However, my personal opinion is that up-to-date education with equal accessibility in combination with decent governance that focuses on the opportunities of technological advancements would be a proper way to prepare for the future.

6.4 Limitations and future work

While predicting jobs and industries at risk of automation provides us with interesting insights, the study has several limitations. First, the predictions do not reflect the future 100% accurately. We only calculated the substitution risk and did not take into account that new jobs will emerge with the implementation of the new technologies. Even though new jobs or tasks are expected to emerge in the near future, it is not possible to know their impacts beforehand. As discussed earlier, the diffusion process may be slow (in the 80s, the time between the innovation and sales was 14.4 years (Gort and Klepper, 1982)) and today's some promising technologies may not be widely used for economic, social or legal reasons.

The second limitation is related to data. FO conducted their study in 2013 and the part of PIAAC data that we are interested in was collected between 2011 and 2012. As we used FO's labelled occupations and engineering bottlenecks and combine them with the PIAAC answers of skill use at work, our results are consistent. However, the findings may be outdated. If there has been an opportunity to conduct the expert assessment and the skill survey today, the labels and skill answers could have been different. Also, there are only a limited number of skill questions and since the data is static it prevents us from including new technology-related skills. Still, engineering bottleneck-related skills are not expected to be *fully* automated in the near future. The third limitation is the definition of risk. The risk is accepted as significantly high if it is higher than 70%. However, there is no solid explanation of why the threshold set as 70%.

All the limitations mentioned above can be considered as possible future work. For the first limitation, with a system dynamics model, a certain level of uncertainty can be included in predicting the risk of automation. To overcome the second limitation, the analysis can be replicated with a different survey or a survey can be conducted from scratch by adding new skills. The research area can be narrowed down to a country-, industry- or company-level to investigate the effects of automation in more detail. Besides, the analysis is reproducible as PIAAC is currently working on the second cycles of the survey. With the updated data, the analysis can provide more up-to-date results.

Bibliography

Acemoglu, D. (1998). Why do new technologies complement skills? directed technical change and wage inequality. *The Quarterly Journal of Economics*, *113*(4), 1055–1089. https://doi.org/10.1162/003355398555838

Acemoglu, D. (2002). Technical change, inequality, and the labor market. *Journal of Economic Literature*, *40*(1), 7–72. https://doi.org/10.1257/jel.40.1.7

Acemoglu, D. (2008). Structural change and economic growth. In *Introduction to modern economic growth* (pp. 814–814). Princeton University Press.

Acemoglu, D., & Autor, D. (2011). *Skills, tasks and technologies: Implications for employment and earnings*. National Bureau of Economic Research. https://economics.mit.edu/files/7006

Acemoglu, D., & Restrepo, P. (2017). The race between machine and man: Implications of technology for growth, factor shares and employment. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2781320

Acemoglu, D., & Restrepo, P. (2018). Low-skill and high-skill automation. *Journal of Human Capital*, *12*. https://doi.org/10.3386/w24119

Aghion, P. (2002). Schumpeterian growth theory and the dynamics of income inequality. *Econometrica*, *70*(3), 855–882. https://doi.org/10.1111/1468-0262.00312

Aghion, P., & Howitt, P. (1994). Growth and unemployment. *Review of Economic Studies*, *61*(3), 477–494. https://EconPapers.repec.org/RePEc:oup:restud:v:61:y:1994:i:3:p:477-494.

Agrawal, A., Gans, J., & Goldfarb, A. (2018). Economic policy for artificial intelligence. https://doi.org/10.3386/w24690

Agrawal, A., Gans, J., & Goldfarb, A. (2019). *The economics of artificial intelligence: An agenda*. The University of Chicago Press.

Amadeo, K. (2020). Human capital and how it shapes america's future. https://www.thebalance.com/human-capital-definition-examples-impact-4173516

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Antonczyk, D., Deleire, T., & Fitzenberger, B. (2010). Polarization and rising wage inequality: Comparing the u.s. and germany. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1589537

Arntz, M., Gregory, T., & Zierahn, U. (2016). The risk of automation for jobs in oecd countries. *OECD Social, Employment and Migration Working Papers*. https://doi.org/10.1787/5jlz9h56dvq7-en

Arntz, M., Gregory, T., & Zierahn, U. (2017). Revisiting the risk of automation. *Economics Letters*, *159*, 157–160. https://doi.org/10.1016/j.econlet.2017.07.001

Autor, D. (2015). Why are there still so many jobs? the history and future of workplace automation. *Journal of Economic Perspectives*, *29*(3), 3–30. https://doi.org/10.1257/jep.29.3.3

Autor, D., & Dorn, D. (2009). This job is 'getting old:' measuring changes in job opportunities using occupational age structure. *American Economic Review*. https://doi.org/10.3386/w14652

Autor, D., Katz, L., & Kearney, M. (2006). The polarization of the u.s. labor market. https://doi.org/10.3386/w11986

Autor, D., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, *118*(4), 1279–1333. https://doi.org/10.1162/003355303322552801

Autor, D. H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, *103*(5), 1553–1597. https://doi.org/10.1257/aer.103.5.1553

Autor, D. H., & Handel, M. J. (2013). Putting tasks to the test: Human capital, job tasks, and wages. *Journal of Labor Economics*, *31*(S1). https://doi.org/10.1086/669332

Autor, D. H., Katz, L. F., & Kearney, M. S. (2008). Trends in u.s. wage inequality: Revising the revisionists. *Review of Economics and Statistics*, *90*(2), 300–323. https://doi.org/10.1162/rest.90.2.300

Barfield, W. (2018). *Research handbook on the law of artificial intelligence*. Edward Elgar Publishing. https://books.google.nl/books?id=KqV-DwAAQBAJ&pg=PA649&redir_esc=y#v=onepage&q&f=false

Becker, G. S. (1994). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago Press. https://www.nber.org/books/beck94-1#:~:text=Education%2C%20Third%20Edition-,Human%20Capital%3A%20A%20Theoretical%20and%20Empirical%20Analysis%20with%20Special%20Reference,National%20Bureau%20of%20Economic%20Research

Belzer, J., & Kent, A. (2000). *Encyclopedia of computer science and technology*. Dekker.

Bessen, J. (2015). *Learning by doing the real connection between innovation, wages, and wealth*. Yale University Press.

Boden, M. A. (2005). *The creative mind: Myths and mechanisms*. Routledge.

Borjas, G., & Freeman, R. (2019). From immigrants to robots: The changing locus of substitutes for workers. *The Russell Sage Foundation Journal of the Social Sciences*, *5(5)*(22). https://doi.org/10.7758/RSF.2019.5.5.02

Boudarbat, B., Lemieux, T., & Riddell, W. C. (2003). Recent trends in wage inequality and the wage structure in canada.

Bowles, J. (2014). The computerisation of european jobs. https://www.bruegel.org/2014/07/the-computerisation-of-european-jobs/

Bravo, E. (2015). Deskilling, up-skilling or reskilling? the effects of automation in information systems context, In *Amcis*.

Brynjolfsson, E., & McAfee, A. (2012). *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy*. Digital Frontier Press.

Brynjolfsson, E., Rock, D., & Syverson, C. (2017). Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. https://doi.org/10.3386/w24001

Brzeski, C., & Burk, I. (2015). Die roboter kommen folgen der automatisierung für den deutschen arbeitsmarkt. INGDiBa Economic Research. https://ingwb.de/media/1398074/ing-diba-economic-research-die-roboter-kommen.pdf

Bughin, J., Hazan, E., Lund, S., Dahlström, P., Wiesinger, A., & Subramaniam, A. (2018). Skill shift: Automation and the future of the workforce. McKinsey amp; Company. https://www.mckinsey.com/featured-insights/future-of-work/skill-shift-automation-and-the-future-of-the-workforce

Busch Group. (2020). Busch vacuum technology. https://www.buschvacuum.com/ch/fr/vacuum4-0

Caines, C., Hoffmann, F., & Kambourov, G. (2017). Complex-task biased technological change and the labor market. *Review of Economic Dynamics*, *25*, 298–319. https://doi.org/10.1016/j.red.2017.01.008

Carbonero, F., Ernst, E., & Weber, E. (2018). https://www.ilo.org/wcmsp5/groups/public/---dgreports/---inst/documents/publication/wcms_648063.pdf

Card, D., & Dinardo, J. (2002). Skill-biased technological change and rising wage inequality: Some problems and puzzles. *Journal of Labor Economics*, *20*(4), 733–783. https://doi.org/10.1086/342055

Card, D., & Lemieux, T. (1994). Changing wage structure and black-white differentials among men and women: A longitudinal analysis. *The American Economic Review*, *84*. https://doi.org/10.3386/w4755

Chentouf, L., & Ernst, E. (2014). Work organisation and incentives. *Global and local economic review*, *18*, 103–135.

Chui, M., Mayika, J., & Miremadi, M. (2016). Four fundamentals of workplace automation. https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/four-fundamentals-of-workplace-automation

Clifford, C. (2016). Elon musk: Robots will take your jobs, government will have to pay your wage. CNBC. https://www.cnbc.com/2016/11/04/elon-musk-robots-will-take-your-jobs-government-will-have-to-pay-your-wage.html

Cope, D. (1989). Experiments in musical intelligence (emi): Non-linear linguistic-based composition. *Interface*, *18*(1-2), 117–139. https://doi.org/10.1080/09298218908570541

Cousins, S. (2017). Is a "robot tax" really an "innovation penalty"? TechCrunch. https://techcrunch.com/2017/04/22/save-the-robots-from-taxes/?guccounter=1

Dengler, K., & Matthes, B. (2018). The impacts of digital transformation on the labour market: Substitution potentials of occupations in germany. *Technological Forecasting and Social Change*, *137*, 304–316. https://doi.org/10.1016/j.techfore.2018.09.024

Dobbs, R., Manyika, J., & Woetzel, J. (2015). McKinsey amp; Company. https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/the-four-global-forces-breaking-all-the-trends

Dustmann, C., Ludsteck, J., & Schönberg, U. (2009). Revisiting the german wage structure*. *Quarterly Journal of Economics*, *124*(2), 843–881. https://doi.org/10.1162/qjec.2009.124.2.843

Ernst, E., Merola, R., & Samaan, D. (2019). Economics of artificial intelligence: Implications for the future of work. *IZA Journal of Labor Policy*, *9*(1). https://doi.org/10.2478/izajolp-2019-0004

Federal Ministry of Labour and Social Affairs. (2017). https://www.bmas.de/SharedDocs/Downloads/EN/PDF-Publikationen/a883-white-paper.pdf?__blob=publicationFile&v=3

Freeman, R. (2015). Who owns the robots rules the world. https://wol.iza.org/articles/who-owns-the-robots-rules-the-world/long

French, S. (2017). Bill gates says robots should pay taxes if they take your job. MarketWatch. https://secure.marketwatch.com/story/bill-gates-says-robots-should-pay-taxes-if-they-take-your-job-2017-02-17

Frey, C. B., & Osborne, M. (2013). The future of employment: How susceptible are jobs to computerization? https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, *114*, 254–280. https://doi.org/10.1016/j.techfore.2016.08.019

Goldin, C. D., & Katz, L. F. (2009). *The race between education and technology*. Harvard University Press.

Google. (2020). Classification: Roc curve and auc nbsp;—nbsp; machine learning crash course. Google. https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc

Goos, M., & Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in britain. *Review of Economics and Statistics*, *89*(1), 118–133. https://doi.org/10.1162/rest.89.1.118

Goos, M., Manning, A., & Salomons, A. (2009). Job polarization in europe. *American Economic Review*, *99*(2), 58–63. https://doi.org/10.1257/aer.99.2.58

Goos, M., Manning, A., & Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review*, *104*(8), 2509–2526. https://doi.org/10.1257/aer.104.8.2509

Gordon, R. J. (1990). The measurement of durable goods prices. https://doi.org/10.7208/chicago/9780226304601.001.0001

Gort, M., & Klepper, S. (1982). Time paths in the diffusion of product innovations. *The Economic Journal*, *92*(367), 630. https://doi.org/10.2307/2232554

Gosling, A., Machin, S., & Meghir, C. (2000). The changing distribution of male wages in the u.k. *The Review of Economic Studies*, *67*(4), 635–666. https://doi.org/10.1111/1467-937x.00148

Greenwood, J., & Yorukoglu, M. (1997). 1974. *Carnegie-Rochester Conference Series on Public Policy*, *46*(1), 49–95. https://ideas.repec.org/a/eee/crcspp/v46y1997ip49-95.html

Grundke, R., Marcolin, L., Nguyen, T. L. B., & Squicciarini, M. (2018). Which skills for the digital era? *OECD Science, Technology and Industry Working Papers*. https://doi.org/10.1787/9a9479b5-en

Halland, H., Lokanc, M., Nair, A., & Kannan, S. (2015). *The extractive industries sector: Essentials for economists, public finance professionals, and policy makers*.

Handel, M. J. (2012). Trends in job skill demands in oecd countries. *OECD Social, Employment and Migration Working Papers*. https://doi.org/10.1787/5k8zk8pcq6td-en

Hernandez, R. (2018). The fall of employment in the manufacturing sector : Monthly labor review. U.S. Bureau of Labor Statistics. https://www.bls.gov/opub/mlr/2018/beyond-bls/the-fall-of-employment-in-the-manufacturing-sector.htm

Hornstein, A., Krusell, P., & Violante, G. L. (2005). The effects of technical change on labor market inequalities. *Handbook of Economic Growth*, 1275–1370. https://doi.org/10.1016/s1574-0684(05)01020-8

ILO. (2012). Isco-08 part 1: Introductory and methodological notes. https://www.ilo.org/public/english/bureau/stat/isco/isco08/

Jorgenson, D. W., & Vu, K. M. (2016). The ict revolution, world economic growth, and policy issues. *Telecommunications Policy*, *40*(5), 383–397. https://doi.org/10.1016/j.telpol.2016.01.002

Kahn, J. (2020). Learn to love the bot: Understand a.i. logic before using it as a business tool. Fortune. https://web.archive.org/web/20200419222011/https://fortune.com/2019/09/26/ai-business-strategy-management/

Katz, L. F., & Murphy, K. M. (1992). Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics*, *107*(1), 35–78. https://doi.org/10.2307/2118323

Korinek, A. (2019). Labor in the age of automation and artificial intelligence. https://econfip.org/wp-content/uploads/2019/02/6.Labor-in-the-Age-of-Automation-and-Artificial-Intelligence.pdf

Korinek, A., & Stiglitz, J. (2017). Artificial intelligence and its implications for income distribution and unemployment. https://doi.org/10.3386/w24174

Krueger, A. (1993). How computers have changed the wage structure: Evidence from microdata, 1984–1989. *The Quarterly Journal of Economics*, *108*(1), 33–60. https://EconPapers.repec.org/RePEc:oup:qjecon:v:108:y:1993:i:1:p:33-60.

Lehman, H. C. (1953). *Age and achievement*. Princeton University Press.

Marks, G. (2020). Microsoft is replacing employees with ai...and other small business tech news. Forbes Magazine. https://www.forbes.com/sites/quickerbettertech/2020/06/07/microsoft-is-replacing-employees-with-aiand-other-small-business-tech-news/

Mazzucato, M. (2015). *The entrepreneurial state: Debunking public vs. private sector myths*. Anthem Press.

McKinsey. (2017). Jobs lost, jobs gained: Workforce transitions in a time of automation. McKinsey Global Institute. https://www.mckinsey.com/~/

media/McKinsey/Featured%20Insights/Future%20of%20Organizations/What%20the%20future%20of%20work%20will%20mean%20for%20jobs%20skills%20and%20wages/MGI-Jobs-Lost-Jobs-Gained-Report-December-6-2017.ashx

Michaels, G., Natraj, A., & Reenen, J. V. (2014). Has ict polarized skill demand? evidence from eleven countries over twenty-five years. *Review of Economics and Statistics*, *96*(1), 60–77. https://doi.org/10.1162/rest_a_00366

Milgrom, P., & Roberts, J. (1990). The economics of modern manufacturing: Technology, strategy, and organization. *American Economic Review*, *80*, 511–28.

Nedelkoska, L., Neffke, F., & Wiederhold, S. (2015). Skill mismatch and the costs of job displacement.

Nedelkoska, L., & Quintini, G. (2018). Automation, skills use and training. *OECD Social, Employment and Migration Working Papers*. https://doi.org/10.1787/2e2f4eea-en

Neufeind, M., OReilly, J., Ranft, F., & Petropoulos, G. (2019). The impact of artificial intelligence on employment. In *Work in the digital age: Challenges of the fourth industrial revolution* (pp. 119–132). Rowman amp; Littlefield International.

OECD. (2013). *Technical report of the survey of adult skills (piaac)*. https://www.oecd.org/skills/piaac/_Technical%20Report_17OCT13.pdf

OECD. (2017). Survey of adult skills (piaac) - piaac, the oecd's programme of assessment and analysis of adult skills. https://www.oecd.org/skills/piaac/

OECD. (2019). Pensions at a glance 2019: Oecd and g20 indicators: En. http://www.oecd.org/els/public-pensions/oecd-pensions-at-a-glance-19991363.htm

O*NET Resource Center. (2020). About o*net. https://www.onetcenter.org/overview.html

Pajarinen, M., & Rouvinen, P. (2014). Computerization threatens one third of finnish employment. https://www.etla.fi/wp-content/uploads/ETLA-Muistio-Brief-22.pdf

Parasuraman, R., Sheridan, T., & Wickens, C. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *30*(3), 286–297. https://doi.org/10.1109/3468.844354

Phelps, E. S., & Nelson, R. R. (1966). Investment in humans, technological diffusion, and economic growth. *The American Economic Review*, *56*, 69–75. https://doi.org/10.1016/b978-0-12-554002-5.50015-7

Phua, C., Lee, V. C. S., Smith-Miles, K., & Gayler, R. W. (2010). (pdf) a comprehensive survey of data mining-based fraud detection research. https://www.researchgate.net/publication/46887451_A_Comprehensive_Survey_of_Data_Mining-based_Fraud_Detection_Research

Pouspourika, K. (2020). The 4 industrial revolutions. https://ied.eu/project-updates/the-4-industrial-revolutions/

Rousseau, P. L. (2010). Biased and unbiased technological change. *Economic Growth*, 5–8. https://doi.org/10.1057/9780230280823_2

Silkin, L. (2019). Robot tax: The pros and cons of taxing tech - future of work hub. https://www.futureofworkhub.info/comment/2019/12/4/

robot-tax-the-pros-and-cons-of-taxing-robotic-technology-in-the-workplace
Simonton, D. K. (1988). Age and outstanding achievement: What do we know after a century of research? *Psychological Bulletin*, *104*(2), 251–267. https://doi.org/10.1037/0033-2909.104.2.251
Söderbom, M. (2011). Lecture 3: Regression and causality. https://www.soderbom.net/metrix2/lec3.pdf
Spitz-Oener, A. (2006). Technical change, job tasks, and rising educational demands: Looking outside the wage structure. *Journal of Labor Economics*, *24*(2), 235–270. https://doi.org/10.1086/499972
Thaler, R. (2015). Who's afraid of artificial intelligence? https://www.edge.org/response-detail/26083
The Economist. (2016). The Economist Newspaper. https://www.economist.com/special-report/2016/06/23/automation-and-anxiety
Thompson, N. (2019). How ai and machine learning can help combat unconscious bias in humans. https://www.power-technology.com/research-reports/sponsored/how-ai-and-machine-learning-can-help-combat-unconscious-bias-in-humans/
UNESCO Institute for Statistics. (2012). *International standard classification of education: Isced 2011*. UNESCO Institute for Statistics. http://uis.unesco.org/en/topic/international-standard-classification-education-isced
United Nations. (2003). Investment and technology policies for competitiveness: Review of successful country experiences. https://unctad.org/en/Docs/iteipc20032_en.pdf
U.S. Bureau of Labor Statistics. (2018). Standard occupational classification. https://www.bls.gov/soc/
Violante, G. L. (2008). Skill-biased technical change. *The New Palgrave Dictionary of Economics, 2012 Version*. https://doi.org/10.1057/9781137336583.1655
Winick, E. (2020). Self-driving cars endanger nearly four million jobs but could create a $7 trillion industry. MIT Technology Review. https://www.technologyreview.com/2017/12/11/3745/self-driving-cars-endanger-nearly-four-million-jobs-but-could-create-a-7-trillion-industry/
Witteman, J. (2019). Economic historian carl frey: 'discontent about technology can become a new breeding ground for populists'. https://www.volkskrant.nl/economie/economisch-historicus-carl-frey-onvrede-over-technologie-kan-nieuwe-voedingsbodem-voor-populisten-worden~bd046751/?referer=https%3A%2F%2Fwww.google.com%2F
Zuboff, S. (1985). Automatefin-fonnate: The two faces of intelligent technology. *Organizational Dynamics*, *14*(2), 5–18. https://doi.org/10.1016/0090-2616(85)90033-6

List of Figures

1.1 The research structure 3

2.1 The outcomes of the sub-research questions 7

3.1 Waves of technological advancements (Busch Group, 2020) 9
3.2 US Employment in Agriculture, Manufacturing and Services (Acemoglu, 2008) 10
3.3 US Employment by occupation (Handel, 2012) 11
3.4 Changes in US employment by occupational skill percentile 1979-2007 (Acemoglu and Autor, 2011) 14
3.5 Distribution of hourly wage growth for routine and non-routine occupations in the US (Autor and Dorn, 2013) 15
3.6 The distribution of US employment by risk levels (Frey and Osborne, 2013) 21

4.1 The content model of O*NET (O*NET Resource Center, 2020) . 28
4.2 Rounds for the first cycle of PIAAC (OECD, 2017) 29
4.3 Hierarchical levels in the SOC system 31
4.4 Hierarchical levels in the ISCO-08 system 31
4.5 ROC curve (Google, 2020) 36
4.6 Confusion matrix for the selected representative countries 38
4.7 ROC for the selected representative countries 39
4.8 Detailed flow of the research 42

5.1 Partial correlation between the level of education and automation risk 47
5.2 Distribution of automation risk by country 50
5.3 Employment shares by risk levels 51
5.4 Automation risk by occupation 55
5.5 Automation risk by industry 56
5.6 Ratio of people per education degree per major groups 57
5.7 Age vs automation risk distribution per country 60
5.8 Earnings vs automation risk distribution per country 63
5.9 Economic sector vs automation risk 64
5.10 Contract type vs automation risk 65
5.11 Firm size vs automation risk 65

C.1 Educational attainment vs automation risk distribution per country 107

List of Tables

3.1 Decomposition of task shifts into between and within industry components in the US (Autor, Levy, et al., 2003) 12
3.2 Engineering bottleneck related O*NET variables (Frey and Osborne, 2013) 22
3.3 Engineering bottleneck related PIAAC variables (Nedelkoska and Quintini, 2018) 24

4.1 Occupations that are in NQ's match table but not in the actual PIAAC with this ISCO-08 occupational name 33
4.2 Confusion matrix 36
4.3 Performance of logistic regression with various countries; best performance in bold 37
4.4 Independent variables in PIAAC 40

5.1 Logistic regression results on automation risk as a function of bottleneck-related skills in PIAAC 44
5.2 Linear regression results on automation risk as a function of socio-demographic characteristics of workers 46
5.3 Country-level automation risk 48
5.4 Automation risk and skill averages by ISCO major groups for the labor force 53
5.5 Occupations with higher than 70% and lower than 30% automation risk 59
5.6 Automation risk by ISCO major groups for the population younger than 20 61
5.7 Automation risk by ISCO major groups for the population older than 60 62

A.1 SOC-ISCO Correspondence Table 91

B.1 Predicted risk of automation for ISCO coded occupations 106

D.1 Employment shares per risk levels by country 110

Appendix A

SOC-ISCO correspondence table for the labelled occupations

2010 SOC Code	2010 SOC Title	ISCO-08 Code	ISCO-08 Title EN	Label
11-1011	Chief Executives	1112	Senior government officials	0
11-1011	Chief Executives	1113	Traditional chiefs and heads of villages	0
11-1011	Chief Executives	1120	Managing directors and chief executives	0
11-3071	Transportation, Storage, and Distribution Managers	1324	Supply, distribution and related managers	0
11-9031	Education Administrators, Preschool and Childcare Center/Program	1341	Child care services managers	0
11-9151	Social and Community Service Managers	1344	Social welfare managers	0
13-1031	Claims Adjusters, Examiners, and Investigators	3315	Valuers and loss assessors	1
13-1041	Compliance Officers	3351	Customs and border inspectors	0
13-1041	Compliance Officers	3353	Government social benefits officials	0
13-1041	Compliance Officers	3354	Government licensing officials	0
13-1051	Cost Estimators	3339	Business services agents not elsewhere classified	1
13-1074	Farm Labor Contractors	3333	Employment agents and contractors	1
13-1121	Meeting, Convention, and Event Planners	3332	Conference and event planners	0
13-1161	Market Research Analysts and Marketing Specialists	2431	Advertising and marketing professionals	1
13-2011	Accountants and Auditors	2411	Accountants	1
13-2041	Credit Analysts	2413	Financial analysts	1
13-2053	Insurance Underwriters	3321	Insurance representatives	1

13-2072	Loan Officers	3312	Credit and loans officers	1
13-2081	Tax Examiners and Collectors, and Revenue Agents	3352	Government tax and excise officials	1
17-1012	Landscape Architects	2162	Landscape architects	0
17-1022	Surveyors	2165	Cartographers and surveyors	1
17-2051	Civil Engineers	2142	Civil engineers	0
17-2071	Electrical Engineers	2151	Electrical engineers	0
17-3012	Electrical and Electronics Drafters	3118	Draughtspersons	1
17-3022	Civil Engineering Technicians	3112	Civil engineering technicians	1
19-1023	Zoologists and Wildlife Biologists	2131	Biologists, botanists, zoologists and related professionals	0
19-2012	Physicists	2111	Physicists and astronomers	0
19-3011	Economists	2631	Economists	0
21-1011	Substance Abuse and Behavioral Disorder Counselors	2635	Social work and counselling professionals	0
21-1013	Marriage and Family Therapists	2635	Social work and counselling professionals	0
21-2011	Clergy	2636	Religious professionals	0
23-1011	Lawyers	2611	Lawyers	0
23-1012	Judicial Law Clerks	3411	Legal and related associate professionals	1
23-1023	Judges, Magistrate Judges, and Magistrates	2612	Judges	0
23-2011	Paralegals and Legal Assistants	3411	Legal and related associate professionals	1
25-2011	Preschool Teachers, Except Special Education	2342	Early childhood educators	0
27-1022	Fashion Designers	2163	Product and garment designers	0
27-2021	Athletes and Sports Competitors	3421	Athletes and sports players	0
27-3042	Technical Writers	2641	Authors and related writers	1
29-1021	Dentists, General	2261	Dentists	0
29-1069	Physicians and Surgeons, All Other	2211	Generalist medical practitioners	0
29-1069	Physicians and Surgeons, All Other	2212	Specialist medical practitioners	0
29-1141	Registered Nurses	2221	Nursing professionals	0
29-9099	Healthcare Practitioners and Technical Workers, All Other	3222	Midwifery associate professionals	0
29-9099	Healthcare Practitioners and Technical Workers, All Other	3230	Traditional and complementary medicine associate professionals	0
29-9099	Healthcare Practitioners and Technical Workers, All Other	3259	Health associate professionals not elsewhere classified	0
31-9094	Medical Transcriptionists	3344	Medical secretaries	1
35-1011	Chefs and Head Cooks	3434	Chefs	0
35-2011	Cooks, Fast Food	9411	Fast food preparers	1
35-3031	Waiters and Waitresses	5131	Waiters	0
35-9021	Dishwashers	9412	Kitchen helpers	1
37-2012	Maids and Housekeeping Cleaners	9111	Domestic cleaners and helpers	0
37-2012	Maids and Housekeeping Cleaners	9112	Cleaners and helpers in offices, hotels and other establishments	0

39-3011	Gaming Dealers	4212	Bookmakers, croupiers and related gaming workers	1
39-5012	Hairdressers, Hairstylists, and Cosmetologists	5141	Hairdressers	0
39-5012	Hairdressers, Hairstylists, and Cosmetologists	5142	Beauticians and related workers	0
39-6012	Concierges	4224	Hotel receptionists	0
39-9011	Childcare Workers	5311	Child care workers	0
41-2011	Cashiers	5230	Cashiers and ticket clerks	1
43-2011	Switchboard Operators, Including Answering Service	4223	Telephone switchboard operators	1
43-4041	Credit Authorizers, Checkers, and Clerks	4312	Statistical, finance and insurance clerks	1
43-4071	File Clerks	4415	Filing and copying clerks	1
43-4161	Human Resources Assistants, Except Payroll and Timekeeping	4416	Personnel clerks	1
43-5021	Couriers and Messengers	4412	Mail carriers and sorting clerks	1
43-5021	Couriers and Messengers	8321	Motorcycle drivers	1
43-5021	Couriers and Messengers	9331	Hand and pedal vehicle drivers	1
43-5021	Couriers and Messengers	9621	Messengers, package deliverers and luggage porters	1
43-5041	Meter Readers, Utilities	9623	Meter readers and vending-machine collectors	1
43-9021	Data Entry Keyers	4132	Data entry clerks	1
45-3021	Hunters and Trappers	6224	Hunters and trappers	0
45-3021	Hunters and Trappers	6340	Subsistence fishers, hunters, trappers and gatherers	0
47-2152	Plumbers, Pipefitters, and Steamfitters	7126	Plumbers and pipe fitters	0
47-2211	Sheet Metal Workers	7213	Sheet-metal workers	1
51-2022	Electrical and Electronic Equipment Assemblers	8212	Electrical and electronic equipment assemblers	1
51-4011	Computer-Controlled Machine Tool Operators, Metal and Plastic	7223	Metal working machine tool setters and operators	1
51-6031	Sewing Machine Operators	8153	Sewing machine operators	1
53-2031	Flight Attendants	5111	Travel attendants and travel stewards	0
53-3021	Bus Drivers, Transit and Intercity	8331	Bus and tram drivers	1
53-3033	Light Truck or Delivery Services Drivers	8322	Car, taxi and van drivers	1
53-3041	Taxi Drivers and Chauffeurs	8322	Car, taxi and van drivers	1
53-5022	Motorboat Operators	8350	Ships' deck crews and related workers	1
53-6021	Parking Lot Attendants	9629	Elementary workers not elsewhere classified	1
53-7051	Industrial Truck and Tractor Operators	8344	Lifting truck operators	1

Table A.1: SOC-ISCO Correspondence Table

Appendix B

Occupations and predicted probabilities

Probability	Label	ISCO08 code	Occupation	Weight
0.13		74	electrical and electronic trades workers	772
0.16		233	secondary education teachers	5830
0.17		232	vocational education teachers	2418
0.18		234	primary school and early childhood teachers	1638
0.18		1	managers	1287
0.19		7541	underwater divers	2580
0.20		1345	education managers	4330
0.20		1343	aged care services managers	2184
0.21		3155	air traffic safety electronics technicians	2827
0.21		2320	vocational education teachers	3633
0.21		111	legislators and senior officials	605
0.21		2330	secondary education teachers	3221
0.21		11	chief executives, senior officials and legislators	402
0.22		2356	information technology trainers	1913
0.22	0	1341	child care services managers	1626
0.23	0	2636	religious professionals	3979
0.23		2310	university and higher education teachers	3074
0.23		323	traditional and complementary medicine associate professionals	542
0.24		3422	sports coaches, instructors and officials	3799
0.24		2341	primary school teachers	2876
0.24		110	commissioned armed forces officers	2811
0.24		122	sales, marketing and development managers	1120
0.25	0	1344	social welfare managers	2373
0.25		132	manufacturing, mining, construction, and distribution managers	681
0.25		1342	health services managers	3244
0.25		235	other teaching professionals	3092
0.25		1223	research and development managers	4360

0.25		23	teaching professionals	1358
0.25		1321	manufacturing managers	3835
0.25		1212	human resource managers	5347
0.25		134	professional services managers	1851
0.25		1322	mining managers	1055
0.26		133	information and communications technology service managers	3957
0.26		2355	other arts teachers	3331
0.26		2353	other language teachers	3363
0.26		2352	special needs teachers	2367
0.26		1349	professional services managers not elsewhere classified	3559
0.26		1213	policy and planning managers	4036
0.27		322	nursing and midwifery associate professionals	573
0.27		51	personal service workers	484
0.27		3121	mining supervisors	2347
0.27		231	university and higher education teachers	3519
0.27		5165	driving instructors	4945
0.27		222	nursing and midwifery professionals	550
0.27		315	ship and aircraft controllers and technicians	548
0.27		1323	construction managers	4152
0.27		1431	sports, recreation and cultural centre managers	3125
0.27		112	managing directors and chief executives	9059
0.28		121	business services and administration managers	3376
0.28		2424	training and staff development professionals	2470
0.28		342	sports and fitness workers	4542
0.28		1114	senior officials of special-interest organizations	6610
0.28		3123	construction supervisors	4891
0.28		2359	teaching professionals not elsewhere classified	4258
0.28		2354	other music teachers	2726
0.28		511	travel attendants, conductors and guides	1636
0.28		1330	information and communications technology service managers	4285
0.29	0	1324	supply, distribution and related managers	2756
0.29		12	administrative and commercial managers	9253
0.29		3253	community health workers	3704
0.29		1111	legislators	2203
0.29		2351	education methods specialists	1686
0.29		210	non-commissioned armed forces officers	7342
0.29		1439	services managers not elsewhere classified	3599
0.29		1346	financial and insurance services branch managers	4298
0.29		221	medical doctors	2790
0.29		3122	manufacturing supervisors	4832
0.29		1219	business services and administration managers not elsewhere classified	4147
0.30	0	2342	early childhood educators	3754
0.30		21	science and engineering professionals	1375
0.30		3423	fitness and recreation instructors and program leaders	3345
0.30		1311	agricultural and forestry production managers	3185
0.30		141	hotel and restaurant managers	3216
0.30		2654	film, stage and related directors and producers	3791

0.30		2656	announcers on radio, television and other media	2392
0.30	0	1120	managing directors and chief executives	3067
0.30		96	refuse workers and other elementary workers	8926
0.31	0	2212	specialist medical practitioners	3819
0.31		2	professionals	5110
0.31	0	2211	generalist medical practitioners	3611
0.31		3152	ships' deck officers and pilots	2584
0.31		31	science and engineering associate professionals	1022
0.31		226	other health professionals	546
0.31		24	business and administration professionals	481
0.31		2653	dancers and choreographers	6564
0.31		333	business services agents	617
0.32		413	keyboard operators	157
0.32		2263	environmental and occupational health and hygiene professionals	2470
0.32	0	1112	senior government officials	4289
0.32		2143	environmental engineers	2926
0.32		214	engineering professionals (excluding electrotechnology)	2580
0.32		541	protective services workers	1606
0.32		1222	advertising and public relations managers	7117
0.32		531	child care workers and teachers' aides	712
0.32		1221	sales and marketing managers	4714
0.32	0	3222	midwifery associate professionals	5682
0.33		1411	hotel managers	1561
0.33		262	librarians, archivists and curators	604
0.33		314	life science technicians and related associate professionals	4080
0.33		1420	retail and wholesale trade managers	3063
0.33	0	2635	social work and counselling professionals	4475
0.33		1312	aquaculture and fisheries production managers	4792
0.33		242	administration professionals	647
0.33	0	3230	traditional and complementary medicine associate professionals	96
0.33		1211	finance managers	3258
0.34		3	technicians and associate professionals	4647
0.34		2421	management and organization analysts	2840
0.34		261	legal professionals	3554
0.34		142	retail and wholesale trade managers	5772
0.34		2146	mining engineers, metallurgists and related professionals	7583
0.34		5113	travel guides	4725
0.34		2423	personnel and careers professionals	3371
0.34	0	2221	nursing professionals	3153
0.35		3412	social work associate professionals	2709
0.35		5413	prison guards	4380
0.35		2114	geologists and geophysicists	1158
0.35		2634	psychologists	3659
0.35		2265	dieticians and nutritionists	4486
0.35		3257	environmental and occupational health inspectors and associates	3800

0.35		3355	police inspectors and detectives	2453
0.35		1412	restaurant managers	2903
0.35		263	social and religious professionals	1336
0.35	0	1113	traditional chiefs and heads of village	5734
0.35		312	mining, manufacturing and construction supervisors	669
0.36		2269	health professionals not elsewhere classified	4287
0.36		3341	office supervisors	3401
0.36		2264	physiotherapists	2568
0.36		5222	shop supervisors	4405
0.36	0	3434	chefs	3024
0.36		2655	actors	4364
0.37	0	3332	conference and event planners	4727
0.37		2141	industrial and production engineers	3341
0.37		211	physical and earth science professionals	2139
0.37		310	armed forces occupations, other ranks	2384
0.37	0	2611	lawyers	4013
0.37		331	financial and mathematical associate professionals	3882
0.37	0	2261	dentists	5510
0.37	0	2142	civil engineers	3866
0.37		813	chemical and photographic products plant and machine operators	1067
0.37		41	general and keyboard clerks	4879
0.37		54	protective services workers	360
0.37		811	mining and mineral processing plant operators	836
0.37		0	armed forces occupations	5600
0.38		213	life science professionals	564
0.38		2152	electronics engineers	5080
0.38		2632	sociologists, anthropologists and related professionals	2488
0.38		265	creative and performing artists	4305
0.38	1	4416	personnel clerks	3503
0.38		2164	town and traffic planners	2654
0.38		2145	chemical engineers	2063
0.38		3154	air traffic controllers	1699
0.38		241	finance professionals	682
0.38		952	street vendors (excluding food)	863
0.38		343	artistic, cultural and culinary associate professionals	726
0.38		311	physical and engineering science technicians	2731
0.38		2222	midwifery professionals	3156
0.38		4323	transport clerks	3569
0.38		2149	engineering professionals not elsewhere classified	5980
0.38		25	information and communications technology professionals	879
0.38		3221	nursing associate professionals	3434
0.38		2529	database and network professionals not elsewhere classified	3581
0.38		341	legal, social and religious associate professionals	607
0.38		2132	farming, forestry and fisheries advisers	4700
0.38	0	2111	physicists and astronomers	3846
0.39		131	production managers in agriculture, forestry and fisheries	588
0.39		2422	policy administration professionals	3202

0.39		3153	aircraft pilots and related associate professionals	2986
0.39		2511	systems analysts	2835
0.39		2642	journalists	4324
0.39		2144	mechanical engineers	3759
0.39		7312	musical instrument makers and tuners	6738
0.39		243	sales, marketing and public relations professionals	951
0.39		313	process control technicians	1434
0.39	1	3333	employment agents and contractors	2564
0.39		754	other craft and related workers	463
0.39		2266	audiologists and speech therapists	1786
0.39		7	craft and related trades workers	7377
0.39		334	administrative and specialised secretaries	682
0.39		5412	police officers	3630
0.39		516	other personal services workers	750
0.39		5411	fire-fighters	4600
0.40		2240	paramedical practitioners	2920
0.40		2434	information and communications technology sales professionals	6190
0.40	0	2162	landscape architects	5465
0.40		4322	production clerks	3960
0.40		2112	meteorologists	432
0.40		2161	building architects	4364
0.40		3143	forestry technicians	1670
0.40		321	medical and pharmaceutical technicians	580
0.40	0	6224	hunters and trappers	2774
0.40		264	authors, journalists and linguists	2860
0.40		2433	technical and medical sales professionals (excluding ict)	4558
0.40		5312	teachers' aides	2109
0.40		9997	don't know	1204
0.40		3131	power production plant operators	7421
0.40		2633	philosophers, historians and political scientists	4766
0.40		7215	riggers and cable splicers	12600
0.40		3142	agricultural technicians	4597
0.40	0	5311	child care workers	3566
0.40	0	2612	judges	2716
0.40		43	numerical and material recording clerks	259
0.40	0	2151	electrical engineers	3495
0.40		3134	petroleum and natural gas refining plant operators	3249
0.40	0	2131	biologists, botanists, zoologists and related professionals	3085
0.41		751	food processing and related trades workers	11215
0.41		532	personal care workers in health services	886
0.41		3119	physical and engineering science technicians not elsewhere classified	4356
0.41		5151	cleaning and housekeeping supervisors in offices, hotels and other establishments	3367
0.41	0	3259	health associate professionals not elsewhere classified	4178
0.41		3413	religious associate professionals	3691
0.41		2523	computer network professionals	2232
0.41		5321	health care assistants	4040
0.41		2512	software developers	3996

0.42		325	other health associate professionals	918
0.42		2153	telecommunications engineers	3141
0.42		816	food and related products machine operators	483
0.42		2622	librarians and related information professionals	3739
0.42		324	veterinary technicians and assistants	457
0.42		9999	not stated or inferred	4177
0.42		3513	computer network and systems technicians	2134
0.42		3139	process control technicians not elsewhere classified	2651
0.42		143	other services managers	603
0.42		2432	public relations professionals	2826
0.42	1	2431	advertising and marketing professionals	3123
0.42		2619	legal professionals not elsewhere classified	2620
0.42		335	regulatory government associate professionals	3757
0.42		351	information and communications technology operations and user support technicians	5544
0.42		224	paramedical practitioners	26269
0.42		2113	chemists	2415
0.42		2659	creative and performing artists not elsewhere classified	1635
0.42		2621	archivists and curators	1416
0.43		2519	software and applications developers and analysts not elsewhere classified	2433
0.43		2230	traditional and complementary medicine professionals	7799
0.43		3334	real estate agents and property managers	5017
0.43	0	5111	travel attendants and travel stewards	2670
0.43		3522	telecommunications engineering technicians	5182
0.43		332	sales and purchasing agents and brokers	675
0.43		215	electrotechnology engineers	456
0.43		3435	other artistic and cultural associate professionals	2692
0.43		7316	sign writers, decorative painters, engravers and etchers	4174
0.43		3258	ambulance workers	3175
0.43		42	customer services clerks	5711
0.43		2522	systems administrators	4469
0.43		514	hairdressers, beauticians and related workers	3498
0.43		3323	buyers	3253
0.43		2133	environmental protection professionals	4359
0.43		3115	mechanical engineering technicians	5632
0.44	1	3112	civil engineering technicians	4400
0.44		4419	clerical support workers not elsewhere classified	2952
0.44		4411	library clerks	2243
0.44		2262	pharmacists	5077
0.44		741	electrical equipment installers and repairers	2602
0.44		441	other clerical support workers	2856
0.44		251	software and applications developers and analysts	1052
0.44		3343	administrative and executive secretaries	3962
0.44	1	2411	accountants	3086
0.44		8112	mineral and stone processing plant operators	4120
0.44		5419	protective services workers not elsewhere classified	3925
0.44		7113	stonemasons, stone cutters, splitters and carvers	4721
0.44		2120	mathematicians, actuaries and statisticians	2156
0.44		3135	metal production process controllers	5769

0.44		711	building frame and related trades workers	3663
0.44		3311	securities and finance dealers and brokers	2798
0.44		3255	physiotherapy technicians and assistants	6682
0.44		2513	web and multimedia developers	3227
0.44		53	personal care workers	667
0.44		7121	roofers	3717
0.44		5241	fashion and other models	7047
0.44	1	2165	cartographers and surveyors	3085
0.44	1	2413	financial analysts	2305
0.44		5221	shop keepers	7073
0.44		621	forestry and related workers	477
0.44		5164	pet groomers and animal care workers	3429
0.45		6122	poultry producers	4952
0.45		3256	medical assistants	3056
0.45		731	handicraft workers	413
0.45	1	3339	business services agents not elsewhere classified	3735
0.45	1	3321	insurance representatives	4675
0.45		2412	financial and investment advisers	1768
0.45		2250	veterinarians	3222
0.45		3359	regulatory government associate professionals not elsewhere classified	2751
0.45		712	building finishers and related trades workers	7040
0.45		7232	aircraft engine mechanics and repairers	10133
0.45		225	veterinarians	9107
0.45		3331	clearing and forwarding agents	3689
0.45		252	database and network professionals	820
0.45		8113	well drillers and borers and related workers	4759
0.45		3521	broadcasting and audio-visual technicians	6405
0.45		216	architects, planners, surveyors and designers	665
0.45		3511	information and communications technology operations technicians	4116
0.45		2652	musicians, singers and composers	3956
0.45		8132	photographic products machine operators	1871
0.45	0	3421	athletes and sports players	2364
0.45		3141	life science technicians (excluding medical)	2821
0.45		3512	information and communications technology user support technicians	4320
0.46		4229	client information workers not elsewhere classified	3086
0.46	0	2163	product and garment designers	4442
0.46		524	other sales workers	661
0.46	1	3352	government tax and excise officials	5273
0.46		834	mobile plant operators	2520
0.46		7542	shotfirers and blasters	3263
0.46	0	3354	government licensing officials	2023
0.46		432	material-recording and transport clerks	1423
0.46		7133	building structure cleaners	2407
0.46	1	3312	credit and loans officers	4166
0.46		5169	personal services workers not elsewhere classified	3500
0.46	1	3344	medical secretaries	1924
0.46		3211	medical imaging and therapeutic equipment technicians	2776

0.46		3432	interior designers and decorators	3593
0.46		3133	chemical processing plant controllers	2920
0.46		5242	sales demonstrators	5683
0.46		611	market gardeners and crop growers	2612
0.46		512	cooks	5783
0.46		9998	refused	4825
0.47		412	secretaries (general)	5863
0.47		713	painters, building structure cleaners and related trades workers	2002
0.47	0	3353	government social benefits officials	3784
0.47		6129	animal producers not elsewhere classified	4678
0.47		5163	undertakers and embalmers	5075
0.47		8111	miners and quarriers	3198
0.47		7514	fruit, vegetable and related preservers	5283
0.47	0	3351	customs and border inspectors	2593
0.47		5322	home-based personal care workers	2971
0.47		613	mixed crop and animal producers	545
0.47		3322	commercial sales representatives	4856
0.47		3433	gallery, museum and library technicians	5253
0.47	1	3411	police inspectors and detectives	3702
0.47		7321	pre-press technicians	4895
0.47		632	subsistence livestock farmers	11089
0.47		742	electronics and telecommunications installers and repairers	1687
0.47		8171	pulp and papermaking plant operators	6452
0.47		7124	insulation workers	3480
0.47		3113	electrical engineering technicians	4809
0.47		3431	photographers	4461
0.48		3111	chemical and physical science technicians	1893
0.48		4225	enquiry clerks	3208
0.48		7127	air conditioning and refrigeration mechanics	3938
0.48		3240	veterinary technicians and assistants	6522
0.48		7315	glass makers, cutters, grinders and finishers	8756
0.48		7543	product graders and testers (excluding foods and beverages)	4547
0.48		2521	database designers and administrators	3317
0.48		3251	dental assistants and therapists	4865
0.48	1	8350	ships' deck crews and related workers	7225
0.48		7119	building frame and related trades workers not elsewhere classified	4604
0.48		5329	personal care workers in health services not elsewhere classified	3883
0.48		3254	dispensing opticians	4151
0.48		4	clerical support workers	4606
0.48		5152	domestic housekeepers	5159
0.48		411	general office clerks	895
0.48		7411	building and related electricians	3806
0.48		6222	inland and coastal waters fishery workers	11366
0.48	1	3315	valuers and loss assessors	2646
0.48		723	machinery mechanics and repairers	3472

0.48		8181	glass and ceramics plant operators	4789
0.48		933	transport and storage labourers	5590
0.49		3314	statistical, mathematical and related associate professionals	2419
0.49		93	labourers in mining, construction, manufacturing and transport	8597
0.49		3151	ships' engineers	1509
0.49		2267	optometrists and ophthalmic opticians	2679
0.49		7314	potters and related workers	9104
0.49	0	2631	economists	3041
0.49		4321	stock clerks	3339
0.49		515	building and housekeeping supervisors	629
0.49		6210	forestry and related workers	3803
0.49		7413	electrical line installers and repairers	12835
0.49		3114	electronics engineering technicians	5092
0.49	0	5141	hairdressers	4972
0.49		5112	transport conductors	7577
0.49		5162	companions and valets	9561
0.49		3116	chemical engineering technicians	2942
0.49		4227	survey and market research interviewers	4116
0.49		962	other elementary workers	772
0.49		4221	travel consultants and clerks	3862
0.49		4214	debt-collectors and related workers	4076
0.49	0	5142	beauticians and related workers	4401
0.49		4120	secretaries (general)	2416
0.49	1	4312	statistical, finance and insurance clerks	3129
0.49		814	rubber, plastic and paper products machine operators	2371
0.49		212	mathematicians, actuaries and statisticians	720
0.49		631	subsistence crop farmers	602
0.49		421	tellers, money collectors and related clerks	4001
0.49		5414	security guards	7285
0.49		3514	web technicians	3763
0.49		941	food preparation assistants	5860
0.50		3117	mining and metallurgical technicians	2498
0.50	1	3118	draughtspersons	4491
0.50	1	2641	authors and related writers	3283
0.50		7132	spray painters and varnishers	3561
0.50		431	numerical clerks	3492
0.50		622	fishery workers, hunters and trappers	4513
0.50		5120	cooks	5123
0.50		3132	incinerator and water treatment plant operators	4342
0.50	0	7126	plumbers and pipe fitters	5558
0.50		7115	carpenters and joiners	3467
0.50		9996	valid skip	3974
0.50		7322	printers	5094
0.50		422	client information workers	2969
0.50		8131	chemical products plant and machine operators	6621
0.50		7122	floor layers and tile setters	3746
0.50		7111	house builders	3232
0.50		513	waiters and bartenders	2934

0.50		9216	fishery and aquaculture labourers	1473
0.50		7131	painters and related workers	4514
0.50		722	blacksmiths, toolmakers and related trades workers	532
0.50		4222	contact centre information clerks	5544
0.50		8159	textile, fur and leather products machine operators not elsewhere classified	3792
0.50		2643	translators, interpreters and other linguists	4034
0.50		7422	information and communications technology installers and servicers	3608
0.51		7112	bricklayers and related workers	4990
0.51	0	4224	hotel receptionists	3057
0.51		7123	plasterers	4127
0.51		3324	trade brokers	1834
0.51		721	sheet and structural metal workers, moulders and welders, and related workers	8852
0.51		7233	agricultural and industrial machinery mechanics and repairers	3506
0.51		4414	scribes and related workers	1152
0.51		4110	general office clerks	6073
0.51		7221	blacksmiths, hammersmiths and forging press workers	4107
0.51		732	printing trades workers	5102
0.51	1	4223	telephone switchboard operators	2784
0.51		831	locomotive engine drivers and related workers	10381
0.51		7412	electrical mechanics and fitters	3184
0.51		752	wood treaters, cabinet-makers and related trades workers	12056
0.51		9311	mining and quarrying labourers	4444
0.51		7521	wood treaters	5937
0.51		3252	medical records and health information technicians	5210
0.51		6113	gardeners, horticultural and nursery growers	3918
0.51		72	metal, machinery and related trades workers	5716
0.51		7544	fumigators and other pest and weed controllers	7832
0.51		3212	medical and pathology laboratory technicians	2746
0.51		6111	field crop and vegetable growers	6695
0.51		7125	glaziers	3816
0.52		6112	tree and shrub crop growers	3456
0.52		3214	medical and dental prosthetic technicians	4119
0.52		7214	structural-metal preparers and erectors	3190
0.52		4413	coding, proof-reading and related clerks	3492
0.52		3342	legal secretaries	2709
0.52		3213	pharmaceutical technicians and assistants	3607
0.52		4313	payroll clerks	3474
0.52		5249	sales workers not elsewhere classified	4082
0.52		2166	graphic and multimedia designers	3193
0.52		5153	building caretakers	4583
0.52		8312	railway brake, signal and switch operators	4370
0.52		5243	door to door salespersons	4743
0.52		8343	crane, hoist and related plant operators	5198
0.52		7231	motor vehicle mechanics and repairers	4668
0.52		5161	astrologers, fortune-tellers and related workers	4890

0.52		2514	applications programmers	2341
0.52		612	animal producers	1550
0.52		821	assemblers	4581
0.52		4226	receptionists (general)	4087
0.52		4213	pawnbrokers and money-lenders	1154
0.52		5211	stall and market salespersons	6227
0.52		7522	cabinet-makers and related workers	2958
0.52		812	metal processing and finishing plant operators	6159
0.52		7515	food and beverage tasters and graders	4586
0.52		7511	butchers, fishmongers and related food preparers	5144
0.53		8342	earthmoving and related plant operators	2982
0.53		6130	mixed crop and animal producers	3331
0.53		753	garment and related trades workers	2407
0.53		7512	bakers, pastry-cooks and confectionery makers	5098
0.53		5246	food service counter attendants	6507
0.53		71	building and related trades workers, excluding electricians	10445
0.53		4211	bank tellers and related clerks	4589
0.53		633	subsistence mixed crop and livestock farmers	44632
0.53		6223	deep-sea fishery workers	7309
0.53		522	shop salespersons	653
0.53		7311	precision-instrument makers and repairers	6075
0.53		2651	visual artists	3612
0.53		4311	accounting and bookkeeping clerks	3593
0.53		9611	garbage and recycling collectors	3991
0.53		932	manufacturing labourers	7172
0.53	0	5131	waiters	4273
0.53		9312	civil engineering labourers	3012
0.53		8142	plastic products machine operators	4437
0.53	1	9629	elementary workers not elsewhere classified	3345
0.53	1	4415	filing and copying clerks	3301
0.53		7114	concrete placers, concrete finishers and related workers	3577
0.54		6121	livestock and dairy producers	2656
0.54		7421	electronics mechanics and servicers	5610
0.54		8	plant and machine operators, and assemblers	3781
0.54		931	mining and construction labourers	3620
0.54		3313	accounting associate professionals	3594
0.54		9622	odd job persons	3699
0.54		8152	weaving and knitting machine operators	7922
0.54		9129	other cleaning workers	8332
0.54		911	domestic, hotel and office cleaners and helpers	2402
0.54		7549	craft and related workers not elsewhere classified	3386
0.54		8189	stationary plant and machine operators not elsewhere classified	3264
0.54		8211	mechanical machinery assemblers	5553
0.54	1	4212	bookmakers, croupiers and related gaming workers	2828
0.54		7212	welders and flamecutters	4425
0.54		8122	metal finishing, plating and coating machine operators	3897
0.54		5244	contact centre salespersons	3634
0.54		7234	bicycle and related repairers	2308

0.54		6114	mixed crop growers	9315
0.54		33	business and administration associate professionals	581
0.54	1	4132	data entry clerks	4905
0.54		9313	building construction labourers	4280
0.54		833	heavy truck and bus drivers	751
0.54		6330	subsistence mixed crop and livestock farmers	2474
0.54		8141	rubber products machine operators	4335
0.54		82	assemblers	3018
0.54		835	ships' deck crews and related workers	622
0.54		5132	bartenders	3404
0.54		8160	food and related products machine operators	3656
0.54	1	7213	sheet-metal workers	4023
0.55		7222	toolmakers and related workers	4927
0.55		4131	typists and word processing operators	4301
0.55		9214	garden and horticultural labourers	3125
0.55		8154	bleaching, dyeing and fabric cleaning machine operators	10188
0.55		9334	shelf fillers	2452
0.55		9333	freight handlers	5391
0.55		5223	shop sales assistants	4112
0.55		81	stationary plant and machine operators	2843
0.55		9215	forestry labourers	2420
0.55		8151	fibre preparing, spinning and winding machine operators	3518
0.55		9122	vehicle cleaners	3798
0.55	1	9411	fast food preparers	6728
0.55		352	telecommunications and broadcasting technicians	755
0.55		7513	dairy-products makers	3029
0.55		8172	wood processing plant operators	3941
0.55		7536	shoemakers and related workers	7583
0.55		7534	upholsterers and related workers	2616
0.55		9213	mixed crop and livestock farm labourers	5117
0.55		817	wood processing and papermaking plant operators	5099
0.55		52	sales workers	2069
0.55		815	textile, fur and leather products machine operators	11050
0.55		9613	sweepers and related labourers	2815
0.56		7317	handicraft workers in wood, basketry and related materials	3917
0.56		6123	apiarists and sericulturists	2795
0.56		8143	paper products machine operators	3643
0.56		8311	locomotive engine drivers	5159
0.56		7523	woodworking-machine tool setters and operators	1421
0.56	1	8344	lifting truck operators	3267
0.56	1	7223	metal working machine tool setters and operators	3903
0.56		7211	metal moulders and coremakers	11174
0.56	1	9621	messengers, package deliverers and luggage porters	4960
0.56		8183	packing, bottling and labelling machine operators	4167
0.56		8121	metal processing plant operators	4977
0.56		9624	water and firewood collectors	224
0.56		7323	print finishing and binding workers	4029

0.57		7313	jewellery and precious-metal workers	5139
0.57		8114	cement, stone and other mineral products machine operators	7454
0.57		9520	street vendors (excluding food)	7754
0.57		9211	crop farm labourers	5597
0.57		9329	manufacturing labourers not elsewhere classified	6190
0.57	1	4412	mail carriers and sorting clerks	3316
0.57		8219	assemblers not elsewhere classified	5508
0.57		6221	aquaculture workers	5379
0.57	1	9331	hand and pedal vehicle drivers	4415
0.57		9510	street and related service workers	2629
0.57	0	9112	cleaners and helpers in offices, hotels and other establishments	3078
0.57	1	9412	kitchen helpers	3623
0.57	1	8212	electrical and electronic equipment assemblers	4506
0.58		7318	handicraft workers in textile, leather and related materials	6137
0.58		9123	window cleaners	1413
0.58		9612	refuse sorters	3245
0.58		7535	pelt dressers, tanners and fellmongers	3382
0.58	1	8153	sewing machine operators	4583
0.58		8341	mobile farm and forestry plant operators	4570
0.58		7224	metal polishers, wheel grinders and tool sharpeners	4838
0.58		83	drivers and mobile plant operators	1790
0.58		8182	steam engine and boiler operators	5071
0.58		9321	hand packers	3965
0.58	1	8331	bus and tram drivers	4487
0.59		523	cashiers and ticket clerks	3601
0.59	1	8322	car, taxi and van drivers	8298
0.59	1	5230	cashiers and ticket clerks	3072
0.59		832	car, van and motorcycle drivers	630
0.59		8157	laundry machine operators	4043
0.59		7532	garment and related pattern-makers and cutters	4680
0.59		951	street and related service workers	3431
0.59	0	6340	subsistence fishers, hunters, trappers and gatherers	4276
0.59		8156	shoemaking and related machine operators	7552
0.59		6310	subsistence crop farmers	9453
0.59	1	9623	meter readers and vending-machine collectors	4881
0.59		8332	heavy truck and lorry drivers	4385
0.59		818	other stationary plant and machine operators	7526
0.59	0	9111	domestic cleaners and helpers	4281
0.60		9121	hand launderers and pressers	3812
0.60		7319	handicraft workers not elsewhere classified	4795
0.60		91	cleaners and helpers	14444
0.60		7516	tobacco preparers and tobacco products makers	1044
0.60		9332	drivers of animal-drawn vehicles and machinery	259
0.60		5245	service station attendants	4032
0.60		8155	fur and leather preparing machine operators	8640
0.60		961	refuse workers	6258
0.60		7533	sewing, embroidery and related workers	7134

0.61		9212	livestock farm labourers	4117
0.61		5212	street food salespersons	9269
0.61		92	agricultural, forestry and fishery labourers	23397
0.61		9	elementary occupations	2719
0.61		921	agricultural, forestry and fishery labourers	3078
0.62		7531	tailors, dressmakers, furriers and hatters	3678
0.63		912	vehicle, window, laundry and other hand cleaning workers	332
0.64		75	food processing, wood working, garment and other craft and related trades workers	905
0.65		61	market-oriented skilled agricultural workers	4582
0.68		6320	subsistence livestock farmers	2705
0.68		34	legal, social, cultural and related associate professionals	939
0.69	1	8321	motorcycle drivers	8718
0.74		22	health professionals	495

Table B.1: Predicted risk of automation for ISCO coded occupations

Appendix C

Education level per country vs automation risk

Figure C.1: Educational attainment vs automation risk distribution per country

Appendix D

Employment shares per risk levels by country

Country	Significant risk	High risk	Medium risk	Small risk
Norway	0.06	0.44	0.32	0.18
Finland	0.06	0.44	0.32	0.18
Sweden	0.08	0.44	0.30	0.17
Denmark	0.09	0.48	0.29	0.15
Japan	0.09	0.55	0.27	0.10
United States	0.10	0.43	0.29	0.18
New Zealand	0.10	0.42	0.30	0.18
Belgium	0.10	0.54	0.24	0.12
Netherlands	0.10	0.50	0.27	0.13
United Kingdom	0.11	0.48	0.25	0.16
Canada	0.11	0.49	0.25	0.14
Turkey	0.12	0.62	0.19	0.08
Estonia	0.12	0.51	0.25	0.12
Korea	0.12	0.50	0.26	0.12
Hungary	0.12	0.60	0.20	0.08
Singapore	0.12	0.46	0.26	0.16
Austria	0.13	0.52	0.24	0.12
Israel	0.13	0.50	0.23	0.14
France	0.13	0.57	0.20	0.10
Ireland	0.14	0.48	0.23	0.15
Germany	0.15	0.57	0.21	0.07
Italy	0.15	0.57	0.19	0.09
Czech Republic	0.15	0.53	0.23	0.09
Poland	0.16	0.56	0.19	0.09
Russia	0.16	0.54	0.21	0.09
Spain	0.17	0.55	0.19	0.09
Slovenia	0.17	0.53	0.18	0.11
Chile	0.18	0.49	0.22	0.11
Lithuania	0.18	0.61	0.15	0.06

Cyprus	0.20	0.46	0.23	0.11
Greece	0.20	0.51	0.21	0.08
Mexico	0.23	0.50	0.18	0.09
Slovak Republic	0.27	0.51	0.14	0.08
All countries	**0.14**	**0.51**	**0.23**	**0.12**

Table D.1: Employment shares per risk levels by country