ROLE OF HIGHER EDUCATION IN THE ERA OF AUTOMATION: A TASK-BASED APPROACH

Hannah Massenbauer (11801877)
Economics, Applied track
Vienna University of Economics and Business
Advisor: Florian W. Szücs

October 8, 2023

Contents

1 Introduction **1**

2 Literature Review **2**

3 Methodology **5**

3.1 Data 5

3.2 Empirical Method 9

4 Results **12**

4.1 Static results 12

4.2 Dynamic Results 19

5 Robustness **22**

5.1 Heterogeneities 22

5.1.1 Heterogeneity in Gender 22

5.1.2 Heterogeneity in Income 26

5.2 Model Specification 28

5.2.1 Binary model 28

5.2.2 Quantile Regression 30

6 Discussion **32**

7 Conclusion **33**

8 References **36**

9 Appendix **39**

1 Introduction

In an era characterized by rapid technological change, the labor market faces some major transformations. The integration of technology into our work lives has raised concerns about the future of work and the role of education in shaping individuals' occupation prospects. Higher education and the skills acquired through it are increasingly decisive for labor market outcomes. Technology either replaces workers without the necessary skills or, if they are sufficiently skilled, technology is used complementary. This leads to an emerging gap between low- and high-skilled workers (Katz et al., 1999).

This research examines the impact of higher education on the degree of routine task intensity (RTI) of individuals in their occupations. The RTI exploits a task-based approach and serves as a measure to assess the robustness of an occupation against automation. Additionally, variations across different countries are covered to shed light on regional disparities in the process of adapting to labor market changes. An additional step is undertaken by examining not only static effects, but also dynamic effects, to capture short-term trends across Europe.

This analysis investigates the role of education by examining heterogeneities in subsamples such as gender and income levels. Furthermore, expanding the Ordinary Least Squares (OLS) estimation with a logit and a quantile regression allows to assess the robustness of the estimator. Revealing these heterogeneities ensures a deeper understanding of the role of education.

Finally, I will contextualize the results in a discussion. Given that changes in the labor market significantly influence individual's livelihoods and enduring consequences for countries, it is necessary to address the inherent political implications and potential policy actions. The labor market is changing fundamentally, and this research aims to examine the role of education in this process.

2 Literature Review

New technologies are tied to disruptions in the labor market that will raise concerns for politicians and society (Gallego and Kurer, 2022 & Autor, 2022). Literature has tried to address several concerns and has formulated hypothesis such as the skill-biased technology change (SBTC) and routine-biased technology change (RBTC) hypotheses. I will first clarify their characteristics and, after setting up the base, elaborate how my research contributes to this growing branch of literature.

In light of technological change, the SBTC hypothesis poses that new technologies are biased toward highly skilled workers (Acemoglu, 1998 & Bittarello et al., 2018). Higher skills are approximated by educational attainment. In the past, the rise of technologies has resulted in an increase in the demand for highly educated people, and the share of higher education experienced growth (Goldin and Katz, 1996). In particular, in post-industrialized states, tertiary education shares have experienced notable growth. Based on this observation, the canonical approach emerged in the labor economics literature (Acemoglu and Autor, 2011). This led me to take advantage of a dichotomous approach that divides the population into workers with higher education (high-skilled) and workers who have attained maximum secondary education (low-skilled).

Technology can exhibit multifaceted effects. New technologies pose a significant threat to low-skilled workers, functioning as substitutes while skilled workers tend to profit. The profit can be attributed to the fact that skilled workers can use technology complementary, which constitutes a further driver of the disparities between these two groups. Acemoglu and Autor (2011), Bessen (2016) and Balsmeier and Woerter (2019) analyze this emerging gap from the perspective of employment-shares and wages. This approach is interesting, however, it neglects the importance of tasks, since a large stream of literature argues that a task approach is necessary to capture the depth of modern technological transition (David, 2013).

Thus, the next step is to turn from the SBTC to a task-biased technology change hypothesis (Maier, 2022). This hypothesis distinguishes workers not based on their skill level but on the tasks that are included in their occupations. Going deeper into this approach leads us to the routine biased technology change (RBTC) hypothesis. The RBTC describes that new technologies are biased to replace routine tasks more than non-routine tasks. It also explains the wage polarization, which causes a hollowing out of middle-wage occupations, and further captures the essence of current technology waves (Haslberger, 2021, Autor et al., 2003, Acemoglu and Autor, 2011 & Manning et al., 2009). All processes that follow rules, either manual or cognitive, can be replaced. Therefore, the change in employment structures is not due to the choice of occupation, but rather to the occupation's composition. In contrast, the SBTC considers the characteristics of workers, while the RBTC considers the routine tasks of an occupation (Haslberger, 2021). This insight is the reason why I am working with the task-based approach. I am not analyzing the employment shares or their wages, but the routine task intensity in peoples occupations.

Choosing the task-based approach entails certain challenges. First, a decision must be made as to which task measurement to use. Starting with some common ground, the advantages and disadvantages of the selected measure get discussed. As introduced by Autor et al. (2003) there are five categories of tasks, which are widely adopted by the authors (Spitz-Oener, 2006, Mihaylov and Tijdens, 2019 & Dengler et al., 2014). The measure of RTI is constructed by the chosen task items. I rely on Mihaylov and Tijdens (2019), since they have conducted a cross-study analysis where they show that their measurement corresponds closely to others. Additionally Mihaylov and Tijdens (2019) are discussing the drawbacks that can occur when using the Occupational Information Network (O*NET). The O*NET database contains the linking information between tasks and occupations, which is necessary to achieve a task-based approach. The O*NET database is constructed based on US occupations, and experts have assigned tasks to occupations. Criticism of this task measurement concept has been expressed, as it is designed by experts based on US occupations and neglects within occupation variations and potential differences between countries (Haslberger, 2022). However, since there exists no other dataset that contains this many countries, I am conducting my research while acknowledging these drawbacks. On the other hand, it is practical to have a consistent definition of RTI across countries, since the methodology and definition of certain tasks are standardized in this way. Albeit I see the point of critique, it seems not to be as worrying in my context as the aim is not to examine within occupation variation but the variation between occupations.

What sets my research apart is the novel data source and methodology. While most related papers examined this topic using the Labor force survey (Górka et al., 2017, Maier, 2022 & Hardy et al., 2018), EUKLEMS (Michaels et al., 2014), European skills and jobs survey (Pouliakas, 2018), Program for the International Assessment of Adult Competencies (PIAAC), or German sources such as BERUFENET, BIBB/IAB and BIBB/BAuA Employment Surveys (Mihaylov and Tijdens, 2019), I decided to work with the European Social Survey (ESS). The ESS provides unique data for many different socioeconomic characteristics and a broad set of countries from 2002 onward. This combination allows me to achieve new insights into population heterogeneities across Europe.

In terms of methodology, my approach is stratified at the individual and country level. The first level is the individual level. To the best of my knowledge, there is no other paper, except the work of Spitz-Oener (2006), in which the characteristics of the individual and the occupation are regressed against each other. However, she uses each category of tasks as an outcome variable. This is due to her goal of capturing the changes in different task categories carried out by high-, middle-, or low-skilled workers. My approach uses a one-dimensional routine task intensity index as the dependent variable. This allows me to comprehensively observe, on the individual level, the intertwining of higher education and routine task intensity. In other words, this research aims to obtain the potential RTI depending on one's own educational level.

The core of my analysis takes a further step by observing the cross-country perspective. The view from a country level enables a comparison of the efficiency of human capital allocation between countries. This comparison provides a broader perspective on the ongoing labor market transformation. The importance of the cross-country level is given by its inherent consequences not only for the countries, but also for the European Union. Brain drain and regional political conflicts are possible repercussions. The country level analysis consists of a static and dynamic part.

The existing literature discusses the impact of technology that can replace high-skill tasks (Frey and Osborne, 2017 & Frey and Osborne, 2013). However, in this research, the focus remains on technologies that existed in the period between 2012-2018 and, thus, excludes the possibilities of technologies that replace non-routine work. Thus, a premise of this research is that technologies are only capable of replacing routine tasks, including routine cognitive tasks.

3 Methodology

3.1 Data

Before diving into the methodology, I want to provide information on the data used in this work. The main data source is the European Social Survey (ESS). Specifically, four waves were used in the period from 2012 to 2018. The time frame choice is contingent upon the consideration of unchanged occupation classification and crosswalk between occupation and their task composition. The decision to reduce the number of selected countries is based on the data availability constraints of certain countries. Based on this criterion, 18 countries are left [1]. From the ESS I use several individual characteristics, including gender, age, mother's education, birthplace[2], and household net income. Furthermore, my sample is restricted to the working population between the ages of 25 and 60. The purpose of this selection of variables is to cover the most important socioeconomic characteristics.

[1] Belgium (BE), Switzerland (CH), Czech Republic (CZ), Germany (DE), Estonia (EE), Spain (ES), Finland (FI), France (FR), Great Britain (GB), Hungary (HU), Ireland (IE), Lithuania (LT), Netherlands (NL), Norway (NO), Poland (PL), Portugal (PT), Sweden (SE), Slovenia (SI)

[2] Birthplace is a binary variable indicating if the current place of residence is equal the country of birth. Where 1 is a yes and 2 a no.

Supplementary, I want to clarify how I defined certain variables. Beginning with the treatment variable, education, the ESS provides information based on the ISCED-classification. To reduce the dimensions of the variable, I classified all ISCED levels equal to five or higher as tertiary education[3], which results in a binary variable. Similarly, defined is mothers' education. To reduce the dimensions of mother's education, a binary variable was created analogously to the binary classification of individuals' education, whether the mother had obtained tertiary education or not.

In addition to individual characteristics, a consideration of country-specific characteristics is required. Important country-level characteristics related to the research question are the share of people who have attained higher education in a country and the share of government investment in research and developement (R&D). The R&D variable is measured in per capita terms and accounts for purchasing power parity. These two country-level variables are intended to control for structural differences in education and research expenditures, respectively, and indirectly for technological progress in a country. The country-specific variables are extracted from EUROSTAT for the same years as the ESS waves took place.

A third source of data is the occupational information network (O*NET). Since I follow a task-based approach using the RTI to determine the exposure of an occupation's risk to automation, I had to augment the occupations of the individuals in the ESS dataset with further information about their task composition. This crosswalk is available at O*NET. The merging of occupations, defined by the ISCO08-classification, with their related tasks is a fundamental step in my analysis. Therefore, I will now elaborate in more detail on the considerations and decisions I made.

The tasks provided in O*NET range from abilities to activities and beyond. When faced with the decision of how to map the sheer volume of task information to occupations, I decided to lean on Mihaylov and Tijdens (2019). As my research is not concerned with the details of the mapping procedure but with the role of education in this rapidly

[3] Level 5: short-cycle tertiary programs, level 6: Bachelor, level 7: Master, level 8: Doctoral

changing labor market. To enable a profound understanding of my model, I will now explain how the mass of task information was transformed into a one-dimensional RTI index, capturing the extent to which an occupation can be automated. The paper by Mihaylov and Tijdens (2019) clearly explains how they classify 3,264 tasks into five task categories and then bridge them to 427 occupations. Additionally, they discuss potential crosswalk pitfalls. Therefore, I use the direct classification of Mihaylov and Tijdens (2019) obtained from their Appendix B, as they incorporate a cross-study comparison of different definitions of routine intensity.

The following provides a brief explanation of the RTI index. According to the state of literature, tasks are split into five categories. The five categories are routine manual (RM), routine cognitive (RC), non-routine manual (NRM), non-routine abstract (NRA), and non-routine interpersonal (NRI). Each task category comprises various tasks. Therefore, Mihaylov and Tijdens (2019) construct in a first step a measure T_{gk}, which captures how many tasks j are assigned to a certain task category g divided by all the tasks j existing in occupation k.

$$T_{gk} = \frac{j_g}{\sum_{g=1}^{5} j_g} \tag{1}$$

This results in five different task categories of T_{gk} (RM, RC, NRM, NRA, NRI), which are precisely the above-mentioned categories. The sum of routine tasks minus the non-routine tasks yields a task composition index, the so-called routine task intensity (RTI).

$$RTI = RM + RC - NRM - NRA - NRI \tag{2}$$

This index serves as the dependent variable in my model. An important note at this point is that the distribution of the RTI remains between -1 and 1 through its standardization achieved by the first step of creating this index. This standardization ensures the comparability of the RTI across occupations.

In the next step, an intuitive understanding of the RTI is provided. If the index takes the value -1, it implies that no routine tasks are executed in this occupation. The opposite case, where the RTI is equal to 1, implies that an occupation consists merely of routine tasks. Therefore, a more negative value implies a reduced risk of automation.

In table 3.1, a descriptive overview of the data is shown to provide a first impression of how characteristics are allocated depending on the educational level. The first column shows the mean of the selected variable if people are not highly educated. The second column shows the mean of individuals with higher education. In the third column, the difference between the educational levels is calculated to facilitate the understanding of the disparities between the two groups.

Table 1: Comparison of Means

	Non-High	High	Difference
Age	44.53	42.33	-2.20
Gender	1.50	1.57	0.07
Mother's educ.	0.15	0.29	0.14
Birthplace	1.10	1.12	0.02
HH total net income	5.15	6.80	1.65
RTI	-0.36	-0.59	-0.23
NRA	0.12	0.35	0.23
NRI	0.16	0.31	0.15
NRM	0.41	0.14	-0.27
RC	0.22	0.18	-0.04
RM	0.10	0.02	-0.07
Observations	37450.00	33026.00	70472.00

It quickly becomes evident that there exist differences. For example, people who have attained higher education are twice as likely to have a mother with higher education. The bottom half of the table reveals the distribution of the RTI as well as its component distribution among the subgroups in my population. Unsurprisingly, people with higher education work more often in occupations consisting of non-routine tasks. Interestingly, the differences for routine tasks are smaller than in the non-routine tasks. This suggests that individuals with a higher level of education possess the ability to perform both routine and non-routine tasks, whereas those without higher education are less successful in carrying out non-routine tasks.

After this section, I will outline the empirical method and how the data are used to pursue the research question.

3.2 Empirical Method

The next step contains an outline of the empirical method. First, a regression analysis is carried out to obtain insights on the importance of higher education for automation risk. A person who has attained tertiary education is classified as highly skilled, whereas a person without tertiary education is classified as low skilled. Therefore, identification tries to establish what the difference in outcome, the RTI, of one's occupation is, depending on the level of education (Angrist and Krueger, 1999)

$$RTI_{i,g} = \beta_1 \text{heduc}_i + \epsilon_i \tag{3}$$

Heduc is a binary variable denoting the attained level of education. The outcome variable RTI depends on the individual i and the occupation g individual i is engaged in. However, it is apparent that the level of education will depend on numerous factors that are yet hidden in the error term. To avoid omitted variable bias, I introduce additional covariates to directly control for the yet-unobserved confounders.

$$RTI = \alpha + \beta_1 \text{heduc} + \beta_2 X_{individual} + \beta_3 X_{country} + \epsilon_i \tag{4}$$

Starting with individual characteristics, $X_{individual}$ includes the variables gender, age, mother's education, household's net income and birthplace. Controlling for socioeconomic characteristics is crucial, since the level of educational attainment is closely related to socioeconomic status (Chmielewski, 2019). Second, the variable $X_{country}$ includes country-specific variables such as the share of the population who have attained higher education and the government expenditure invested in R&D. The country-specific controls aim to capture structural differences between countries.

The model will be extended incrementally. Starting with an OLS estimation, this will deliver first insights. The next steps are to append the model with a clustered error term and to implement fixed effects. These are necessary steps, as it is likely that the OLS model suffers from several potential econometric issues such as heteroskedasticity and omitted variable bias.

Firstly, I address the issue of heteroskedastic errors. Heteroskedasticity leads to a violation of the Gauss-Markov assumption of homoskedasticity and will yield an estimator that might be inefficient (Verbeek, 2008). To account for the likely scenario that errors are distributed differently depending on the country, year or industry, I cluster the error based on these three variables (Cameron and Miller, 2015). Adding multiple clusters, called multiway clustering, is applied when robust errors fail, due to non-nested clusters (Cameron and Miller, 2015). A non-nested form of clusters is what I am dealing with, as year and country are not sublevels of each other.Clustering deals with the issue of heteroskedasticity without imposing any further assumptions.

Secondly, there exists unobserved variation across countries and years, which should be addressed. Absorbing the unobserved time-invariant variation within countries and years can be achieved with a fixed effects estimation. Consequently, I will absorb country and year specific characteristics that are time invariant to obtain a statistically valid result of the variables' effects.

After this initial step, the analysis will turn to its actual purpose. Insights into possible challenges and disparities between European countries are gathered in light of the transformation of the labor market. Therefore, I extend the presented base model by adding an interaction between the higher education variable and the country variable, which comprises the 18 countries.

$$RTI = \alpha + \beta_1 heduc * country + \beta_2 X_{individual} + \beta_3 X_{country} + \epsilon_i \tag{5}$$

Introducing an interaction term allows for splitting the prior obtained average effect into its country-specific components. As a result, country-specific differences can be examined. The country-level analysis includes only year-fixed effects, because otherwise the effect on a country-level would get absorbed due to multicollinearity. The obtained coefficient will reflect not only the importance of education on the occupation's RTI in one country, but also, due to its relative comparison between countries, differences in the efficiency of human capital allocation.

Efficient allocation of human capital is fundamental to achieve technological and consequently societal progress. When comparing the higher education coefficient between countries, the interpretation of the coefficient changes as the level of observation changes. Instead of the individual level, where the emphasis lies on the importance of education, it reflects the efficiency of human capital allocation at the country level. The focus is now on a comparative element, where the interpretation is based on the relative outcome. To be precise, a more negative coefficient in contrast to other countries reflects a more efficient allocation of human capital. This interpretation is derived from the assumption that a larger magnitude is reflecting a larger outcome gap of an occupation's RTI depending on the person's education level. A larger difference in the allocation of workers reflects a more efficient allocation, since the occupations of people are better matched to their skills. On the downside, it also implies that the gap in occupations executed by high- and low-skilled workers could expand, which could amplify existing disparities.

Of course, the status quo can deliver interesting insights, but an additional piece of information can be extracted by considering the time-trend of RTI across countries. Therefore, the last step of my core analysis consists of a dynamic analysis. This dynamic analysis examines the relation between the initial average value of a country's RTI in 2012 and its average change rate until 2018. Through the utilization of this dynamic analysis, it is possible to gain deeper insights.

At this point, it is relevant to elaborate on the similarity between my approach and the convergence theory of economic growth. My approach resembles in some way an unconditional β-convergence, but there are some crucial differences. The convergence theory of economic growth is embedded in a neoclassical framework that explains long-term growth (Barro and Sala-i-Martin, 1990). My analysis runs in the short run, and hence it is not possible to state any long-term trajectories. In this sense, I cannot make any statement about convergence or divergence that could be interpreted causally. I remain in the field of correlation-related analysis and do not want to claim causality.

4 Results

4.1 Static results

In the following, I present my results. Figure 1 provides a first indication of how education divides the allocation of RTI performed by workers in their occupations.

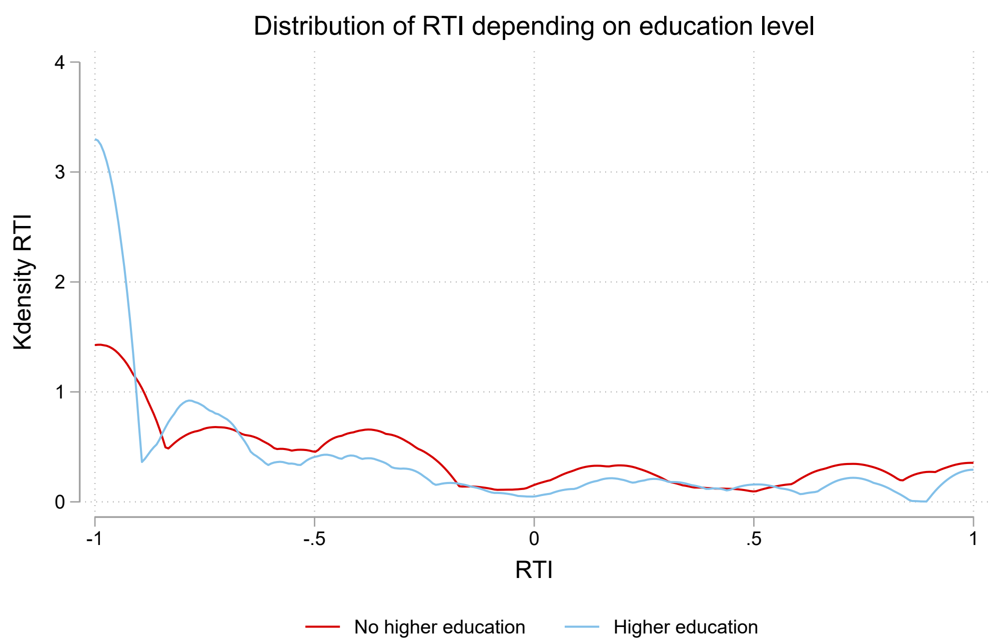


Figure 1

Figure 1 shows a gap in the allocation of RTI between workers with and without higher education. Working in an occupation without routine tasks is determined by educational attainment. As an interim result, it can be stated that people with higher education have a lower risk of being replaced in their occupation, simply because their share of routine tasks is lower. The subsequent step quantifies this gap numerically.

Moving stepwise through the table 2, I start by explaining what can be seen in each column. As a reminder, the dependent variable is the RTI, where a lower value implies less routine work. Thus, if a coefficient shows a negative value, it can be taken as a positive outcome. Firstly, column (1) shows the OLS estimation results. It is striking that most of the coefficients are statistically significant. This can be derived from the comparably lower standard deviation that the OLS estimate exhibits.

In column (2) from table 2, I added clusters for country, year and industry on a NACE 1 level, to account for the possible heteroskedasticity in the error term. Multi-way clustering accounts for each industry in each country and year. After accounting for heteroskedasticity in the error term, almost no variable is significant anymore, except mother's education. The change in standard errors can be explained by the circumstance that the OLS estimation is inefficient due to heteroskedasticity. The violation of homoskedasticity results in a downward bias toward zero in the standard errors, which often leads to highly significant estimates (Moulton, 1990). Accounting for heteroskedasticity is necessary to draw inference reliably. From now on, my model implies this cluster specification.

As my data set exhibits a stratified structure with individuals from different countries, it seems intuitive to account for country-specific effects. First, I include country-fixed effects to absorb variations between countries. After this modification, the treatment variable, higher education, yields a significant result again. To make a further step in my analysis, I add year fixed effects in the fourth column. In a fifth step, since the implementation of both fixed effects is reasonable, I append the model with country and year fixed effects. The fifth model is the core model on which I will base my further analysis, as it considers all possible concerns that could result in invalid results due to misspecified models.

Having explained the setup, I will now elaborate on the interpretation of the magnitude of the coefficients in table 2. The outcome variable takes on values between -1 and 1, thus a coefficient of -0,2163 corresponds to a decrease in RTI of 10,82%. The result indicates that a person who has completed tertiary education works in an occupation that has on average 10,82% less routine work than a person without tertiary education. The magnitude of the coefficient remains relatively consistent over the five different models. The largest variation I observe is in the standard errors. This variation is responsible for the level of significance across the models. In general, people with higher education are less at risk of automation.

Table 2: Incremental regression

	(1) OLS	(2) Clustering	(3) Country FE	(4) Year FE	(5) Country-Year FE
Heduc	-0.2009***	-0.2147	-0.2164*	-0.2147	-0.2163*
	(0.0056)	(0.0740)	(0.0630)	(0.0706)	(0.0618)
Age	-0.0006*	-0.0007	-0.0008	-0.0007	-0.0007
	(0.0003)	(0.0006)	(0.0006)	(0.0005)	(0.0005)
Gender	0.1071***	0.1203	0.1178	0.1203	0.1178
	(0.0052)	(0.0967)	(0.0940)	(0.0949)	(0.0918)
Mother Heduc	-0.0508***	-0.0539*	-0.0522*	-0.0537	-0.0517
	(0.0065)	(0.0164)	(0.0153)	(0.0171)	(0.0163)
Birthplace	-0.0218**	-0.0117	-0.0084	-0.0114	-0.0084
	(0.0085)	(0.0202)	(0.0261)	(0.0195)	(0.0252)
HH netincome	-0.0095***	-0.0102	-0.0105	-0.0102	-0.0104
	(0.0010)	(0.0069)	(0.0058)	(0.0067)	(0.0058)
Heduc %	-0.0046***	-0.0050	-0.0016	-0.0053*	0.0050
	(0.0004)	(0.0018)	(0.0041)	(0.0015)	(0.0201)
R&D	-0.0000***	-0.0000	-0.0000	-0.0000*	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0002)
Cons	-0.2341***	-0.2298	-0.3643***	-0.2184	-0.5950
	(0.0218)	(0.1411)	(0.0068)	(0.1513)	(0.7141)
N	58049	55225	55225	55225	55225

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This general effect of education on the RTI provides a first impression with respect to sign and magnitude. However, the analysis will not conclude here, as it will investigate further how the effect varies across different European countries. The division by country aims to illustrate that the impact of education is not purely endogenous but can also be enhanced or deteriorated by structural factors and policy decisions. Therefore, to enhance the understanding of country-level differences, I will now explore the question of how efficiently countries allocate their human capital resources.

The previous model from column (4) in table 2 is utilized, while extending it with country-specific effects, which are implemented directly in the model by interacting education with the country variable as shown in equation 5. Through this interaction, the comparison between countries becomes central. Therefore, the coefficients determined by the interaction term indicate the efficiency of the allocation of human capital within a country. A more negative value suggests that countries allocate their highly skilled workers more strongly to occupations that contain fewer routine tasks. This results in a more efficient allocation of human capital, since workers are utilizing their skills effectively.

A prerequisite to be able to allow a more distinguished matching between non-routine occupations and high-skilled workers is a labor market that provides sufficient occupations for high-skilled workers. Another requirement is that the labor market comprises occupations that consist of non-routine tasks. Through the latter channel, countries obtain the possibility of influencing their labor market towards more automation robust occupations. Therefore, the disparities in the following results will reflect structural differences between countries. As extracting information from a table of 18 countries is rather cumbersome, I present the regression coefficients graphically in figure 2:

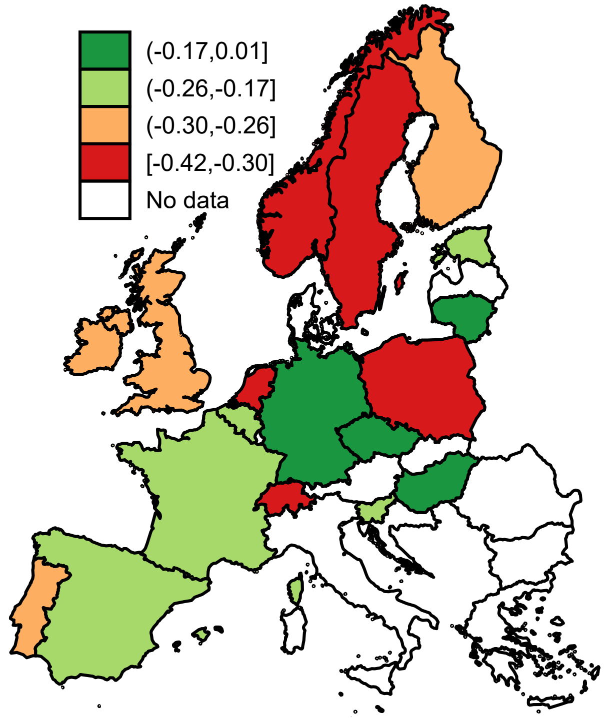


Figure 2: Coefficients of the importance of education obtained for each country

This map shows the magnitudes of the coefficients from regression 5. The red and orange countries exhibit a more negative coefficient. A more negative coefficient reflects a greater impact of higher education. Setting it in the context of the cross-country comparison, the red and orange countries exhibit a more efficient allocation of human capital. Countries colored in green reflect a lower importance of higher education, indicating that the labor market provides fewer opportunities for high-skilled workers.

In general, several countries exhibit a coefficient that could have been assumed on the basis of well-known regional European differences (Szymańska, 2021). Scandinavian countries display a more efficient allocation of human capital, as well as Great Britain and Ireland. Two other leading countries in higher education, Switzerland and the Netherlands, are as well among the countries with the lowest coefficient. The highest coefficients were found for Germany, the Czech Republic, and Hungary, indicating a less efficient allocation of human capital. One result is striking; Poland's coefficient is on the same level as the Scandinavian ones, except that Poland's coefficient is not significant. In conclusion, there exist patterns with respect to efficiency of human capital allocation, which perpetuate already existing divides in the European Union between North and certain wealthy countries, like Switzerland, versus Eastern and Southern countries.

The observation of regional patterns induced me to construct country bins. In particular, creating bins of similar countries represents the next step. The aim of this binned analysis is to emphasize that for most countries a spatial contiguity occurs. The country bins consist of North (Finland, Sweden, Norway), South (Spain, Portgual), West (Belgium, Germany, France, Netherlands, Swiss), Islands (United Kindom, Ireland) and East (Slovenia, Hungary, Chzechia, Lithuania, Estonia, Poland).

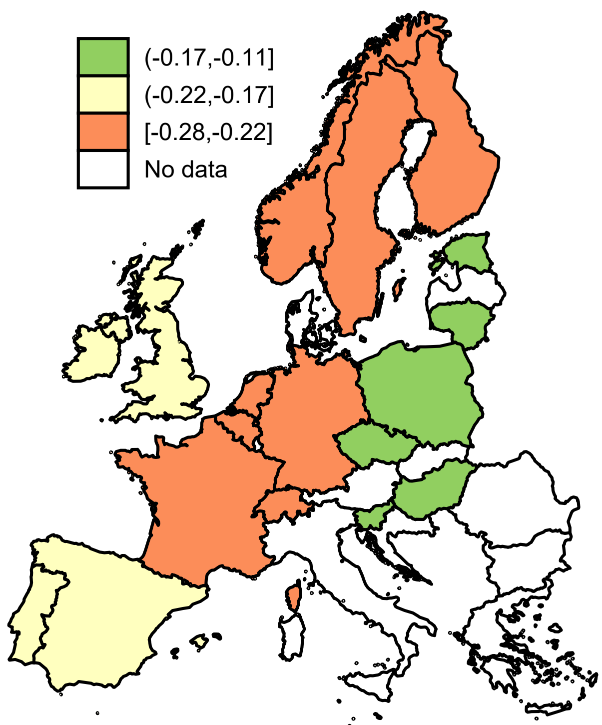


Figure 3: Coefficients from pooled regression

Figure 3 shows the variation of the coefficients between the country bins. Overall, the bin coefficients fluctuate around the value obtained prior in the OLS estimation, which was around -0.21. To elaborate on the results in more detail, North and West display a slightly larger coefficient with -0,28, respectively, -0,23, whereas the Islands and East have a lower coefficient at -0,21. To translate the magnitude into a percentage, high-skilled workers in the North experience on average a 14% lower RTI than low-skilled workers. East shows the lowest value with an effect of education of -0,14 or 7%[4]. Thus, in countries in the East, the effect of workers' education is only half as large as in countries in the North. An insight from this analysis is that differences in the EU persist in regional patterns.

One detail deserves more attention here. This binned approach conceals some heterogeneity. Three outliers are now hidden and exert an upward influence on their neighbors. Namely, Switzerland, The Netherlands, and Poland. It is important to make a comment alongside the results, as, particularly in western Europe, the divergence is more pronounced than initially perceived in figure 3. In figure 2 the heterogeneity in West Europe is revealed, where Germany is quite contrary to Switzerland or the Netherlands. Overall, it is evident that regional patterns exist, except for the mentioned cases, where some regions outperform other neighboring countries in terms of efficiency in human capital allocation.

This analysis provides a more comprehensive understanding of the extent to which countries contribute to the overall treatment effect. Additionally, it gives us a first hint on which countries possess a more resilient labor market in anticipation of future changes.

4.2 Dynamic Results

Knowing the static position of the labor market performance of a country is already interesting, but even more information can be extracted by exploring the dynamics. The dynamics are examined by plotting the relation between the RTI of a country in 2012 and its average change rate until 2018. The aim is to record changes across countries to gain a deeper understanding of current labor market trends.

[4] Result table can be found in Appendix

The dynamic analysis is depicted by figure 4, which shows on the horizontal axis the initial average value of the RTI and on the vertical axis the average change rate from 2012 to 2018. Perhaps a bit counterintuitive at this point is that a negative change is the preferred outcome. Intuitively, a decrease in the RTI is favorable; thus, a negative change rate describes an improvement. To provide some context on how to interpret the actual change rate, the RTI ranges from -1 to 1 and a 0,02 average change equates to a 1% increase. This might seem like a negligible change, but even marginal changes can yield substantial differences in the long run.

To illustrate an example, Finland had a rather low RTI value in 2012 and has experienced a decline, indicating a change in the labor market toward more occupations, including a low RTI. Therefore, workers with higher education can utilize their skills more effectively, by opting for an occupation that is better performed by a highly educated worker. The explanation why the decrease in Finland's RTI necessitates a shift in occupations can be attributed to the consistent allocation of RTI to occupations. The composition of tasks within an occupation remains unchanged throughout the observed period, thus a change in a country's RTI is caused by a change in occupations. In contrast, the Czech Republic had a high RTI value in 2012 and experienced a further increase until 2018. This suggests that education is not as distinguishing for the RTI of workers. Furthermore, there is an increasing share of occupations in the Czech Republic with more routine tasks, indicating a worsening of the situation.

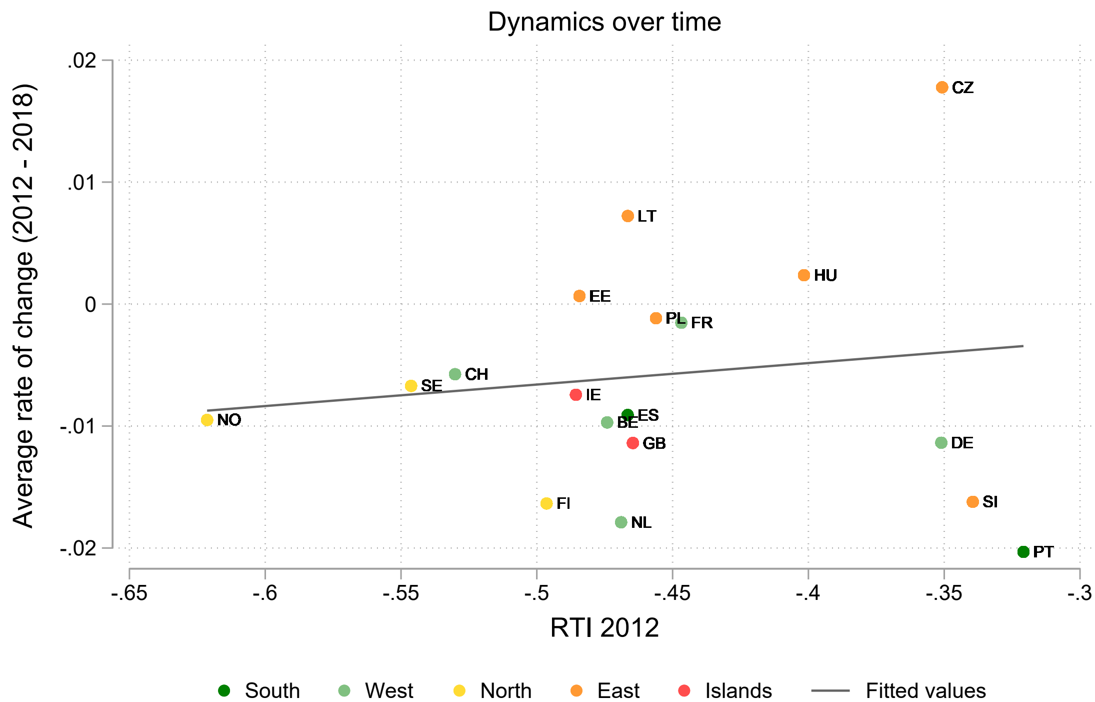


Figure 4: Dynamic analysis of changes in the RTI across countries. Countries are pooled geographically.

Another insight provided by figure 4 is that, in general, there exists a positive correlation between the RTI value in 2012 and its average change until 2018. Consequently, a higher value of RTI in 2012 is more likely to lead to a further increase in RTI. On the other hand, a lower RTI value in 2012 led to a decrease in RTI until 2018. Three exceptions are Germany, Slovenia, and Portugal. They have experienced a decrease in RTI, even though they started with a higher RTI in 2012.

To emphasize the occurrence of spatial contiguity, the country bins are highlighted in different colors. This facilitates the observation that regional patterns occur not only in static analysis but also in dynamic analysis, in which certain clusters are noticeable. For instance, the Scandinavian countries exhibit the lowest RTI values in 2012 and are undergoing a further decline in RTI. In proximity to Scandinavian countries, we can find

Switzerland and the Netherlands, which have been outliers among western countries. Ireland and Great Britain, constituting the Island bin, are found in the middle. In orange, we see the countries in the Eastern bin. They are predominantly allocated in the upper right corner, indicating a higher value of RTI in 2012 and a subsequent increase in RTI. On the other hand, countries in the West bin form the most scattered group in terms of their starting value, but all have experienced a decrease in RTI.

One limitation of this analysis pertains to the short time frame. As discussed in this section, the actual change rates were relatively modest. However, a major shift in the RTI in such a short period would have been suspicious. Capturing the shifts in labor market structures, caused by new technologies, requires long-term observations. What might seem minor now, can turn out substantial in the long run.

5 Robustness

5.1 Heterogeneities

The results so far delivered insights into differences between countries, but open for investigation is whether they are robust. Thus, various robustness checks are conducted in the next section. The first step consists of examining population heterogeneities to find out whether certain groups of society are especially prone to the impacts of automation (Imbens and Wooldridge, 2009). This is of particular interest since potential policies can be targeted more precisely, and thus more efficiently. The emphasis will remain on individuals with different socioeconomic characteristics, such as gender and net income of households. In a second step, an alternative estimation strategy will be employed to test the validity of the results obtained.

5.1.1 Heterogeneity in Gender

Heterogeneities among individuals have not been taken into account. This has implicitly imposed the assumption that individuals are homogeneous, which may prove unrealistic. The existence of the gender wage gap is no longer a secret. Therefore, it could become relevant to control for heterogeneities in gender. Investigating whether such structural differences also occur in the RTI by comparing the means of the RTI between men and women constitutes the first step.

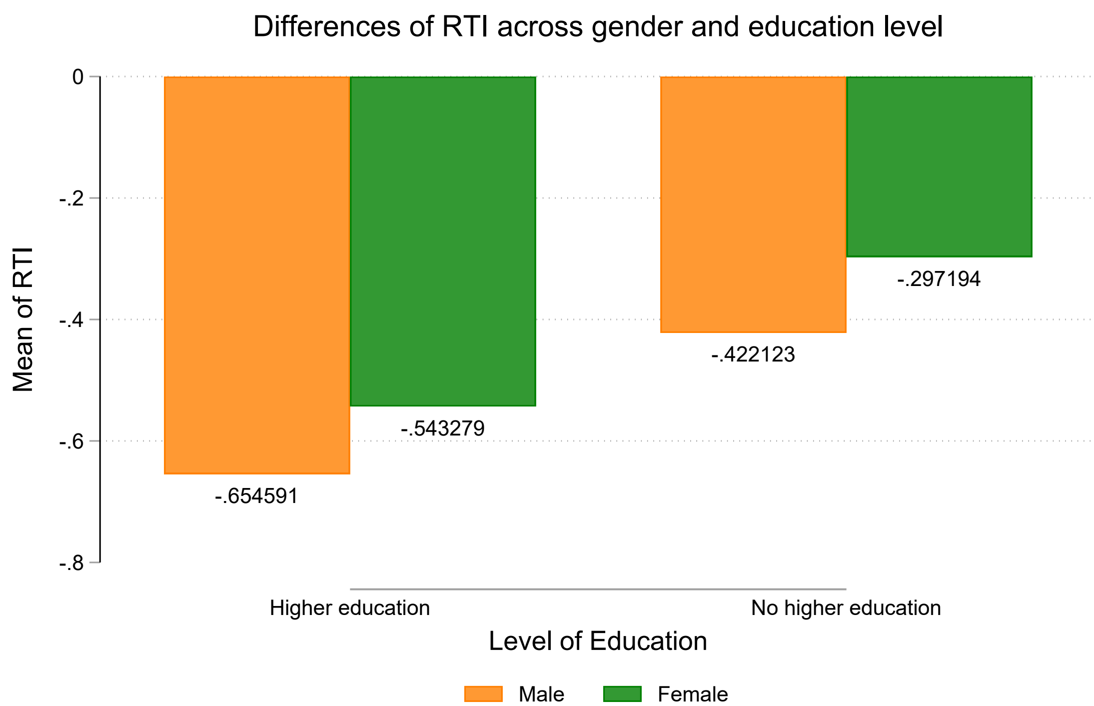


Figure 5: Bar graph of the average RTI of men and women across levels of education

Figure 5 shows the different results of RTI between men and women depending on the level of education. Women, regardless of their level of education, have a higher RTI than men. Although women have caught up and even overtaken in higher education degrees, they are behind with respect to RTI (Goldin, 2006). The RTI among highly educated workers is 0,11 higher for women than for men. In the non-highly educated population the disparity is even slightly greater with 0,12. This consistent gap is hinting towards the existence of factors, other than education, that are responsible for the level shift between men and women.

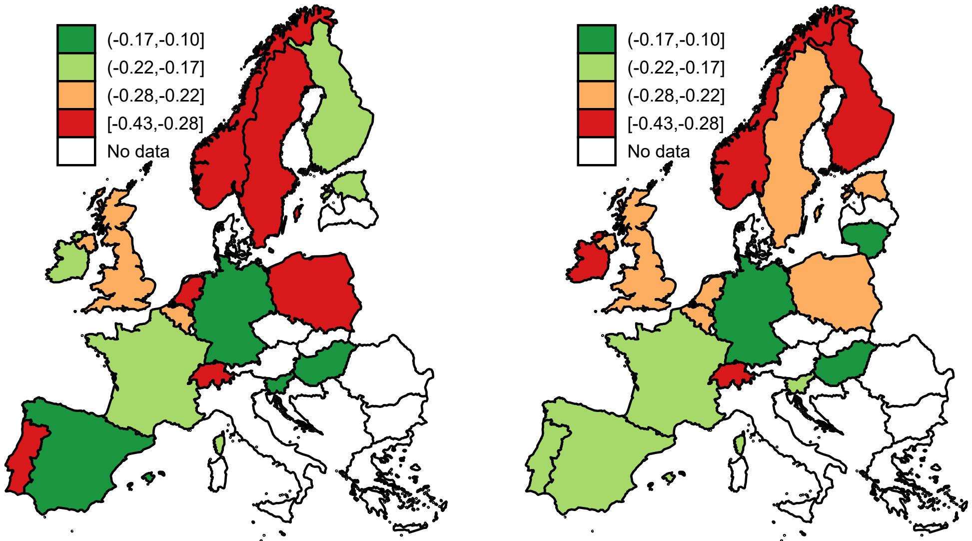
Given that a simple mean comparison neglects other circumstances, matching can help to account for some observable factors. In particular, propensity score matching is utilized in this case to obtain the average treatment effect of being identified with a certain gender. I use the propensity score, which calculates a person's probability of having a particular outcome based on a set of covariates and distinguishes people based on a treatment variable, in my case gender. The set of covariates consists of the same variables as used in equation 4. Individuals with narrow propensity scores are matched, and the resulting difference in RTI is the average treatment effect [5].

Based on the findings of the propensity score matching, a difference of 0,1064 remains between men and women after accounting for observable individual characteristics. Set in context with the range of RTI, the difference corresponds to a gap of 5,32% between men and women. These preliminary results suggest that there is presence of heterogeneity between men and women.

But this research aims not only to understand the individual level but also the country level. The main point of this analysis is to show that addressing the unequal access to the labor market requires regional considerations. Hence, it is of interest whether the impact of education on the RTI varies across countries. This is examined using a regression analysis as in equation 5, but run for women and men separately. The results for women are presented on the left side and those for men are presented on the right side. Recall that red implies a stronger reduction in RTI due to higher education, while green, and especially darker green, implies a smaller impact of education on the RTI. The change in perspective opens the door for a comparative analysis where we can now figure out if men and women experience different impacts in different countries.

After taking a look at the maps, it can be observed that there are variations across Europe with respect to the impact of education by gender. For example, in Portugal there exists a particularly pronounced gap. Women experience a strong impact of education, whereas

[5] Further information regarding the matching procedure can be found in the Appendix section about Heterogeneities 9.



(a) Coefficients for women

(b) Coefficients for men

Figure 6: Regression results for interacted gender, education and country

men only a moderate impact. The reverse occurs in Ireland, where women experience a smaller impact of education than men. As the coefficient values are grouped into four bins, some variation is absorbed in these graphs[6]. Countries like Ireland and Portugal, which show coefficients two bins apart, indicate that the impact of education is unequally distributed between men and women.

To summarize, there exists a considerable disparity in the representation of men and women along the RTI. Various methodological approaches have been employed, but the results consistently indicated a level difference in RTI between men and women. Even after controlling for several observable characteristics, a gap of 0,1064 remains in favor of men. In the second step, it could be shown that heterogeneities in the importance of higher education between men and women occur at the country level. When thinking about the impact of automation, it is important to bear in mind that the effect apparently affects men and women in different countries unevenly.

[6]A detailed table of the results can be found in the Appendix see table 4.

5.1.2 Heterogeneity in Income

The second part of my heterogeneity analysis consists of controlling for the relation between households' income, education, and RTI. Household income might play a role, as having access to more financial means provides people with the opportunity to get a better quality of education, which is usually more costly. High-quality education is often an indispensable step in achieving an automation resistant occupation. If one looks at the impact of education on RTI along the deciles, it indeed appears like household income decreases the RTI.

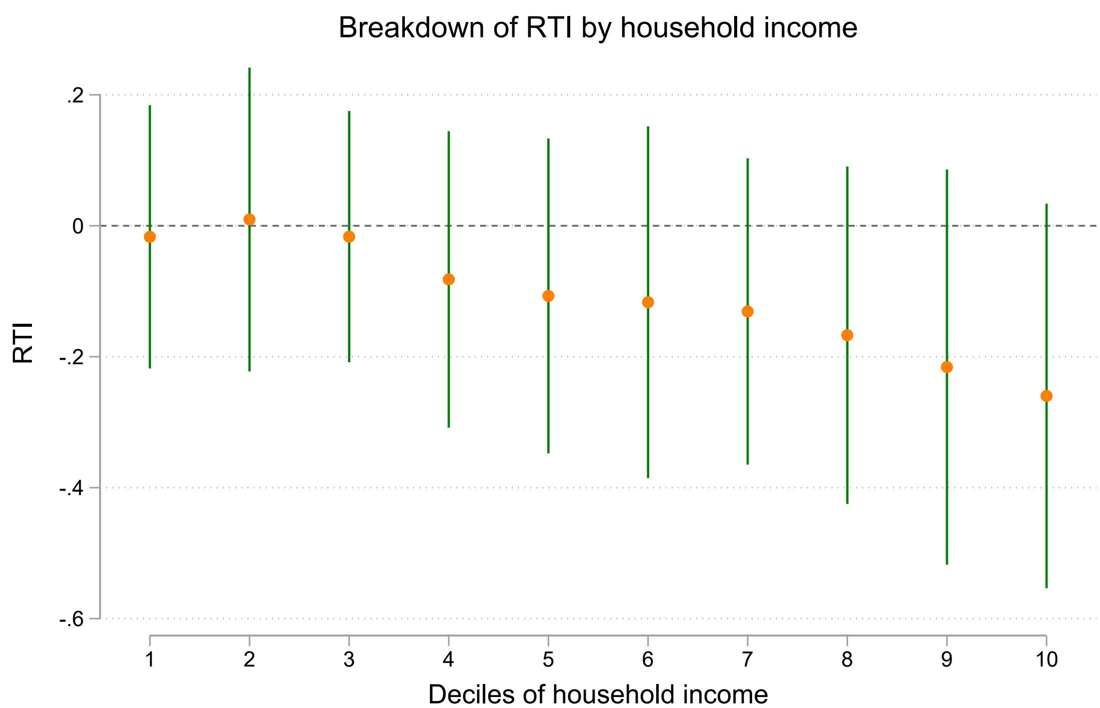


Figure 7: Results from regression interacting education with household income

Based on the result from figure 7, one could assume that higher income generally correlates with lower RTI. To control for this possibility, I turn back to the main question of my work, which is whether education matters for the risk of being exposed to automation.

Therefore, the sample is divided into people with tertiary education and those without. Of each subsample, the RTI over the different income deciles is plotted. Therefore, the different impact of income levels depending on the level of education on RTI can be observed.

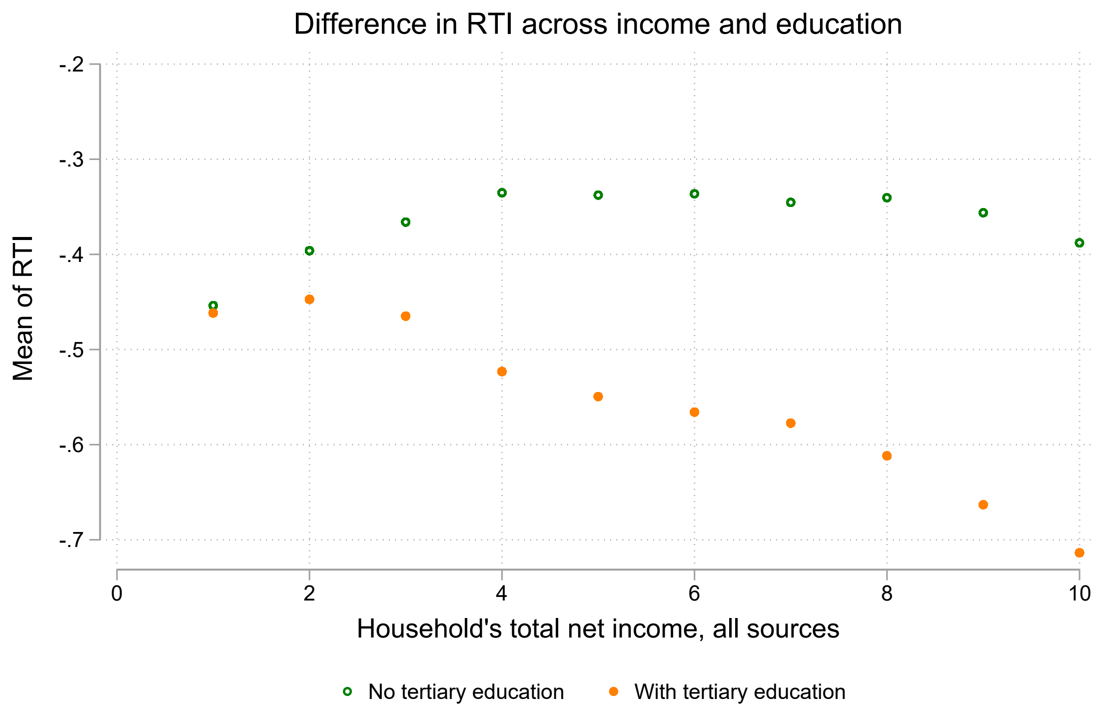


Figure 8: Average RTI depending on the level of household income and education

The insight here is that if a person has attained tertiary education, the RTI is decreasing while increasing household income. On the other hand, the RTI of a person without tertiary education remains almost constant across income levels. Interestingly, both education groups commence at the lowest deciles around the same RTI and are just diverging with an increasing income level. It indicates that highly educated individuals in the higher deciles are considerably less affected than those lacking tertiary education or highly educated workers with a lower level of household income.

In conclusion, it seems like there is a glass ceiling for people without tertiary education, whereas people with tertiary education experience a decrease in RTI when household income increases. The downward trend in RTI of highly educated people with an increasing household income indicates that higher income correlates negatively with RTI, but only if they are highly educated.

5.2 Model Specification

5.2.1 Binary model

In researching the question of the importance of education for the risk of automation of occupations, I also want to ensure that my results are robust to different specifications. The distribution of the outcome variable is skewed, with 38% of the RTI values being allocated at -1 [7] Therefore, the next step consists of testing a binary specification.

Creating a binary variable that takes on the value 1 if RTI is -1 and 0 otherwise. To deal with the binary character of the dependent variable, I use a logit model. The included variables remain the same as in equation 4 to ensure comparability.

The higher education coefficient gives the probability of not having any routine tasks involved in your occupation if you have higher education, in contrast to, no higher education. As the logit model exhibits a nonlinear link, the effect is likely not identical for all individuals, therefore, an average is formed. The coefficient obtained is the so-called average marginal effect. The coefficient is based on the average of all predicted changes by a one-unit change of everyone. In my case, it is the probability of being in an occupation without any routine tasks if you have attained tertiary education.

[7] Figure 14 shows the distribution of the RTI and can be found in Appendix 9.

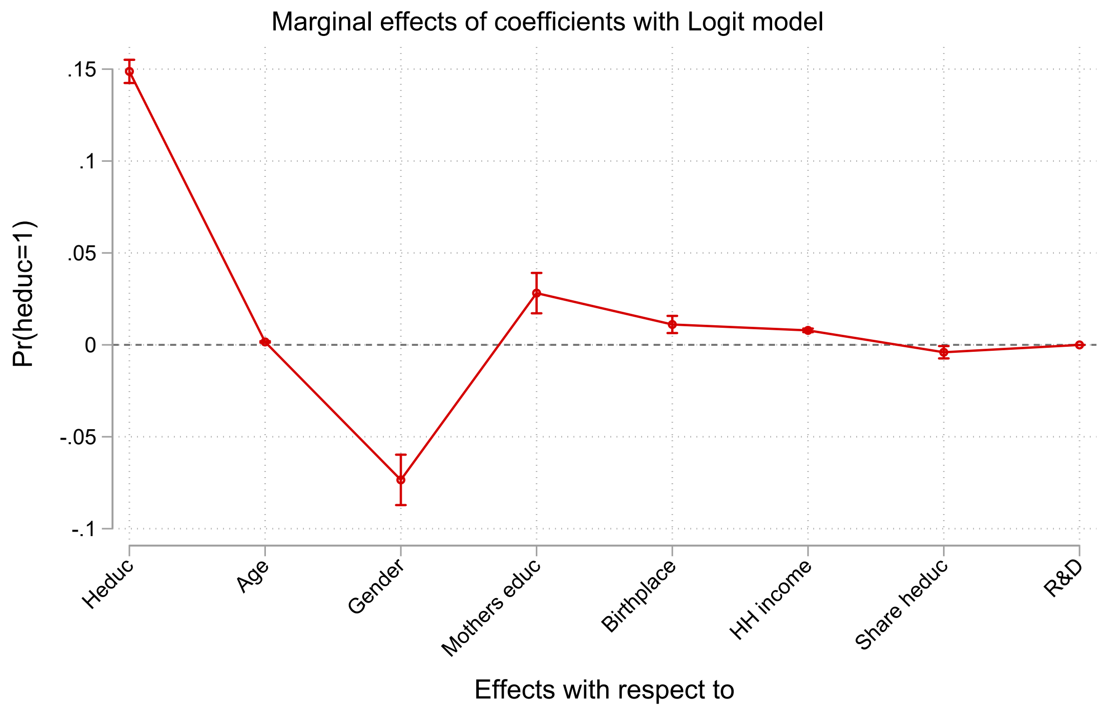


Figure 9: Marginal effects at the average of covariates in a binary dependent model specification using a logit model.

It is noticeable that higher education has the strongest positive impact on the probability of having an occupation that contains no routine tasks. The model specification changed the outcome variable, and consequently the sign. The positive coefficient in the binary specification is in its meaning equal to the negative coefficient in the OLS model. The magnitude is now slightly lower than before, displaying a value of 0,1428. In contrast to the OLS specification, the magnitude is now a third smaller. Except for gender, no other variable shows considerable impact. Albeit the effect is a bit smaller than in the OLS model, the results are still pointing in the same direction.

5.2.2 Quantile Regression

The next and last step in my model specification section is a quantile regression. A coefficient in a quantile regression does not represent the mean but the median. Thus, the coefficient is more robust against outliers. As mentioned above, the data are not evenly distributed, but heavily skewed. Therefore, this model specification can address this concern. In contrast to OLS, where residuals are minimized, quantile regression targets minimizing the sum of absolute errors (Koenker and Bassett Jr, 1978 & Petscher and Logan, 2014).

In table 3 are the numerical results of the model specifications presented. Column (1) shows the OLS model, and column (2) shows the OLS model including fixed effects and robust errors. These are already known from the base model. Column (3) yields the results from the logit model using the binary specification and column (4) is the quantile regression output.

The first three columns show the mean, whereas the last column shows the effect for individuals situated around the median. As my outcome variable is clustered around -1, the mean is susceptible to outliers. Initially, it might seem concerning that the logit model in column (3) results in contrast to the first two specifications in a smaller coefficient for higher education. However, the discrepancy can be attributed to its distinct interpretation based on the characteristics of the logit model. Namely, it indicates the probability change of the RTI being 1 if an individual is highly educated compared to those who are not. On the contrary, the coefficient of the quantile regression reveals a stronger impact of higher education. This can be explained by the characteristics of the model, where the median is less sensitive to outliers.

In general, the OLS model yields an effect somewhere between the logit model and quantile regression. The additional models demonstrate that the magnitude of the OLS model is positioned between the two alternative estimators. This leads to the conclusion that the OLS estimator does not produce implausible results.

Table 3: Overview of model specifications

	(1) OLS	(2) OLS+FE	(3) Logit	(4) Quantile
Heduc	-0.2009***	-0.2163*	0.1428***	-0.2867***
	(0.0056)	(0.0618)	(0.0163)	(0.0069)
Age	-0.0006*	-0.0007	0.0015***	-0.0011***
	(0.0003)	(0.0005)	(0.0004)	(0.0003)
Gender	0.1071***	0.1178	-0.0705***	0.1138***
	(0.0052)	(0.0918)	(0.0131)	(0.0055)
Mother Heduc	-0.0508***	-0.0517	0.0270***	-0.0284***
	(0.0065)	(0.0163)	(0.0059)	(0.0056)
Birthplace	-0.0218**	-0.0084	0.0107	-0.0153*
	(0.0085)	(0.0252)	(0.0104)	(0.0071)
HH netincome	-0.0095***	-0.0104	0.0076***	-0.0112***
	(0.0010)	(0.0058)	(0.0014)	(0.0011)
Heduc %	-0.0046***	0.0050	-0.0039	-0.0029***
	(0.0004)	(0.0201)	(0.0043)	(0.0005)
R&D	-0.0000***	-0.0000	-0.0000	-0.0000***
	(0.0000)	(0.0002)	(0.0000)	(0.0000)
Cons	-0.2341***	-0.5950		-0.4768***
	(0.0218)	(0.7141)		(0.0246)
N	58049	55225	58049	58049

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6 Discussion

In the following, the results obtained are discussed in a broader context along with the fundamental concerns they raise. Furthermore, I will elaborate on how this study provides an advance in handling the consequences of automation with a greater awareness of potential obstacles.

The type of tasks executed is closely associated with an individual's level of skill. Over recent decades, the importance of education has increased, as new technologies demand higher skills. In the last few years, the importance of non-routine tasks have gained attention. New technologies can replace an increasing number of routine tasks. As shown in this research, the amount of RTI in an occupation is linked to educational attainment. Therefore, an intertwined effect of SBTC and RBTC could be assumed. The new developments predicted by the RBTC possess the capability to widen the gap between high- and low-skilled workers. One potential consequence of this divergence is a political destabilization of countries and the European Union (Kurer and Palier, 2019). This argument is based on an emerging income gap, which can create a feeling of being left behind among those who experience a relative income loss.

The individual-level analysis showed that higher education results in a significant reduction in RTI. Although there exist variations between the individual's level of RTI, such as between men and women, education consistently led to a decrease in the routine intensity. Since gender RTI gaps vary between countries, they likely reflect differences in society's norms of gender. On the contrary, other socioeconomic variables, such as the level of household income, show an impact only if an individual is highly educated. By this analysis, it is shown that education has a reducing impact on the RTI, but the level at which the effect takes place varies.

The country-level analysis provides another perspective on the situation in Europe. Differences in higher education coefficients in Europe reflect varying levels of efficiency in the allocation of high-skilled workers. The differences could be explained by at least two channels. First, there is a lack of investment in the infrastructure required for the implementation of new technologies. The demand for infrastructure is unlikely to diminish in the future. This results in an outdated labor market, with a lower wage growth rate, which will not attract productive workers and could trigger brain drain toward more prosperous occupation outlooks. The second channel is the lack of high-skilled labor. Even if the supply in occupations with more non-routine tasks rises, it can occur that there is a lag until sufficient people receive the necessary skills (Bardhan et al., 2013). This could reinforce long-term struggles for countries that are not catching up. The dynamic analysis has indicated such a trend: Countries with a higher RTI have experienced a growth in RTI until 2018. Ultimately, divergence among European countries can pose a threat to cohesion in the EU. These results should be considered as preliminary, since labor market developments take time and at this point many questions cannot be solved yet.

Higher education enriches workers' skills, allowing them an effective execution of non-routine tasks. On an individual level, the RTI will be a distinctive feature in determining differences in labor market outcomes. From a country-level perspective, it seems advisable to invest in higher education to meet the demand of the labor market. Simultaneously, investing in the infrastructure to provide a sufficient supply of automation-resistant occupations is indispensable. High-skilled workers will be decisive for the competitiveness of a country's economic productivity in the future.

7 Conclusion

In summary, this research highlights the importance of higher education and provides insights that contribute to the growing literature, it can inform policy decisions, and ultimately contribute to creating a more resilient and equitable labor market for the future. On the basis of the SBTC and RBTC hypotheses, the results underline the importance of higher education. I have examined the effects at the individual and country level, as well as static and dynamic effects.

At the individual level, the results reveal that higher education leads to an average reduction in RTI of 10,82%. The wave of new technologies affects people heterogeneously. For example, women's occupations have on average an RTI that is 5% higher than that of men. The effect persisted even after turning from an OLS model to a logit model and a quantile regression approach. The results of this study corroborate that higher education has a reducing effect on RTI.

A deeper examination of the role of education was achieved by extension at the country level. The results of the static and dynamic analysis revealed regional disparities between the EU member states. Some countries face challenges in providing automation-resilient occupations that may impede their progress. On the contrary, other countries have already developed labor markets that ensure future competitiveness. This potentially results in a divergence in economic prospects between EU countries. Consequently, a lack of cohesion within the EU could be exacerbated.

A brief note on the anticipated evolution of the impact of education is provided. A notable characteristic of the current wave of technology is the shift of tasks rather than the replacement of entire occupations. There is an inherent potential to enhance productivity, but sufficient human capital is an essential requirement to handle the new technologies. Although an increasing share of tasks can be automated, there are still technological bottlenecks in tasks that cannot be substituted yet (Frey and Osborne, 2017). These bottlenecks contain complex non-routine tasks that are performed mainly by high-skilled workers. Therefore, higher education plays a detrimental role, either through the channel of using technology complementary or by maintaining irreplaceability. Therefore, higher education is expected to gain further importance in the future (Autor, 2015).

As we face an uncertain future characterized by ongoing technological changes, the importance of higher education for labor market outcomes becomes increasingly evident. To mitigate adverse effects such as brain drain and regional disparities, decisive action is imperative. The goal should be a rapid transition to a labor market characterized by a substantial share of non-routine occupations and effort to upskill people with higher education.

8 References

Acemoglu, D. (1998). Why do new technologies complement skills? directed technical change and wage inequality. *The quarterly journal of economics*, *113*(4), 1055–1089.

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics* (pp. 1043–1171). Elsevier.

Angrist, J. D., & Krueger, A. B. (1999). Empirical strategies in labor economics. In *Handbook of labor economics* (pp. 1277–1366). Elsevier.

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.

Autor, D. (2022). The labor market impacts of technological change: From unbridled enthusiasm to qualified optimism to vast uncertainty. *National Bureau of Economic Research.*

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.

Balsmeier, B., & Woerter, M. (2019). Is this time different? how digitalization influences job creation and destruction. *Research policy*, *48*(8), 103765.

Bardhan, A., Hicks, D. L., & Jaffee, D. (2013). How responsive is higher education? the linkages between higher education and the labour market. *Applied Economics*, *45*(10), 1239–1256.

Barro, R. J., & Sala-i-Martin, X. (1990). Economic growth and convergence across the united states. *National Bureau of Economic Research.*

Bessen, J. E. (2016). How computer automation affects occupations: Technology, jobs, and skills. *Boston Univ. school of law, law and economics research paper*, (15-49).

Bittarello, L., Kramarz, F., & Maitre, A. (2018). *The task content of occupations* (tech. rep.). IZA Discussion Papers.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of human resources*, *50*(2), 317–372.

Chmielewski, A. K. (2019). The global increase in the socioeconomic achievement gap, 1964 to 2015. *American sociological review*, *84*(3), 517–544.

David, H. (2013). The "task approach" to labor markets: An overview. *Journal for Labour Market Research*, *46*(3), 185–199.

Dengler, K., Matthes, B., & Paulus, W. (2014). Occupational tasks in the german labour market. *FDZ Methodenreport*, *12*.

Frey, C. B., & Osborne, M. (2013). The future of employment. *Oxford Martin Programme on Technology and Employment*.

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.

Gallego, A., & Kurer, T. (2022). Automation, digitalization, and artificial intelligence in the workplace: Implications for political behavior. *Annual Review of Political Science*, *25*, 463–484.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American economic review*, *96*(2), 1–21.

Goldin, C., & Katz, L. F. (1996). Technology, skill, and the wage structure: Insights from the past. *The American Economic Review*, *86*(2), 252–257.

Górka, S., Hardy, W., Keister, R., & Lewandowski, P. (2017). Age, tasks and skills in european labour markets.

Hardy, W., Keister, R., & Lewandowski, P. (2018). Educational upgrading, structural change and the task composition of jobs in europe. *Economics of Transition*, *26*(2), 201–231.

Haslberger, M. (2021). Routine-biased technological change does not always lead to polarisation: Evidence from 10 oecd countries, 1995–2013. *Research in Social Stratification and Mobility*, *74*, 100623.

Haslberger, M. (2022). Rethinking the measurement of occupational task content. *The Economic and Labour Relations Review*, *33*(1), 178–199.

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of economic literature*, *47*(1), 5–86.

Katz, L. F., et al. (1999). Changes in the wage structure and earnings inequality. In *Handbook of labor economics* (pp. 1463–1555). Elsevier.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, 33–50.

Kurer, T., & Palier, B. (2019). Shrinking and shouting: The political revolt of the declining middle in times of employment polarization. *Research & Politics*, *6*(1), 2053168019831164.

Maier, T. (2022). Change in occupational tasks and its implications: Evidence from a task panel from 1973 to 2011 for western germany. *Quality & Quantity*, *56*(3), 889–921.

Manning, A., Goos, M., & Salomons, A. (2009). Job polarization in europe. *American Economic Review: Papers and Proceedings*, *99*(2).

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

Mihaylov, E., & Tijdens, K. (2019). Measuring the routine and non-routine task content of 427 four-digit isco-08 occupations. *Tinbergen Institute Discussion Paper*.

Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The review of Economics and Statistics*, 334–338.

Petscher, Y., & Logan, J. A. (2014). Quantile regression in the study of developmental sciences. *Child development*, *85*(3), 861–881.

Pouliakas, K. (2018). Determinants of automation risk in the eu labour market: A skills-needs approach. *IZA discussion paper*.

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.

Szymańska, A. (2021). Reducing socioeconomic inequalities in the european union in the context of the 2030 agenda for sustainable development. *Sustainability*, *13*(13), 7409.

Verbeek, M. (2008). *A guide to modern econometrics*. John Wiley & Sons.

9 Appendix

The appendix provides further information on the country-level analysis, but consists mostly of notes on the robustness part of this research.

Results

In the figure below, the numerical coefficients are shown in ascending order for each country from the interacted regression 5. The numerical results of the coefficient plot are presented in Table 4. The coefficients show the result for the interaction between higher education and the country and the individual- and country-level covariates.

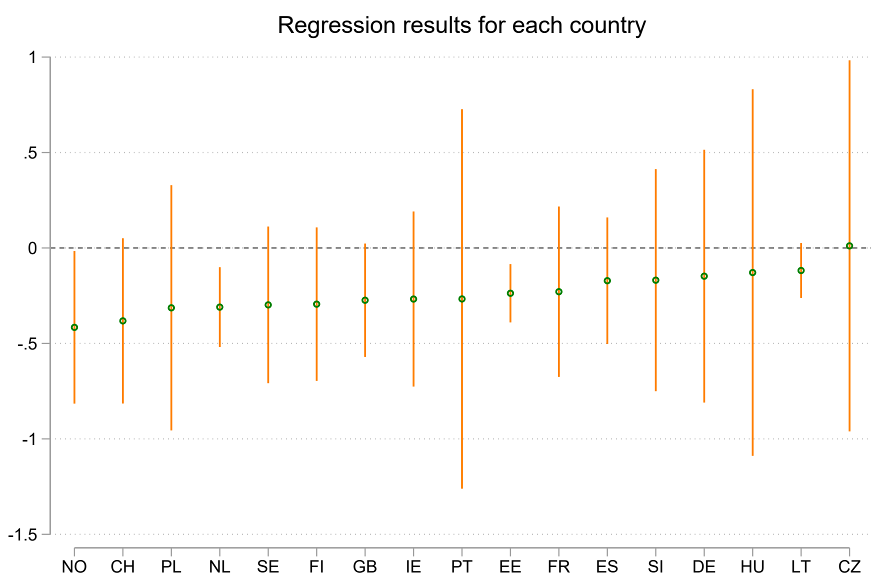


Figure 10: Plot of coefficients of higher education across countries

Table 4: Results from interacted regression between education and country

RTI			
Belgium	-0.2548**	Norway	-0.4158*
	(0.0409)		(0.1256)
Switzerland	-0.3819	Poland	-0.3133
	(0.1360)		(0.2018)
Czech Republic	0.0112	Portugal	-0.2670
	(0.3054)		(0.3123)
Germany	-0.1477	Sweden	-0.2981
	(0.2081)		(0.1290)
Estonia	-0.2372*	Slovenia	-0.1687
	(0.0480)		(0.1828)
Spain	-0.1714	Age	-0.0008
	(0.1041)		(0.0006)
Finland	-0.2942	Gender	0.1167
	(0.1263)		0.0930
France	-0.2291	Mothers educ.	0.0526
	(0.1402)		(0.0189)
Great Britain	-0.2738	Birthplace	-0.0102
	(0.0933)		(0.0282)
Hungary	-0.1287	HH net income	-0.0102
	(0.3017)		(0.0059)
Ireland	-0.2675	Share Heduc	0.0053
	(0.1441)		(0.0203)
Lithuania	-0.1180	R&D pc ppp	-0.00
	(0.0451)		(0.0002)
Netherlands	-0.3097*		
	(0.0656)		
Observations	55225		

Robustness

Heterogeneity

The numerical results of the analysis of gender heterogeneity are presented in table 5. The regression is run twice, once for women and once for men.

Table 5: Result for women and men separatly from regression.

	RTI	
	Women	Men
Heduc	-0.2326*	-0.1976
	(0.0575)	(0.0776)
Age	-0.0004	-0.0011*
	(0.0009)	(0.0003)
Mothers educ	-0.0538	-0.0486
	(0.0206)	(0.0159)
Birthplace	-0.0273	0.0097
	(0.0358)	(0.0253)
HH net income	-0.0042	-0.0170
	(0.0038)	(0.0080)
Share heduc	0.0008	0.0092
	(0.0252)	(0.0187)
R&D pc ppp	-0.0000	0.0000
	(0.0002)	(0.0002)
Constant	-0.2335	-0.5993
	(0.8671)	(0.6478)
N	28590	26635

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

It is striking that women experience a stronger and more significant effect than men. One conjecture that explains why this result occurs is the relative disadvantage experienced by lower educated women. The relative disadvantage emerges through the larger gap between less educated men and women than between highly educated people.

Table 4 comprises the coefficients of regression 5 that interact country with education, but are run separately by gender. The coefficients in the table reflect differences by gender between countries. These coefficients are the numerical values of the maps 6.

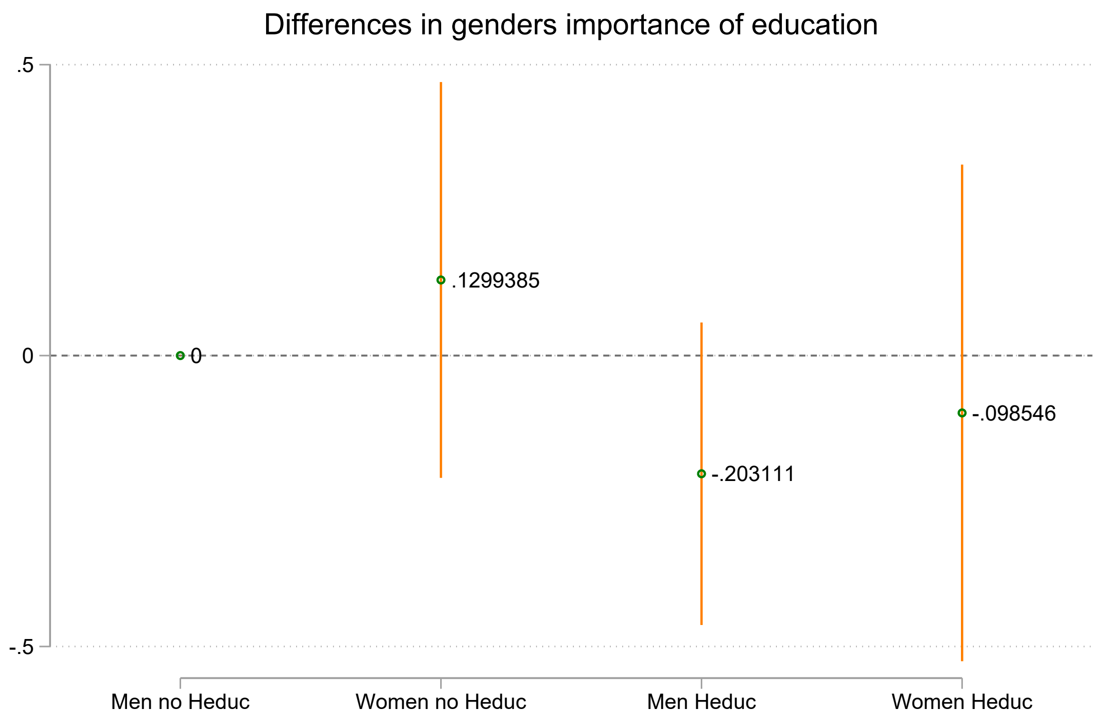


Figure 11: Plot of coefficients obtained by regression including interaction term between education and gender

Table 6: Impact of education for women and men across countries

	Men	Women
Belgium	-0.2503*	-0.2432**
	(0.0613)	(0.0351)
Switzerland	-0.3770*	-0.3957*
	(0.1064)	(0.0802)
Czech Republic	0.0197	0.0041
	(0.2815)	(0.3771)
Germany	-0.1373	-0.1340
	(0.1744)	(0.2334)
Estonia	-0.2691*	-0.2038
	(0.0533)	(0.0741)
Spain	-0.1826	-0.1577
	(0.1019)	(0.1009)
Finland	-0.3655	-0.2019
	(0.1166)	(0.1336)
France	-0.2175	-0.2026
	(0.1028)	(0.1040)
Great Britain	-0.2700	-0.2385
	(0.1053)	(0.1080)
Hungary	-0.1667	-0.1011
	(0.2653)	(0.3551)
Ireland	-0.3445	-0.1672
	(0.1456)	(0.1825)
Lithuania	-0.1697	-0.0466
	(0.0684)	(0.0776)
Netherlands	-0.2469*	-0.3434*
	(0.0740)	(0.0632)
Norway	-0.3902	-0.4321
	(0.1242)	(0.1489)
Poland	-0.2651	-0.3429
	(0.1832)	(0.2382)
Portugal	-0.2013	-0.3032
	(0.2744)	(0.3712)
Sweden	-0.2569	-0.2828
	(0.1155)	(0.1134)
Slovenia	-0.1739	-0.1465
	(0.1470)	(0.1992)
Observations	26635	28590

The question of how disparities between men and women vary across Europe has attracted my attention. Therefore, the difference between the coefficients of men and women is calculated and plotted on a map. After figure 12, table 7 with the difference computations is shown.

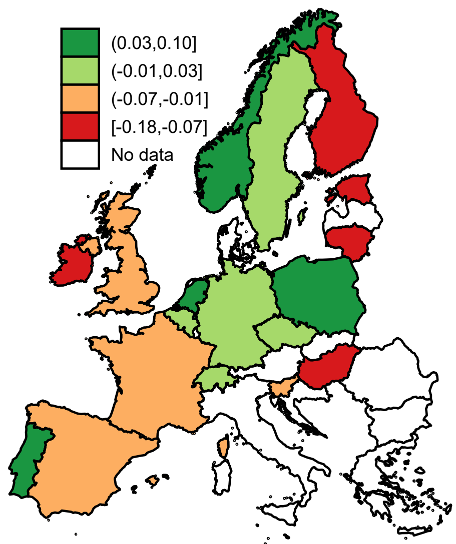


Figure 12: Differences in the impact of higher education for women and men.

In the map, the colors clearly reveal the differences between the countries. If the difference is negative, men experience a stronger impact of education on RTI. Conversley, a dark green colored country signals that the difference is positive and that women experience a greater impact of education. While western and most northern European countries show a stronger importance of education for women, the reverse result emerges for countries colored orange and red. However, regional patterns are more indistinct than the former results.

Table 7: Numerical differences between men and women across countries

Country	Men	Women	Difference
Belgium	-0.2503	-0.2432	-0.0071
Switzerland	-0.3770	–0.3957	0.0187
Czech Republic	0.0197	0.0041	0.0156
Germany	-0.1373	-0.134	-0.0033
Estonia	-0.2691	-0.2038	-0.0653
Spain	-0.1826	-0.1577	-0.0249
Finland	-0.3655	-0.2019	-0.1636
France	-0.2175	-0.2026	-0.0149
Great Britain	-0.2700	-0.2385	-0.0315
Hungary	-0.1667	-0.1011	-0.0656
Ireland	-0.3445	-0.1672	-0.1773
Lithuania	-0.1697	-0.0466	-0.1231
Netherlands	-0.2469	-0.3434	0.0965
Norway	-0.3902	-0.4321	0.0419
Poland	-0.2651	-0.3429	0.0778
Portugal	-0.2013	-0.3032	0.1019
Sweden	-0.2569	-0.2828	0.0259
Slovenia	-0.1739	-0.1465	-0.0274

Matching

In addition, I want to provide some additional information about the matching procedure. Since matching procedures imply many subjective choices, I want to state the choices made explicitly. First, I used the teffects psmatch command in STATA. The propensity score has been constructed based on higher education, higher education of mothers, age groups[8], birthplace, household net income, country, and year. The outcome variable is RTI and the treatment variable gender. Furthermore, I have implemented nearest neighbor matching of one. I tried several other specifications as well, and all displayed a similar magnitude in the treatment effect, indicating that my choices

[8]I have chosen age groups as the exact age would be very limiting. The age groups are clustered into three bins: workers below 36 years of age, workers between 37 and 50 years of age, and workers between 50 and 60 years of age.

are not very influential for the average effect.

The balance plot indicates that matching reduces the differences in observables between men and women in the raw and matched sample. For most RTI values, the outcome is similar for men and women, except around -1, where a larger gap exists. In the matched sample, the gap between men and women becomes less pronounced around -1, as figure 13 shows. This indicates that the matching of observables has absorbed some of the variation that distinguished men and women.

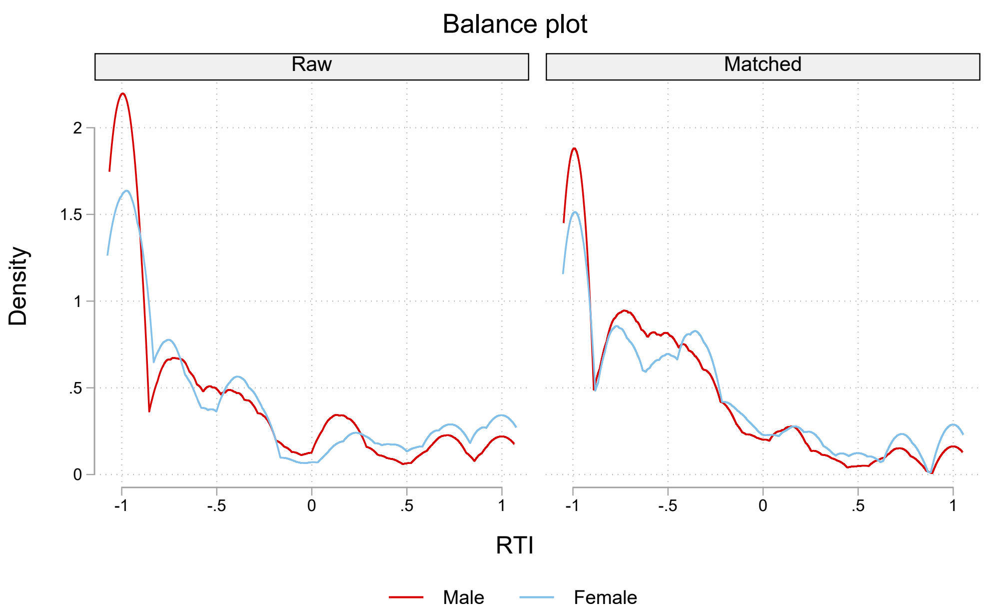


Figure 13: Difference between men and women before and after matching

Below in table 8, the difference in means and variance ratios is presented before and after the matching is conducted. Except in the case of higher education of mothers, all variables could be matched such that the difference is reduced.

Table 8: Numerical balance between men and women before and after matching

	Standardized diff.		Variance ratio	
	Raw	Matched	Raw	Matched
Heduc	.1402946	.004437	1.015231	1.000461
Mothers educ	-.0000443	.0002146	.999934	1.000322
Age groups	.0075817	.0027619	.9927149	.9936625
Birthplace	-.014034	.0069862	.9660753	1.018813
HH net income	-.1090616	.0050686	1.018092	.9893241
Country	.0505811	.0030022	.9650234	.995583
Year	.0062797	-.0005835	1.002701	1.004231

Specification

Below in figure 14, the distribution of the dependent variable RTI is depicted. This skewed distribution is the reason why the binary model specification is applied.

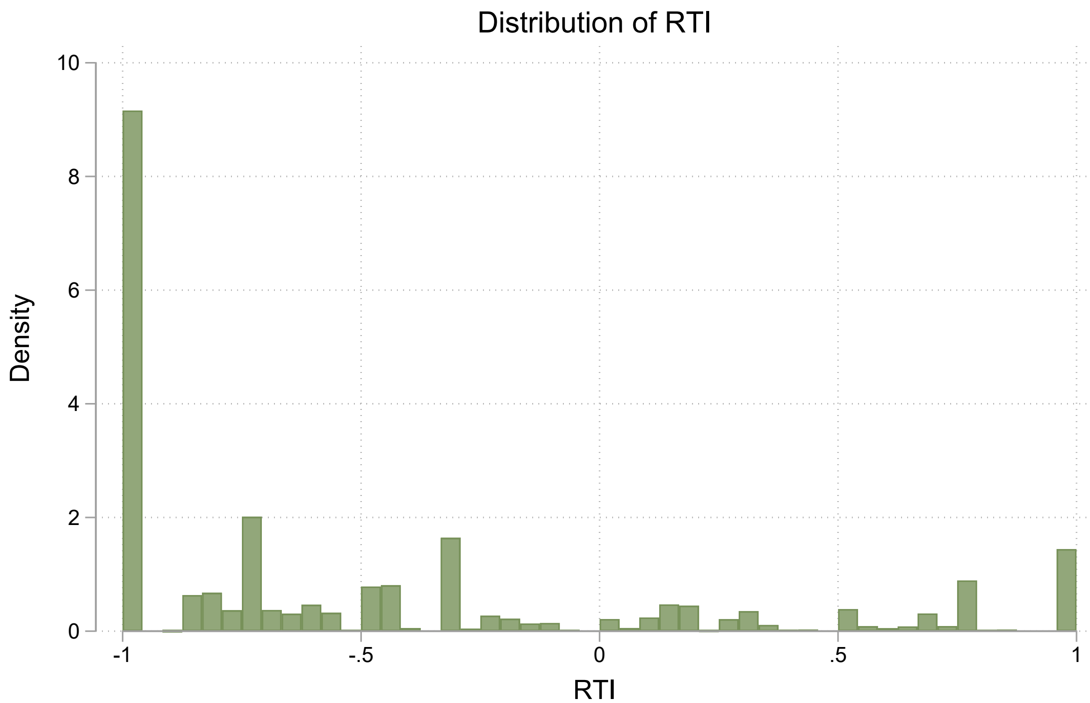


Figure 14: The frequency of RTI is is not equally distributed.

As a last point, I would like to add the following graph 15 of the quantile regression estimation.

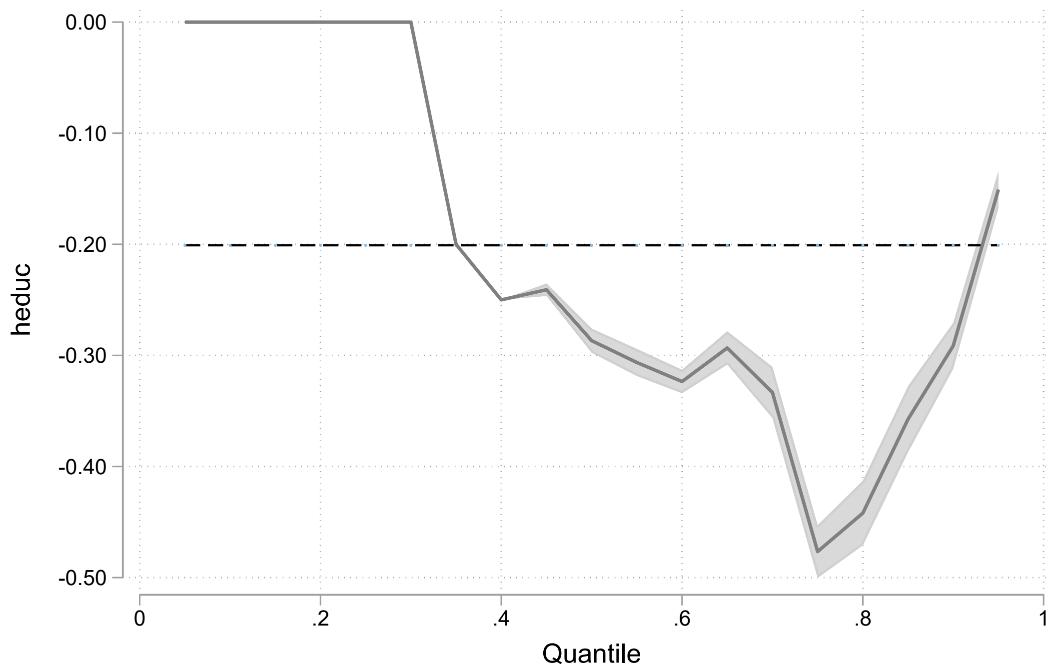


Figure 15: Effect of higher education on RTI across all percentiles

What can be seen here is the predicted value of the coefficient over different percentiles of the dependent variable. At first sight, it becomes evident that the effect of education is not equally distributed across the percentiles. Since my variable is skewed around -1, the quantile regression fails to estimate the effect at -1. But what becomes more evident is that the effect gets stronger for individuals with a higher RTI. This suggests that, all else equal, an individual in the 75^{th} percentile with higher education has a significantly lower RTI compared to an individual without higher education in the same percentile. It indicates that in occupations with higher routine shares, higher education has a greater impact.

Even though the quantile regression plot reveals that the highly skewed distribution of the RTI raises some issues for the quantile regression at the bottom quartile, the effect at the median can still be obtained. Observing the median in the applied context aims to provide an estimate that is less prone to outliers.