University of Tartu

Faculty of Economics and Business Administration

Using causal random forest to evaluate vocational training programmes on employment probability in Estonia

Master's thesis

Martin Piir

Supervisors: Andres Võrk, Taavi Unt

2024

Abstract

This thesis analyzes the impact of vocational training programmes on employment probability in Estonia using the Causal Random Forest method, which is a machine learning algorithm designed to effectively estimate treatment effects. In order to conduct the analysis, the thesis uses anonymized registry data from the Estonian Unemployment Insurance Fund from 2015 to 2023. A moderate positive effect of vocational training on employment probability is found, with an average increase of 9.7 percentage points. Looking at the effects for different subgroups, it is evident that the effects of vocational training are larger for those who belong in disability or language risk groups or have less education, but they are currently underrepresented in the treatment group.

Keywords: vocational training, employment probability, unemployment, causal random forest, Estonia

JEL Classification: J24, J64, J65

CERCS Classification: S220, S230

Table of Contents

Introduction

It is important to evaluate the effectiveness of labour market training programmes, since productive training programmes help people improve their skills and knowledge to either get out of unemployment or to find work that is more complex and thus also gives a higher salary. On the other hand, ineffective or suboptimally assigned programmes are a waste of money and also a waste of time for those attending them. In order to analyze the effectiveness of the vocational trainings for Estonian data, the Causal Random Forest method is used, which is a machine learning algorithm designed to estimate treatment effects among different subgroups in a dataset. (Tibshirani J. , Athey, Sverdrup, & Wager, 2019) The algorithm uses decision trees built on separate subsets of data, making it ideal for this sort of research, as it minimizes overfitting.

While it could at first glance be thought that any kind of vocational training is sure to show positive effects on the population, a paper summarizing 200 studies of active labour market programmes (Card, Kluve, & Weber, 2018) found that the effects are not so straightforwardly positive as one might expect. For example, public sector employment programmes were shown to have minor or even negative impact. In general, the programmes exhibit small to no effect in the short run, but the effects become more positive as time goes on. As for age, both the youngest and oldest age groups in the workforce were shown to have the least positive results after completing the training programmes. On the other hand, females and people who had been unemployed for long, had the most positive results.

Another meta-analysis (Vooren, Haelermans, Groot, & van den Brink, 2019) looking at 57 different studies concluded that the overall results of active labour market policies (ALMPs) are rather small, with the Cohen’s d value, which indicates the difference of two means divided by a standard deviation for the data, usually falling below 0.1. Contrary to the previous study (Card, Kluve, & Weber, 2018), there were no statistically significant effects found regarding the age or gender of the labour market program participants. Similarly to the previous study (Card, Kluve, & Weber, 2018), small to no impact was found in the short term here as well with the positive effects appearing later. Beyond that, they deduced that enhanced services schemes, which are more individual than average vocational training as they include more one on one meetings with caseworkers and provide assistance while looking for a job, are more effective than other methods in the short run. In the long run, the subsidized labour programs with features such as wage subsidies for employers and targeted employment were found to be most effective.

The results of the metaanalysis papers (Card, Kluve, & Weber, 2018), (Vooren, Haelermans, Groot, & van den Brink, 2019) suggest that it is important to conduct the study. This is because the effects on the population seem to vary quite a lot depending on the sample analyzed. This in turn implies that the risk of creating ineffective or suboptimal labour market programmes in practice without proper research is surprisingly high. Additionally, the vocational trainings could be assigned ineffectively, meaning that people with certain characteristics who don't get a significant increase in employment probability after attending the training are assigned into vocational training too frequently or vice versa. The variations in populations, job markets, and program quality across different countries and times make it challenging to draw conclusions for the Estonian job market from the meta-analysis papers alone. Thus, detailed analysis of regional data is essential to maximize the effect of vocational training.

As such, the thesis at hand uses anonymized registry data on unemployment instances provided by the Estonian Unemployment Insurance Fund from 2015 to 2023 in order to analyze the effects of vocational training and labour market training programmes on employment probability. The value added to the existing academic literature is using the Causal Random Forest on Estonian data. The research problem to solve with this thesis is to figure out how to better distribute opportunities for vocational training programmes in the Estonian job market in order to maximize their effectiveness and get the most amount of people back into the workforce.

The thesis is structured as follows: first a literature review of similar papers done globally, then focusing on reports published by the Estonian Unemployment Insurance Fund in addition to papers on the Estonian labour market. After that, there is a general overview about the nature of vocational training in Estonia, then a more in depth overview of the dataset on which the analysis is conducted on. Next, there is a description of the specific methodology used when conducting the analysis and running the model in R[1] alongside with an overview of the Causal Random Forest method itself. Then, the results of the Causal Random Forest model are presented and discussed. Finally, there is a comparison of the Causal Random Forest model with the more commonly used Propensity Score Matching method. Additionally, a list of each variable in the analyzed dataset is added to the Appendix section, alongside with tables and graphs of the results.

[1] Version 4.3.3

Literature review

The method of using causal random forests (Tibshirani J. , Athey, Sverdrup, & Wager, 2019) in order to analyze the labour market specifically has not been very popular so far in academic literature, as it has mostly been used in the medical field. However, there are a few papers similar to the study at hand, one of which was done in Belgium (Cockx, Lechner, & Joost, 2023) and found that the labour market programmes are especially effective for recent migrants. Differing from the meta-analysis mentioned in the introduction (Card, Kluve, & Weber, 2018), they did not find larger positive effects for women in their study. (Cockx, Lechner, & Joost, 2023) In our data, we do not have a specific value to determine whether someone is a recent migrant or not, but we do have an indicator which pinpoints if the person belongs in Estonian language risk group or not, thus it could be hypothesized that those belonging to that risk group could stand to gain more from vocational training. However, in Estonian context there are also Russian speakers belonging to that risk group who have been living here for their entire life, but still do not speak Estonian well. Another interesting finding from the Belgian study is that shorter vocational training programmes are more effective at getting people to (re-)enter the labour market. When entering any kind of a labour training programme, there is initially a negative effect on employment probability, as people will often be more focused on completing the training programme and not looking for a job. This effect is also called the "lock-in effect." However, for longer term programmes, this drop-off is much more drastic at first and even long-term, two and a half years after starting the programme, the people who enrolled in a short term training programme were more likely to be employed. (Cockx, Lechner, & Joost, 2023, pp. 27-28)

Another similar study was conducted in Germany using the same method. However, in this case, the training programmes were shown to be extremely effective, with positive employment effects showing up right after enrolling in the programmes, meaning that no lock-in effects are exhibited. Despite this, the effects are also long-lasting, meaning that the people are able to stay employed several years after enrolling in the programmes. One reason for this could be that the study was conducted by only looking at the people who were about to run out of unemployment benefits altogether, thus making them extremely motivated to excel in the training and find a job. Another reason could be that as previously seen in the Belgian study where the shorter training programmes had a better outcome at every point in time as compared to the longer training programmes, in the German system a vast majority of the programmes are extremely short. For so-called job training programmes, which are meant to guide recipients into apprenticeship and work, 92,9% of the programmes lasted less than three months and for reducing employment impediments training, the equivalent value was 94,5%. The third reason for the exemplary effectiveness of the German programmes could be that as a mandatory part of the job training, the participants are taught how to find work appropriate for their experience and skill level and how to write CV's and motivation letters in order to apply, thus providing them with not only the know-how on how to work, but also how to find work. (Goller, Lechner, Pongratz, & Wolff, 2023)

A more commonly used method in analyzing this type of problem is Propensity Score Matching (PSM), which is used to match individuals with similar attributes in order to isolate the effects of a specific treatment, in this case the labour market training programme. One study comparing PSM and the regular Random Forest method (Goller, Lechner, Moczall, & Wolff, 2020) found PSM to be more suitable for their dataset and also highlighted some of the potential drawbacks of using regular Random Forest, as in their study, the share of treated units was low and thus the regular Random Forest method could not effectively balance the covariate distribution between the treated group and the control group, thus leading to significant selection bias. Because of this, the regular Random Forest estimates were more biased than those obtained using PSM.

PSM has also been used to analyze the effectiveness of labour market progammes in Romania (Pirciog, Ciuca, & Popescu, 2015), where they were found to have a negative effect on employment instead, where the authors attributed such a result to the lock-in effects of the longer programmes and superficial knowledge gained in the shorter ones, but concluded that better targeting should be done in order to improve the efficiency of the programmes. However, the authors of the Romanian study have not considered the possible general poor quality of the programmes themselves, which in my opinion must certainly be looked at as a possibility after receiving such results. PSM was used in examining the Ugandan youth unemployment crisis, where different types of labour programmes were compared to each other. General vocational training provided by the government turned out to be more effective than offering firms subsidies in order to train the youth directly on the job, likely due to a broader skillset being gained that is applicable to many jobs in many firms, rather than learning to do one specific job in one firm. However, both programmes showed significant positive outcomes nevertheless. (Alfonsi, et al., 2020) In Poland, using PSM it was found that labour training programmes have a slight positive effect, but it is not as large as expected. The author believes that one of the reasons for this result could be that the people who receive the aid are in large part people who are trying to stack up more qualifications for themselves and improve their CV, who would find work without the vocational training programs as well, thus creating a significant amount of deadweight loss in the system. (Wiśniewski, 2022) As Propensity Score Matching is the most popular method used in academic literature for this type of analysis, we will be using it at the end as well to compare it to the results we got using the Causal Random Forest.

A lesser used method in analyzing the effects of labour market programmes is difference-in-differences, (Card & Krueger, 1994) which compares the changes in outcomes over a group that has enrolled in a labour market programme to a group that has not. A study using this method was conducted in Switzerland, researching how employers perceive different potential fictitious candidates who have enrolled in a labour market programme. It turns out that in case of people who had been unemployed for long or had poor general education and had clearly struggled with finding a job, attending vocational training was seen as a positive. However, in the case where the potential applicant was perceived as someone who should not have any trouble finding a job – good education, native, etc, having attended vocational training was instead often seen as a negative aspect, signaling a lack of thrive and motivation to succeed. (Liechti, Fossati, Bonoli, & Auer, 2017)

Vocational training for unemployed in Estonia

In Estonia, there are no recent academic articles analyzing the effectiveness of labour market training programmes specifically, however there are reports published by the Estonian Unemployment Insurance Fund on this matter, with the most recent one being published in January of 2024. It is shown in this report that there is a lock-in effect exhibited in Estonian data, as for the first three months after starting the programme, it is more likely that people who do not attend the programme find work more easily than those who do, as the people enrolled in the programmes tend to be busy with studying and are not as actively searching for working opportunities. However, after the third month, those enrolled in the programme become more likely to find work, as their programmes end or near their end. The effect reaches its peak 12 months after the start of the programme and stabilizes thereafter, with the effect being that those enrolled in the programmes are about 4-5 percentage points more likely to find work than those who did not. It is somewhat surprising to find out that the effect of the programmes for men is three times higher than for women, drastically differing from the extensive meta-analysis of labour market studies, which in general showed larger effects for women.

In Estonian data, this discrepancy is explained by the differences in the fields men and women prefer to complete their programmes in. Out of men, 57% of them enrolling in a labour market training programme do so in the field of transportation, thus obtaining necessary and easily marketable skills in the job market, which are in high demand. However, almost a quarter of the women have enrolled in programmes teaching marketing or beauty services, where the skills obtained are more ambiguous and not in very high demand, thus the effect of these programmes is nonexistent or even negative, which led the Estonian Unemployment Insurance Fund to decide to no longer finance programmes related to providing beauty services. Another possible reason for the discrepancy between genders, that the authors of the report mention, is the different education levels. A higher number of women enrolling in the programmes have a higher education degree compared to the men, thus the effect of additional education for the men could be more significant. Another peculiarity in the Estonian data highlighted in the report is the high effect of the programmes on older people aged 51-60, whereas in the meta-analysis, the effect on older people was typically relatively small. While the potential causes of this are not discussed in the report, it could be the effect of making the older people more acquainted with modern technology used in the workplace. Additionally, the effect of the programmes is larger for people with less previous education and smallest for those living in or around the capital city of Tallinn. The effect of the programmes on wages was found to be positive, but quite tiny, as the wages of those who had enrolled in the programmes were on average less than 1% higher than those who had not. (The Estonian Unemployment Insurance Fund, 2024)

The other recent report relevant to the thesis at hand was published in 2020 and it is on measures taken to prevent unemployment, in which the effects of labour market programmes were analyzed as well. The sample under review in this report has an especially high linguistic focus, as almost 80% of the participants had insufficient Estonian language skills and 68% of the programmes were focused on teaching Estonian. The results of the programmes in this sample were considered successful in this report, as 89,5% of the people who completed the programmes were employed at least 135 out of the next 180 days following the completion of the programme. (The Estonian Unemployment Insurance Fund, 2020) However, it must be noted that vocational training and language courses are analyzed together in the 2020 report, which is not done here, as the main focus in the current thesis is to analyze the impact of vocational training. As such, the results of that study are not very well comparable to our main results, but having received Estonian language training is still one of the variables in the analyzed dataset.

When talking about previous academic papers pertaining to the Estonian labour market, one such study was done by CentAR in 2012, analyzing the effects of both wage subsidies and labour market programmes, using data from 2008 to 2011. In this case it was found that men took part in vocational training more often than women, although during that time period more men than women were registered as unemployed as well. By age, most vocational training programmes were for the age group of 25-54. By education, the most represented segment in vocational training programmes were people with Level 2 education (*corresponding to Level 3 in our dataset*). Altogether, the impact of vocational training on wages was not statistically relevant, but the employment probability for those attending the training increased by 5.8 percentage points 6 months after starting the vocational training compared to similar people who did not attend any training. This effect also lasts over time, staying statistically relevant after a year and further on. Therefore it was concluded that vocational training has significant positive effects on those who receive it. (Anspal, et al., 2012)

The Estonian Unemployment Insurance Fund provides labour market training for those who are looking for a job and are in the unemployment registry. Additionally, they can register a person to and pay for their training if they have not been entered into the unemployment registry yet under the conditions if they have been given their layoff notice, they are an employee of a detained person or are detained themselves. Furthermore, it is possible to receive training while working if they are deemed to have a high risk of unemployment due to their health, field of employment or education. However, the labour market training is only provided under the condition if the skills to be obtained are relevant and in high demand in the current job market in the Unemployment Insurance Fund's assessment. Before being accepted into a labour market training programme, the potential applicants need to analyze their needs, strengths and weaknesses together with a consultant. If the consultant feels that assigning the person to a labour market programme is justified, then the future trainee will be either assigned to a training programme which has already been ordered by the Unemployment Insurance Fund or get given a training card, which can be used to pay for entry of up to 2500 euros in a wider variety of independently run training programmes, which have not been directly ordered or their contents thoroughly verified by the Unemployment Insurance Fund. (The Estonian Unemployment Insurance Fund, n.d.) Nevertheless, there is still a select list of training programmes to choose from, as the programme managers have to first submit an application to partner up with the Unemployment Insurance Fund in order for the trainees to be able to use their training card to pay for the programme. (RUP, 2023)

Data

The analysis was done using anonymized registry data provided by the Estonian Unemployment Insurance Fund. It is flow data, showing all new instances of unemployment, which are followed until exiting the unemployment. The provided dataset includes data from 1st of January 2015 up until 1st of September 2023, with unemployment periods starting before 2015 left out. Furthermore, people for whom the county is unknown, those who last worked in the armed forces occupations or those earning on average over 5000€ a month for the past 12 months are also left out of the data. The dataset is instance-based, meaning that a single person can have multiple instances of unemployment if they were unemployed, got employed and became unemployed again within the timeframe. In order to protect the privacy of those included in the data, meaning to avoid the chance of individual people getting recognized based on the date they were included in the unemployment registry, all unemployment instances have been shifted to the first day of the month. For example, if a person got added to the registry on the 30th of August, the registry data will show as them being added on the 1st of August instead, with the other date data in the registry such as beginning of labour market programmes or end of unemployment being shifted backwards by the according number of days as well. The data includes characteristics such as gender, age, county, education, type of last employment (first digit of ISCO code) and risk of handicap. Also, if two or more of the aforementioned important characteristics for a specific instance are missing, the instance has been already removed from the provided dataset. In addition to this, there are also indicators for the more popular different labour market services, which have been received in at least 1000 instances, such as computer training or language courses for example. A comprehensive list of all characteristics in the dataset alongside with their description is provided in Appendix 1.

Because of previously mentioned reasons, 21 272 instances have been removed, with the final dataset provided having 624 886 different instances. Before further analysis, the following instances were filtered out from the dataset in order to be more clearly able to assess the impact of the labour market training:

- labour training was received more than 180 days after being added into the registry
- labour training was previously received in the last 2 years

- labour training was stopped half way through

This brings the total number of instances down to 571 042, on which we will be doing our further analysis. Out of these instances, 50.4% are women and 49.6% are men. By age, the instances are divided into three groups: 30% are 16-29 years old, 43.2% are 30-49 years old and 26.8% are 50-64 years old. Based on the first digit of the ISCO code, the people who formerly worked in Elementary Occupations (*first digit of ISCO code 9*) make up 21.4% of the unemployment instances, followed by Services and Sales Workers (*5*) with 19.3% and Craft and Related Trades Workers (*7*) with 16.5%, with the rest being around 5-10% each, except for Skilled Agricultural, Forestry and Fishery Workers (*6*) making up only slightly less than 1%. As for education, 37.4% have level 3 education, 25.6% level 4, 21.5% level 1, 12.5% level 5 and 1.6% level 2, while around 1.5% of education levels are also marked as unknown. In the case of Estonian language skills, 25.4% of instances have been classified as falling into Estonian language risk group. Additionally, 17.7% of instances have been classified as having a disability risk. Finally, for two of the utmost important statistics, as they are our treatment and outcome – 7.8% of the instances received vocational training, while the rest did not and in 64.8% instances, the unemployment status ended by entering the workforce.

Table 1: Overview of the most important categorical variables

Variable	Number of observations	Share in data	Share in treatment group	Share in treatment	Share in employment (Y)
Total obs.	**571 042**		**44 754 (7.8%)**	**44 754 (7.8%)**	**369 906 (64.8%)**
Gender					
Men	283 104	49.6%	51.5%	8.14%	62.2%
Women	287 938	50.4%	48.5%	7.54%	67.3%
Age					
16-29	171 326	30.0%	26.1%	6.82%	65.1%
30-49	246 756	43.2%	51.9%	9.42%	67.7%
50-64	152 960	26.8%	22.0%	6.43%	59.6%
Education					
Level 1	122 847	21.5%	14.6%	5.31%	55.9%
Level 2	9 349	1.6%	1.8%	5.64%	59.4%
Level 3	213 403	37.4%	36.4%	7.63%	65.2%
Level 4	146 400	25.6%	29.4%	8.97%	69.2%
Level 5	71 685	12.5%	17.6%	10.99%	71.1%
Unknown	7 358	1.5%	0.9%	5.50%	59.3%
Language risk					
No risk	425 939	74.6%	81.8%	8.59%	66.5%
Risk present	145 103	25.4%	18.2%	5.62%	59.6%
Disability risk					
No risk	470 157	82.3%	88.6%	8.44%	67.1%
Risk present	100 885	17.7%	11.4%	5.04%	54.1%

Methodology

In order to solve the research problem, we have decided to use the Causal Random Forest method, which is a machine learning algorithm developed in order to better estimate heterogeneous treatment effects across different subpopulations within a dataset, making it an excellent method to use in order to analyze the effects of labour market programmes.

The algorithm works by first creating decision trees, which get built by splitting the data between the different subpopulations. At the end of the trees are leaves, where every leaf represents the different characteristics that an observation can have and the singular observations get assigned to the leaf that they belong, according to their characteristics. These trees identify the covariate splits, maximizing the squared difference in the treatment effects of different subpopulations, thus separating the data in a way which best brings out the differences in the treatment effects across subgroups. Once the preferred number of decision trees are created, the forest is formed. It then uses the structure of individual trees to estimate the treatment effects of new observations. (Credit & Lehnert, 2023) This method produces reliable estimates by using separate subsamples for different trees in the forest. Additionally, unlike the regular random forest, the causal random forest method allows dividing the subsamples into two sets, where one set is used to determine where the splits should occur in the tree, while the other set is used to provide the estimations at the end of the tree. Such a divide is used to avoid overfitting. Another way how the causal random forest differs from the regular random forest is that in the regular case, every tree makes a prediction on a test example, like whether the specific person is going to re-enter the workforce or not, and the predictions get combined at the end to produce the output. In our case, by using the grf package, using the test example, a list of neighbouring training examples is made instead, with the list being weighted according to the frequency of how often the neighbouring examples ended up in the same leaf of the tree as the test example. Another benefit of using this method and R package is the built-in robust average treatment estimator, providing more accurate estimators of the average treatment effect than simply averaging treatment effects across different examples, thus it should in the end provide us with accurate estimates of the effects of labour market programmes in Estonia. (Tibshirani J. , Athey, Sverdrup, & Wager, 2019)

In order to analyze the described data set, the causal.forest function in R from the package grf is used. (Tibshirani J. , Athey, Sverdup, & Wager, 2024) This function takes seven main arguments: X, Y, W, number of trees, sample.fraction, mtry and honesty. X is a matrix, which includes the explanatory characteristics for both the participation group, which received treatment and the comparison, which did not receive treatment. Y is the output vector. In our case, the output is the unemployment ending by either becoming employed (1) or not (0). W is the vector which separates the participation group from the comparison group, which in our case is the separation between those who started attending vocational training in the time period of 6 months after being added to the unemployment registry and those who did not receive any vocational training at all. The number of trees is simply the number of trees that will be created in the causal forest, the more trees, the more accurate will the predictions be. Sample.fraction specifies the fraction of the training data which is used to build each individual tree in the forest. In our study, we set the value to 0.33, thus 33% of the training data is used in order to build each tree. The purpose of the sample.fraction parameter is to reduce the dependence between the component trees, which reduces variance of the predictions, enhances the generalization ability of the model and increases computational efficiency. The argument mtry specifies the number of variables, which are considered for splitting at each tree node. In our study, the value was set to 24 (approximately 1/3 of total variables), meaning that the random forest algorithm would select 24 variables at random at each decision point while building the tree and evaluating only among those variables, leading to forests which are more robust. While the default setting for this argument is the square root of all variables (which in our case of 70 would be rounded down to 8), we decided to opt for a larger value. This approach captures more relevant information at each split, which reduces bias at the expense of a slight increase in variance. Finally, the binary honesty argument specifies how data is used during the construction of the forest. In general and in our study as well, this parameter is activated to ensure that the data used to decide where to make the splits within the tree is separate from the predictions at the end of the leaves. This distinction helps to reduce overfitting and reduce bias when estimating the treatment effects.

Results

The main result that we get by running the model specified above is the Average Treatment Effect on Treated (ATT), given by the following formula, where Y_1 refers to the outcome for a person if they received the treatment, Y_0 refers to the outcome for the same person if they did not receive the training. W being equal to 1 shows that the individual has received the treatment.

$$ATT = E[Y_1 - Y_0 \mid W = 1]$$

ATT shows that vocational training leads to an average increase of 9.7 percentage points in employment probability, showing that vocational training has a moderate positive effect on future employment. The Average Treatment Effect (ATE) is given by the following formula:

$$ATE = E[Y_1 - Y_0]$$

ATE shows an impact of 9.9 percentage points. For estimating the treatment effects for the training data using out-of-bag prediction, the average effect is 10.9%.

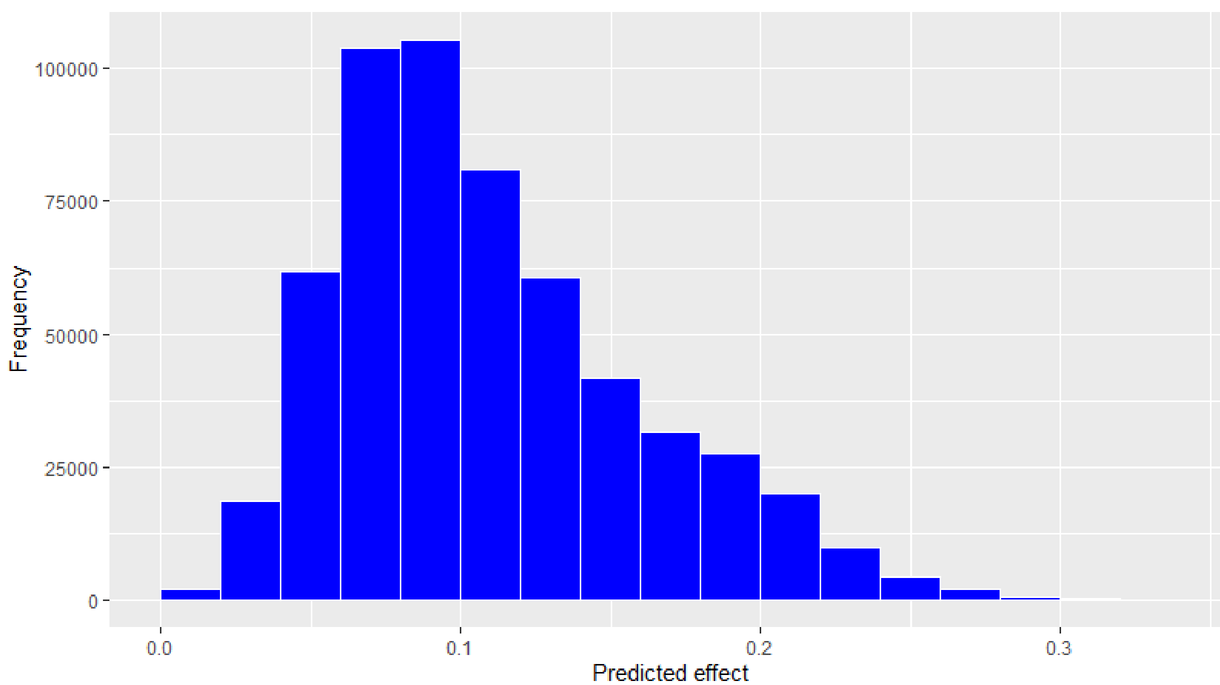

Figure 1: Histogram showing out-of-bag prediction effects

When it comes to conditional average treatment effects, we are using best linear projection, meaning the relationship between treatment effects and covariates. We will first look at how different age groups respond to the treatment. In our data, the age groups are divided into three: 16-29, 30-49 and 50-64. For the 16-29 age group, the estimate is 0.091, showing an increase of 9.1 percentage points in unemployment levels with a standard error of 0.004 and p-value under 0.001 (*Appendix 4*), suggesting that this estimate is highly statistically significant. For the 30-49 age group the estimate is -0.003 with a p-value of 0.51, showing that there is no difference in the treatment effects compared to the 16-29 age group. For the smallest 50-64 age group, the estimate is the highest by far – 0.05 with a p-value under 0.001 for a total of 14 percentage points, meaning that the estimate is highly statistically significant and strongly positive compared to the reference group of 16-29. Overall, this suggests that vocational training has the highest positive effects on the oldest age group, with the effect for the two younger age groups being similar. These effects are surprisingly large, but as previously discussed, could be explained by the older generation gaining a familiarity with technology through vocational training that they have not obtained through everyday life like the younger generations.

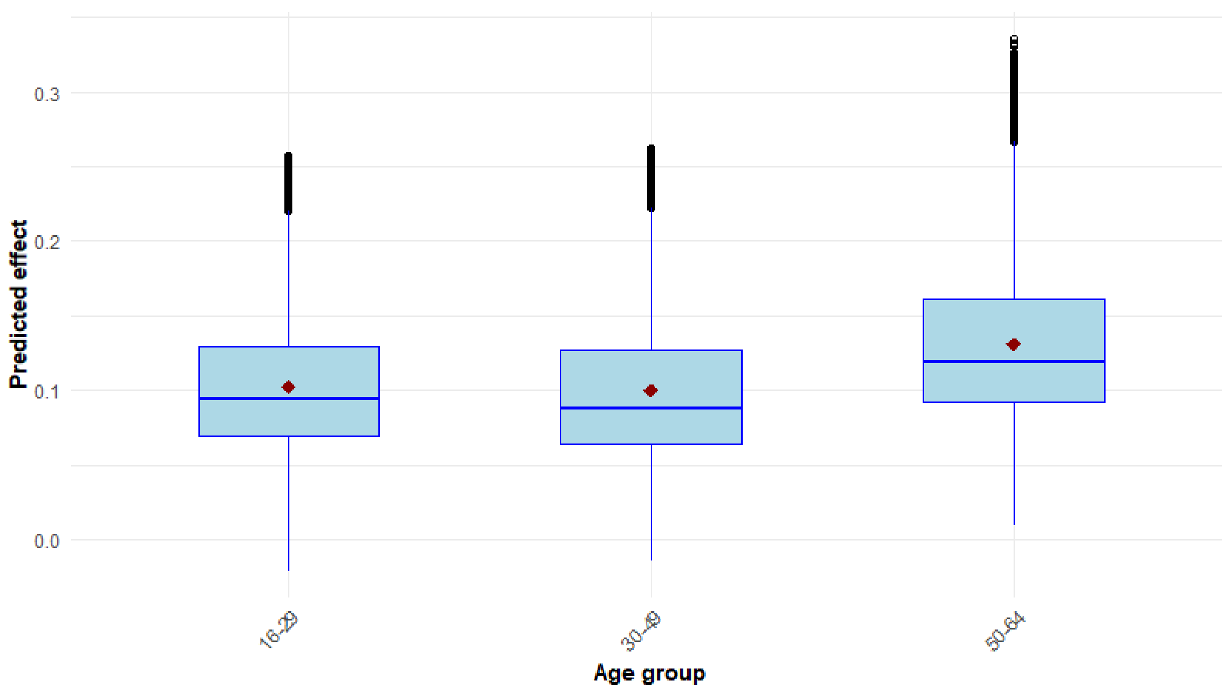

Figure 2: Boxplot showing the predicted effect of vocational training by age groups (diamond showing the mean)

For ISCO codes or previous unemployment the effects are not that large or interesting overall, although it could be noted that the Professional group (*first digit ISCO code 2*) benefits the least from vocational training with their employment probability showing an increase of just 6.72 percentage points, which could be because they already are highly skilled workers, thus the vocational training increases their skills less than for other groups. Similarly, those with a missing ISCO code (who have not been employed before), benefit the most from vocational training, increasing their employment probability by almost 17 percentage points.

When it comes to the amount of times being registered as unemployed before the current instance in the previous 3 years, there is surprisingly not much of an effect. The baseline of 0 has a positive estimate of 9.85 percentage points, being relatively close to the overall average of 9.7. While being registered as unemployed only 1-3 times in the previous 3 years has no effect at all compared to the people registered as umemployed 0 times, being registered 4 times shows a slight positive effect at +1.95 percentage points, but with a p-value of 0.21, thus not being statistically significant. Surprisingly, being registered as unemployed 5 or over 6 times doesn't seem to have a statistically significant negative effect either. This is somewhat unexpected, as falling into unemployment on average twice every year or more often seems to hint at some serious personal problems with health, attitude or motivation, which I would have expected to have been difficult to fix through vocational training.

In the case of the reason for the end of last employment, as already seen in the ISCO code section, the people without any previous employment at all receive a nearly 17 percentage point increase in employment probability from vocational training. It is quite surprising that the next highest increase in employment probability is for the group who had their employment contract terminated for reasons arising from the employee at a 11.5 percentage point increase. It is unexpected, since some of the reasons for being terminated in this faction include inability to perform their duties due to a health condition, but also ignoring the employer's warnings, drinking on the job, stealing or damaging property, but also inability to complete the tasks required. However, if they are unable to complete the tasks required, then generally the employers themselves must offer other work that the employee is able to perform or provide retraining or adapt the workplace. (The Labour Inspectorate of Estonia, 2024) Thus, it seems that firing somebody for their inability to complete the required tasks is relatively difficult. For most of these reasons, it would seem that vocational training would have little effect, as it would not help in the case of health and behavioural issues, but the data peculiarly indicates otherwise. The third group that has a higher than average increase in employment probability are those who mutually with their employer decided to end their working relationship at a 10.7 percentage point increase. This is a more expected result, as one would think that those ending working relationships with their employer cordially would also be more likely to find a new job in the future. Other reasons for last termination of employment of contractual reasons, reasons arising from the employer or termination at the initiative of the employee all have about a 8.5 percentage point increase, so slightly lower than the overall average of 9.7.

When it comes to gender, men are far more likely to receive an increase in employment probability than women with 12.4 percentage point increase for men and just a 7.77 percentage point increase for women. This was also an expected result after reading the most recent report on the matter, (The Estonian Unemployment Insurance Fund, 2024) as that report highlighted this gap between the genders as well, with the reason being that a substantial number of women in the past attended vocational training in fields such as beauty services or marketing, which did not improve their employment prospects significantly. On the other hand, the men often became certified drivers or forklift operators and quickly found a job, as there is a shortage of skilled workers in those fields. However, as the Unemployment Insurance Fund is already aware of this issue and funding for such vocational trainings that are unlikely to have an effect on employment prospects has stopped, this gap between the genders is expected to close at least somewhat in the upcoming years as well.

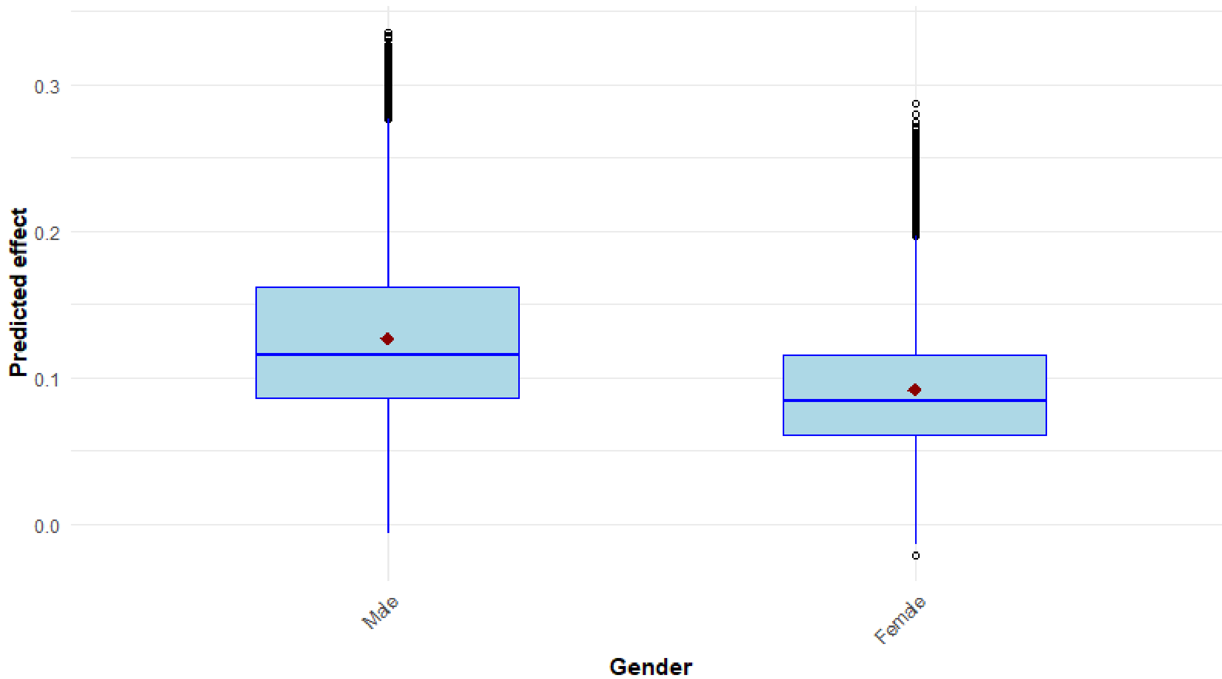

Figure 3: Boxplot showing the predicted effect of vocational training by gender

For the linguistical aspect, there is a positive correlation between falling into the Estonian language risk group and the effect on employment that vocational training has, as the employment probability for those in the risk group increases 2.8 percentage points more compared to those not in the risk group. Although when we look at the descriptive statistics, those with no language risk are much more prevalent in the treatment group. Thus, the first suggestion would be to allocate more vocational training to those with language risk in the future.

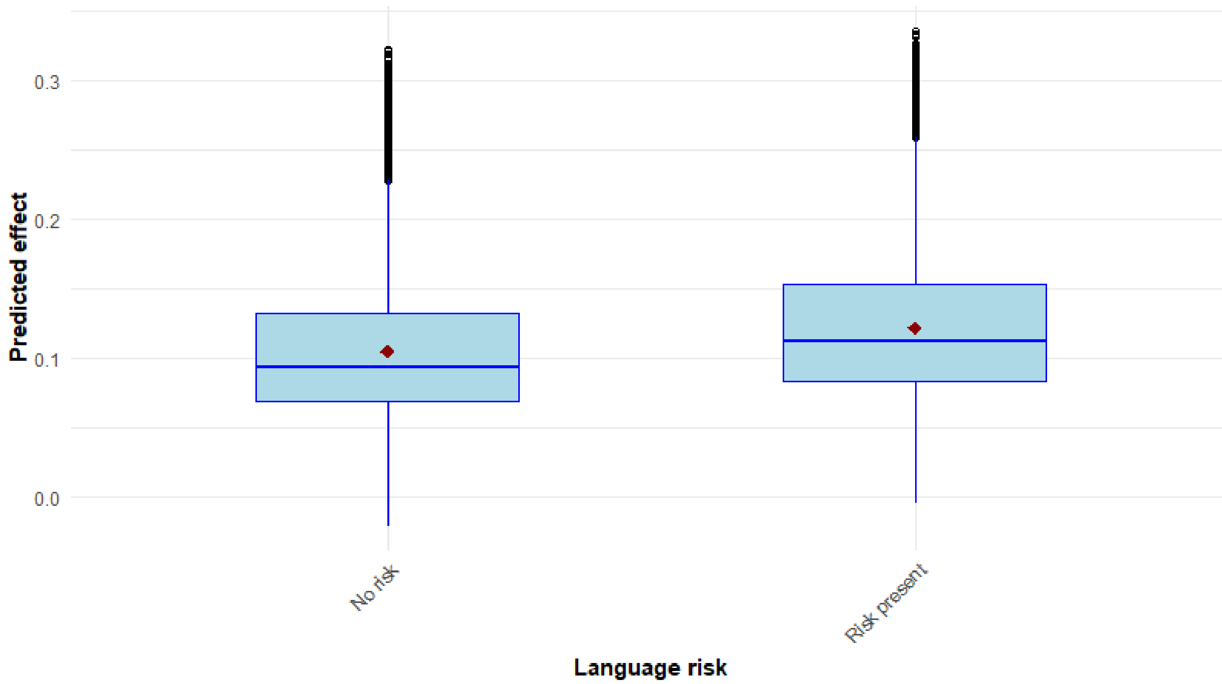

Figure 4: Boxplot showing the predicted effect of vocational training by language skill

The linguistic aspect is also reflected in the regional effects, as the region where vocational training has the biggest effect of 4.9 percentage points over the baseline of Harju county is Ida-Viru county, where there is a large Russian-speaking population who often have poor Estonian language skills despite living in Estonia for a long time. In all other regions of Estonia, vocational training has an effect of around 3 percent more than the baseline of Harju county, where almost half of the total population lives. The reason for this is probably that in Harju county, a large population is condensed into a relatively small area, which means a lot of employment opportunities, which makes finding a job relatively easier compared to other regions even without vocational training. In other areas however, employment prospects are often much more limited and thus vocational training to fill the specific available vacancies becomes more important. When it comes to the comparison of urban and rural areas, expectedly people living in rural areas respond to vocational training better for the same limited employment options reason specified above, but the effect is not as large as I would have expected, with vocational training in rural areas having only a 1.6 percentage points larger effect than in urban areas.

Turning our attention to education, there is a pretty clear pattern of vocational training having a much greater impact on those with lower education and a rather small effect on those with higher education, which is a pattern that we already got a hint of in the previous employment section, where the Professionals were the least affected by vocational training. The baseline of the Level 1 group which has an unspecified education level or missing basic education has an average increase of 12.4 percentage points in employment probability when attending vocational training with Level 2 having the exact same effects. When it comes to Level 3 or Level 4 groups, essentially those who have completed vocational education or high school, their effects are slightly higher than the overall average at around 10 percentage point increase. Finally, for Level 5 group, the people with an university degree, the effects are by far the lowest with their effects coming out to full 5 percentage points below the first group at 7.4 percentage points. Such results are quite logical, as the less education a person has, the more relative gain they are going to have from each bit of additional education. The same types of trends are present when analyzing computer skills as well, as when those have been classified as "Beginner" or "Nonexistent", there are significant increases in employment probability, with 12.9 percentage points for beginners and a whopping 22 percentage point increase for those with no experience with computers at all, 79% of whom also belong in the 50-64 age group, which had a significantly higher increase in employment probability than younger age groups and this generational gap in computer skills explains why vocational training is so effective for the oldest age group. Other levels of computer skills seem to not have much effect and stay around the overall average treatment effect.

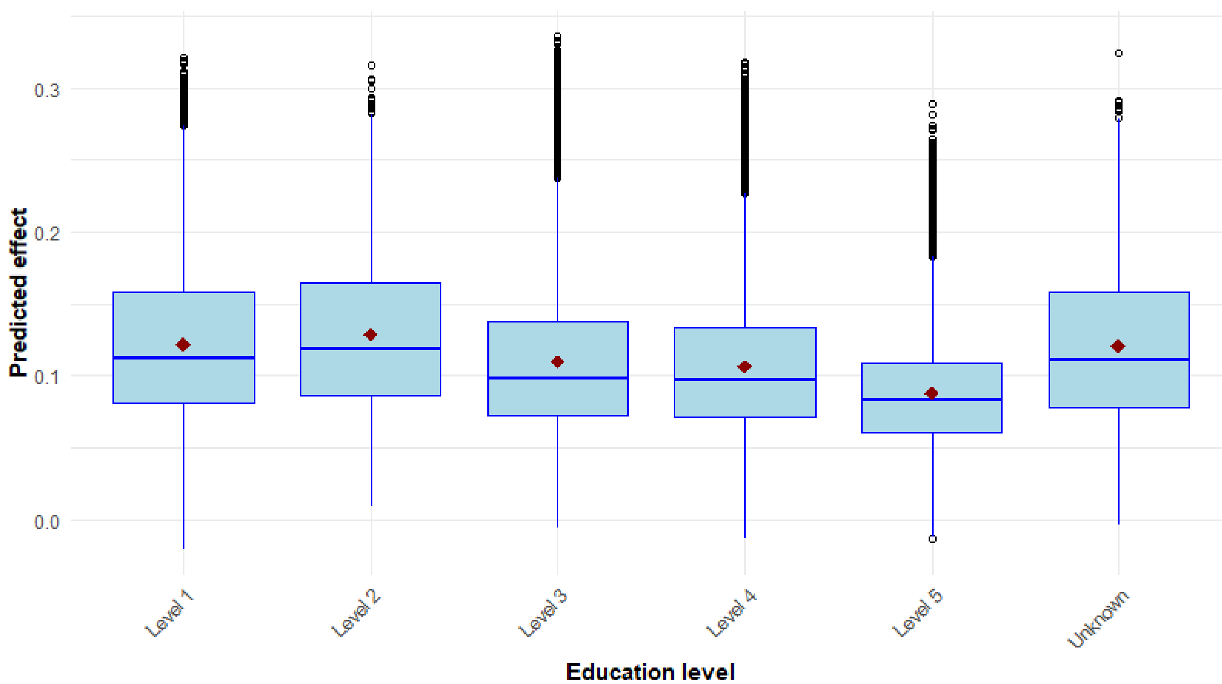

Figure 5: Boxplot showing the predicted effect of vocational training by education level

However, when looking at descriptive statistics for education, we can see that for share in treatment, the more education the person has, the more likely it is that they have received treatment, with only 5.31% of Level 1 educated people receiving vocational training and almost 11% of Level 5 educated people receiving the training. As discussed in the previous paragraph, the effectiveness of the training goes the exact opposite way, with the least educated people benefitting the most and most educated people benefitting the least. Thus, in the near future, the current type of vocational training should be more often assigned to those with a lower education level in order to maximize its effectiveness. Additionally, when looking at the more distant future, vocational training should be improved in order for the programs to be more effective for those with an university education.

Another statistic worth mentioning is the high impact vocational training has on disabled people, as attending vocational training makes their employment probability almost 5 percentage points higher compared to people without disability. However, they are much less likely to receive vocational training than those without disabilities, with their share in the treatment group being

5.04% compared to 8.44% of those with no disability risk. Therefore, vocational training should be assigned to those with disabilities more often in the future.

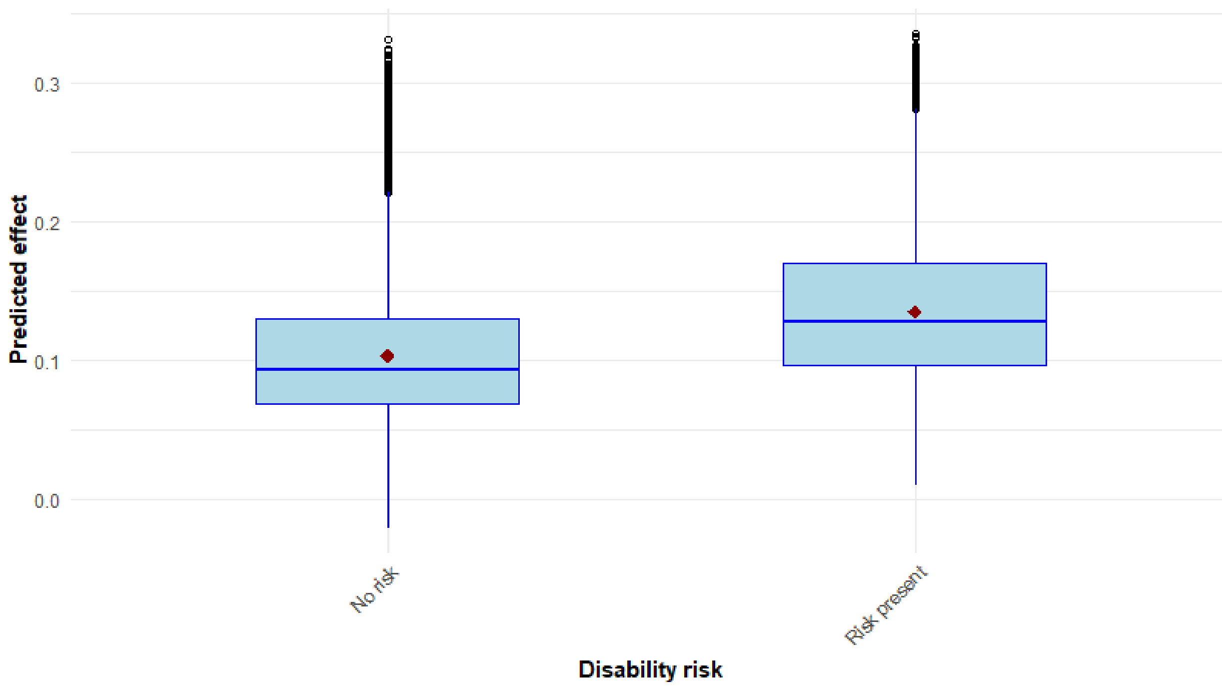

Figure 6: Boxplot showing the predicted effect of vocational training for people with disability risk

For the part of the data which shows whether the person has participated in a previous service or training in the period during the two years preceeding the period of unemployment, most of these seem to have had no statistically relevant effects.

The services received in the past two years that have had a positive statistically relevant effect are firstly job information session, which has a statistically relevant positive effect of 5.7 percentage points with a p-value of 0.002. Thus it could be that this is a valuable service for future jobseekers, but as the service is short and was received a while ago, it could simply show that the people seeking out this service are also more active in looking for potential job opportunities.

The second service with a statistically relevant effect is the subsidy for employment of minors. The effect is +7.9 percentage points at a p-value of 0.02, compared to the reference group of those not having received this service. This is not actually a service for the worker, but rather a subsidy paid to the employer for hiring a youth aged 13-16. (Estonian Unemployment Insurance Fund, 2024) However, the majority of these youths in the database have the lowest level of education (Tase 1), thus extra vocational training goes a long way in helping their job prospects.

Peculiarly, having attended career counseling in the past two years has a statistically relevant negative effect of 1.4 percentage points on employment probability with a p-value of 0.034. However, this does not necessarily mean that this service has a detrimental effect on those who receive it. Since it is on the higher end of statistically relevant p-values, it could be possible that we are dealing with a Type 1 error. Secondly, it could also be possible that the group receiving this service has a much harder time navigating the job market in the first place and as such, the negative effects on employment probability for this group would be much larger without the career counseling service.

Table 2: Services received in the past two years which have a statistically relevant effect

Variable	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	0.1005	0.0020	51.2668	0.0000
Job information session in the past two years	0.0576	0.0184	3.1252	0.0018
(Intercept)	0.1009	0.0020	51.6501	0.0000
Subsidy of employment of minors in the past two years	0.0786	0.0339	2.3204	0.0203
(Intercept)	0.1026	0.0020	50.2983	0.0000
Career counseling in the past two years	-0.0144	0.0068	-2.1202	0.0340

For the trainings received during the current unemployment instance, having received computer training has a statistically relevant positive effect of 2.7 percentage points with a p-value of 0.018, making it statistically significant. However, having received Estonian language training instead has a large negative effect of 4.1 percentage points on employment probability with a p-value of less than 0.001, which is quite surprising and it is not clear why the effects would be so negative in this case, especially as vocational training in general seems to have better results for people who do not speak Estonian well. There are extremely large negative effects for having received the long-term sheltered employment service of negative 20.3 percentage points at a statistically significant p-value of under 0.001, but this makes sense as the people who have received this type of service are incapacited for work to the extent of 80-100% or have severe health issues like brain trauma or significant mental disorders. (Republic of Estonia Social Insurance Board, 2024) Therefore, it is understandable that people who have received this service stay in the program for long as the program name hints at as well and are not able to move on to regular employment often. There are also negative effects for having received career counceling of -1.2 percentage points at a p-value of 0.005 and for having attended an information session (different from job information session) of -1.9 percentage points at a p-value of 0.02. More services showing negative effects at a p-value of less than 0.001 are the the trial work service at -6.8 percentage points, apprenticeship service at -4.5 percentage points and general skills service of -4.6 percentage points. However, in these cases there is almost certainly significant sampling bias, as the people receiving these additional services are likely to have low value in the labour market in the first place and it could be the case that without these services, their employment probability would be even lower, which makes the true impact of these services difficult to properly assess.

Table 3: Statistically important trainings or services received within the current unemployment instance

Variable	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.1003	0.0020	50.6593	0.0000
Computer skills training	0.0274	0.0116	2.3698	0.0178
(Intercept)	0.1025	0.0020	51.7497	0.0000
Estonian language training	-0.0417	0.0112	-3.7218	0.0002
(Intercept)	0.1013	0.0020	51.8930	0.0000
Long-term sheltered employment service	-0.2027	0.0340	-5.9691	0.0000
(Intercept)	0.1052	0.0023	44.8923	0.0000
Career counseling	-0.0119	0.0042	-2.8195	0.0048
(Intercept)	0.1025	0.0020	50.9690	0.0000
Information session	-0.0188	0.0081	-2.3366	0.0195
(Intercept)	0.1037	0.0020	51.8936	0.0000
Trial work service	-0.0685	0.0089	-7.6638	0.0000
(Intercept)	0.1040	0.0020	52.0031	0.0000
General skills service	-0.0465	0.0089	-5.2435	0.0000
(Intercept)	0.1032	0.0020	51.9705	0.0000
Apprenticeship service	-0.0445	0.0101	-4.3939	0.0000

Propensity Score Matching

In order to check the results that were gotten using the Causal Random Forest method, Propensity Score Matching was also used. In this case, for defining the formula, 13 of the 70 variables were used. The methodology for picking the variables to use involved first picking some of the most basic sociodemographic variables and then using the variable importance table to add other variables which showed the most effect. More specifically, the variables used in the formula are age, gender, education, Estonian language risk, disability risk, first digit of last occupation's ISCO code, reason for termination of last employment, average wage for the last 12 months of employment, duration of last employment, assigned unemployment insurance benefit in days, the first daily rate of the determined unemployment insurance benefit divided by the average salary on which the unemployment insurance benefit is based in the previous year, whether the person received practical training during current unemployment instance and whether the person received general skill training during current unemployment instance. For this purpose, the R package MatchIt was used. (Ho, Imai, King, & Stuart, 2011) For the parameters of the model, nearest neighbor matching, logistic regression model for the propensity score and logit link function were used. Each treated subject was matched with 3 controls, matching was done with replacement, average treatment effect on the treated was targeted. A caliper of 0.01 was chosen, caliper was not in standard deviations and controls that could not be matched with the caliper were discarded. The control group had 526 288 observations, of which 91219 were matched, 109053 were matched but unweighted, 417023 were unmatched and 212 discarded. Out of the treated group of 44 754, only 2 observations were unmatched, with the rest being matched.

For this model, the average treatment effect on treated (ATT) is an 11.3 percentage point increase in employment probability.

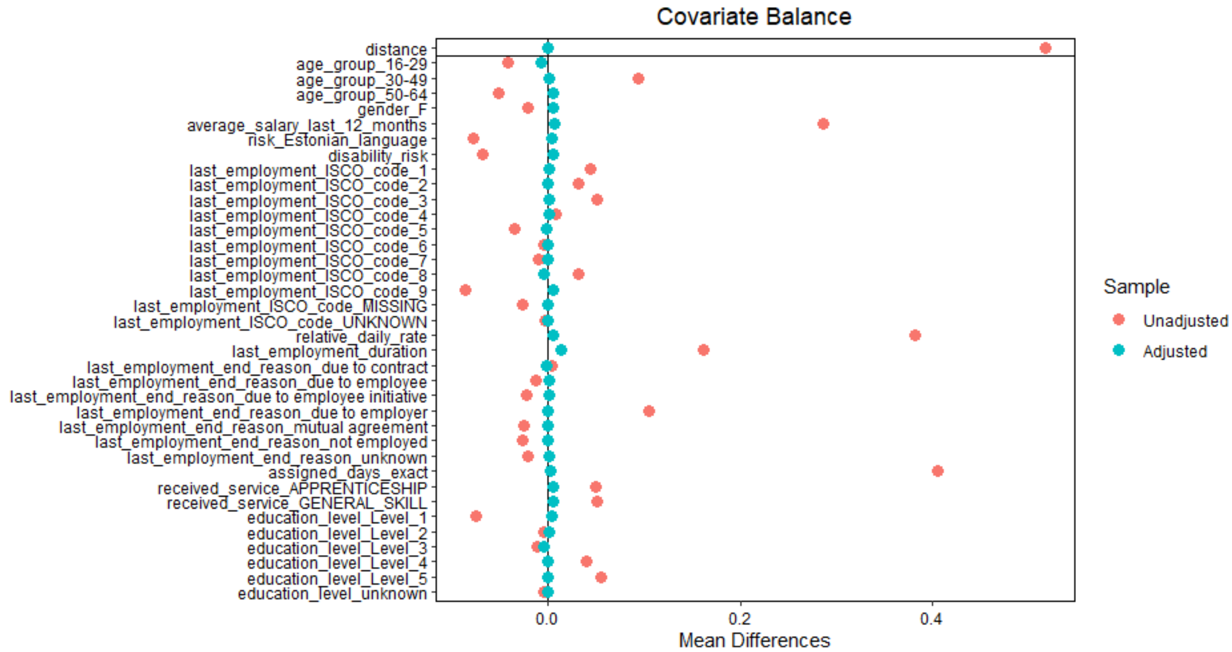

Figure 7: Love plot showing both unadjusted and adjusted covariates

The graph indicates that the matching has greatly reduced the differences between the two groups.

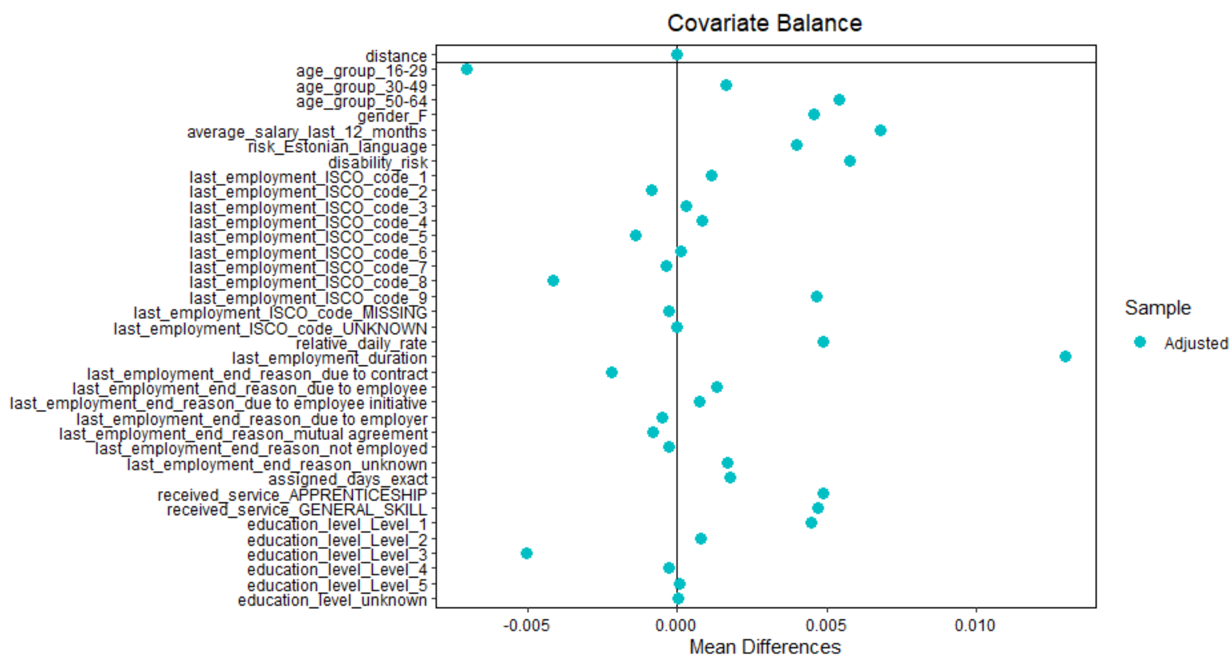

Figure 8: Love plot showing adjusted covariates only

The graphs show that the covariate balance post-matching is good and the imbalances are quite small.

Conclusion

The thesis used the Causal Random Forest method to evaluate the effect of vocational training programs on employment probability using registry data ranging from 2015 to 2023. Overall, the results show that on average vocational training has a moderate positive effect of improving the employment probability around 10-11 percentage points. The effectiveness of the vocational training programmes varies by age and gender, with the oldest age group of 50-64 benefitting the most and men benefitting more than women. The most obvious ways to improve the allocation of the programmes in the future is to firstly give vocational training to people with language and disability risk more often, as they are underrepresented in the treatment group, but the training has a more significant effect on their employment probabilities compared to those outside of the language and disability risk groups. Secondly, the higher the person's education level is, the more likely they are to receive vocational training as well. However, the effectiveness of the vocational training goes the opposite way, with the training having much more effect on the employment probabilities of less educated people and a relatively small effect on highly educated people. Thus, in the future those with less education, language risk or disability risk should be more highly represented in the treatment group to maximize the effectiveness of the vocational training programmes.

Bibliography

Alfonsi, L., Bandiera, O., Bassi, V., Burgess, R., Rasul, I., Sulaiman, M., & Vitali, A. (2020, November). Tackling Youth Unemployment: Evidence from a Labor Market Experiment in Uganda. *ECONOMETRICA, 6*(88), pp. 2369-2414. Retrieved from https://sticerd.lse.ac.uk/dps/eopp/eopp64.pdf

Anspal, S., Järve, J., Jääts, K., Kallaste, E., Maar, K., Mägi, A., & Toots, A. (2012). *Programmi "Kvalifitseeritud tööjõu pakkumise suurendamine 2007-2013" teenuste hindamine.* Eesti Rakendusuuringute Keskus CentAR.

Card, D., & Krueger, A. B. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review, 84*(4), 772-793.

Card, D., Kluve, J., & Weber, A. (2018, June). What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations. *Journal of the European Economic Association, 16*(3), pp. 894-931.

Cockx, B., Lechner, M., & Joost, B. (2023, January). Priority to unemployed immigrants? A causal machine learning evaluation of training in Belgium. *Labour Economics*(80). Retrieved from https://arxiv.org/ftp/arxiv/papers/1912/1912.12864.pdf

Credit, K., & Lehnert, M. (2023, June 14). A structured comparison of causal machine learning methods to assess heterogeneous treatment effects in spatial data. *Journal of Geographical Systems*. Retrieved from https://link.springer.com/article/10.1007/s10109-023-00413-0

Estonian Unemployment Insurance Fund. (2024). *Services and Benefits*. Retrieved from Subsidies for employment of minors: https://www.tootukassa.ee/en/services/employers/subsidies-employment-minors

Goller, D., Lechner, M., Moczall, A., & Wolff, J. (2020, August). Does the estimation of the propensity score by machine learning improve matching estimation? The case of Germany's programmes for long term unemployed. *Labour Economics*(65).

Goller, D., Lechner, M., Pongratz, T., & Wolff, J. (2023, May). Active labor market policies for the long-term unemployed: New evidence from causal machine learning. Retrieved from https://arxiv.org/ftp/arxiv/papers/2106/2106.10141.pdf

Ho, D., Imai, K., King, G., & Stuart, E. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software, 42*(8), 1-28. doi:doi:10.18637/jss.v042.i08

Liechti, F., Fossati, F., Bonoli, G., & Auer, D. (2017, April). The Signalling Value of Labour Market Programmes. *European Sociological Review, 33*(2), pp. 257-274. Retrieved from https://www.jstor.org/stable/44507708

Pirciog, S., Ciuca, V., & Popescu, M. E. (2015). The net impact of training measures from active labour market programs in Romania – subjective and objective evaluation. *Procedia Economics and Finance*(26), pp. 339-344. Retrieved from https://www.sciencedirect.com/science/article/pii/S221256711500862X

Republic of Estonia Social Insurance Board. (2024). *Republic of Estonia Social Insurance Board*. Retrieved March 28, 2024, from Long-term sheltered employment: https://sotsiaalkindlustusamet.ee/en/disability-and-welfare-services/special-welfare/long-term-sheltered-employment

RUP. (2023, November 21). Retrieved from https://www.rup.ee/uudised/tooturg/tootukassa-avab-vastuv-tu-koolituskaardi-koostoopartneriks-saada-soovivatele-koolitajatele

The Estonian Unemployment Insurance Fund. (2020, March). Töötust ennetavad meetmed; Ülevaade perioodist mai 2017-aprill 2019. Retrieved from https://www.tootukassa.ee/web/sites/default/files/2022-01/tootust_ennetavate_meetmete_analuus_loplik.pdf

The Estonian Unemployment Insurance Fund. (2024, January). Ametikoolituste mõjuhindamise tulemused; Lühiülevaade ametikoolituste töötamise tõenäosusele ja keskmisele palgale. Retrieved from https://www.tootukassa.ee/web/sites/default/files/2024-01/Ametikoolituste%20m%C3%B5juhindamise%20tulemused.pdf

The Estonian Unemployment Insurance Fund. (n.d.). *Career and training: Trainings*. Retrieved March 6, 2024, from https://www.tootukassa.ee/et/teenused/karjaar-ja-koolitamine/koolitused

The Labour Inspectorate of Estonia. (15. February 2024. a.). *Working Life Portal*. Allikas: Unilateral termination of employment contract: https://www.tooelu.ee/et/31/toolepingu-uhepoolne-lopetamine#toolepingu-ulesutlemine-tootajast-tuleneval-pohjusel

Tibshirani, J., Athey, S., Sverdrup, E., & Wager, S. (2019). The GRF algorithm. Retrieved from https://grf-labs.github.io/grf/REFERENCE.html#causal-forests-1

Tibshirani, J., Athey, S., Sverdup, E., & Wager, S. (2024). *_grf: Generalized Random Forests_.* Allikas: https://CRAN.R-project.org/package=grf

Vooren, M., Haelermans, C., Groot, W., & van den Brink, H. M. (2019, February). The Effectiveness of Active Labor Market Policies: A Meta-analysis. *Journal of Economic Surveys*, 125-149. Retrieved from https://onlinelibrary.wiley.com/doi/full/10.1111/joes.12269

Wiśniewski, Z. (2022). The effectiveness of Polish active labour market policies. *Bulletin of Geography. Socio-economic Series, 56*(56), pp. 125-132. doi:http://doi.org/10.12775/bgss-2022-0017

Appendixes

Appendix 1: Comprehensive list and description of all characteristics

- **instance_id** – dataset ID number of the period (instance) of being unemployed
- **person_id** – dataset ID number of the person
- **date_of_unemployment_registry_start** – the date of being added to the unemployment registry. Moved to the first day of the month. All other date type features in the dataset are also shifted by the same number of days.
- **date_of_unemployment_end** – the date of the end of being unemployed. If the person has not found a job by the time the data was finalized, then the value is NA.
- **centralized_date_of_unemployment_end** – centralized date of the end of unemployment
- **days_in_registry** – duration of being in the registry in calendar days
- **reason_for_instance_end_employment** – instance in the registry has been ended due to either moving into employment (1) or for another reason (0). In the original data set provided, the instances that continue to persist are marked as NA, in the data set modified for analysis, the NA values have been changed to 0.
- **age_group** – age of the person 34 days after being registered as unemployed, divided into three groups of 16-29, 30-49 and 50-64.
- **last_employment_ISCO_code** – the first number of the ISCO code of the last occupation (value 0 – military personnel were already excluded in the provided dataset, since there were very few of them)
    - 1 - Managers
    - 2 - Professionals
    - 3 - Technicians and Associate Professionals
    - 4 - Clerical Support Workers
    - 5 - Services and Sales Workers
    - 6 - Skilled Agricultural, Forestry and Fishery Workers
    - 7 - Craft and Related Trades Workers
    - 8 - Plant and Machine Operators and Assemblers
    - 9 - Elementary Occupations
    - MISSING - no previous employment
    - UNKNOWN - previous employment unknown
- **last_employment_duration** – duration of last employment in days
- **unemployed_3Yr** – the number of instances registered as unemployed in the three years prior to the current unemployment instance
- **unemployed_days_3Yr** – the number of days registered as unemployed in the three years prior to the current unemployment instance
- **assigned_days_exact** – assigned unemployment insurance benefit in days
- **relative_daily_rate** – the first daily rate of the determined unemployment insurance benefit divided by the average salary on which the unemployment insurance benefit is based in the previous year
- **assigned_benefits_duration** – duration of the specified unemployment benefit
- **last_employment_end_reason** – the reason for the end of last employment

- **average_salary_last_12_months** – the average salary of the 12 months prior to the entry into the account based on TSD (found as the total salary during the previous 12 months divided by 12).
- **gender –** binary variable, either male (M) or female (F)
- **Estonian_language_risk** – indicator which shows whether a person belongs to the Estonian language risk group (1) or not (0).
- **disability risk** – indicator, which shows whether a person belongs to a disability risk group (1) or not (0).
- **village** – whether the person is a resident of a village (1) or not (0).
- **location** – codes for Estonian regions:
    - EE001: Harju county
    - EE004: Hiiu, Lääne, Pärnu and Saare counties
    - EE006: Järva, Lääne-Viru and Rapla counties
    - EE007: Ida-Viru county
    - EE008: Jõgeva, Põlva, Tartu, Valga, Viljandi and Võru counties
- **education_level** – education level, aggregated
    - Level_1: unspecified; either primary or basic education missing
    - Level_2: vocational education without basic education, vocational education without secondary education, basic education with vocational education
    - Level_3: general secondary education, vocational education, vocational secondary education based on basic education
    - Level_4: vocational secondary education based on high school, secondary special education, vocational higher education
    - Level_5: Bachelor's studies, Master's studies, Doctoral studies
- **computer_skills** – level of computer literacy. Logging was started in September of 2017.
- **vocational_training_start** – date of the start of the vocational training. If during the instance the person participated in several vocational trainings, then the first one is selected. If the person did not participate in any vocational training during the condition, the value is NA.
- **vocational_training_result –** result of the vocational training. Either “Lõpetas” (finished the training) or “Katkestas” (stopped attending the training).
- **training_card** – indicator, which shows whether the vocational training has a training card (1) or is a procurement training instead (0).
- **vocational_training_hours** – number of hours that the vocational training was attended
- **vocational_training_plan_hours –** number of planned hours of the vocational training
- **last_2year_*** – indicators, which show whether the person has participated in a service/training during the 2 years preceding the period of unemployment
- **received_service_* –** indicators, which show whether the person has participated in a service/training within the unemployment period

Appendix 2: Additional descriptive statistics

Variable	Number of observations	Share in data	Share in treatment group	Share in treatment	Share in employment (Y)
Total obs.	**571 042**		**44754 (7.8%)**	**44 754 (7.8%)**	**369 906 (64.8%)**
First digit of last employment's ISCO code					
1	30 104	5.28%	9.29%	13.82%	65.8%
2	43 269	7.58%	10.44%	10.80%	73.1%
3	49 550	8.68%	13.36%	12.07%	70.4%
4	28 106	4.92%	5.65%	9.00%	70.1%
5	110 334	19.32%	16.11%	6.54%	68.9%
6	5 542	0.97%	0.57%	4.61%	60.7%
7	94 241	16.50%	15.61%	7.41%	64.1%
8	56 296	9.86%	12.80%	10.17%	65.8%
9	122 025	21.37%	13.44%	4.93%	59.9%
Missing	24 212	4.24%	1.78%	3.29%	37.8%
Unknown	7 363	1.29%	0.94%	5.73%	64.9%

Continuous variables (0 and missing values removed)	Total average	Treated group average	Untreated group average
Wage in previous 12 months of last employment period	916€	1117€	897€
Duration of last employment (days)	832	1052	813

Appendix 3: Variable importance table

Variable	Importance
average_salary_last_12_months	0.1857
relative_daily_rate	0.1060
genderF	0.1056
assigned_days_exact	0.0951
age_group_50-64	0.0729
last_employment_duration	0.0610
received_service_APPRENTICESHIP	0.0405
last_employment_end_reason _unknown	0.0393
received_service_GENERAL_SKILL	0.0381
assigned_benefits_duration	0.0295
last_employment_end_reason _due_to_employer	0.0190
age_group_30-49	0.0175
last_employment_ISCO_code_2	0.0174
education_level_5	0.0171
computer_skill_unknown	0.0150
disability_risk1	0.0147
last_employment_ISCO_code_8	0.0146
unemployed_days_3Yr	0.0134
last_employment_ISCO_code_5	0.0115
locationEE007	0.0107
received_service_CAREER_COUNSELING	0.0062
received_service_INFO_SESSION	0.0056
village1	0.0054
risk_Estonian_language1	0.0051
last_employment_ISCO_code_3	0.0039
last_employment_ISCO_code_7	0.0038
last_employment_end_reason _due_to_employee	0.0037
locationEE008	0.0036
received_service_WAGE_SUBSIDY	0.0036
received_service_WORKSHOP	0.0030
last_2_years_CAREER_COUNSELING	0.0030
unemployed_3Yr1	0.0028
last_employment_end_reason _due_to_employee_initiative	0.0023
computer_skill_INTERMEDIATE	0.0022
locationEE006	0.0020
last_employment_end_reason_by_mutual_agreement	0.0020
last_employment_ISCO_code_4	0.0019
education_level_3	0.0019
education_level_4	0.0018

computer_skill_SPECIALIST_LEVEL	0.0017
last_employment_ISCO_code_9	0.0016
received_service_Estonian_language_training	0.0016
locationEE004	0.0013
received_service_PSYCHOLOGICAL_COUNSELING	0.0010
received_service_TRIAL_WORK	0.0008
unemployed_3Yr2	0.0007
received_service_JOB_INFO_SESSION	0.0006
received_service_EURES_INFORMATION	0.0005
received_service_COMPUTER_TRAINING	0.0004
last_2_years_WORKSHOP	0.0004
received_service_JOB_CLUB	0.0002
received_service_DEBT_COUNSELING	0.0002
unemployed_3Yr3	0.0002
last_employment_end_reason_not_employed	0.0001
last_employment_ISCO_code_missing	0.0001
last_2_years_INFO_SESSION	0.0001
last_2_years_GENERAL_SKILL	0.0001
last_2_years_TRIAL_WORK	0.0001
last_2_years_WAGE_SUBSIDY	0.0001
last_2_years_APPRENTICESHIP	0.0000
received_service_JOB_REHABILITATION	0.0000
last_2_years_COMPUTER_TRAINING	0.0000
last_2_years_EURES_INFORMATION	0.0000
last_2_years_ESTONIAN_LANGUAGE_TRAINING	0.0000
last_2_years_WORK_PRACTICE	0.0000
last_2_years_PSYCHOLOGICAL_COUNSELING	0.0000
unemployed_3Yr4	0.0000
last_2_years_JOB_CLUB	0.0000
computer_skill_EXPERT_LEVEL	0.0000
received_service_TRAINING_FOR_WORKING_ABROAD	0.0000
last_employment_ISCO_code_unknown	0.0000
unemployed_3Yr5	0.0000
unemployed_3Yr6+	0.0000
computer_skill_missing	0.0000
last_2_years_JOB_REHABILITATION	0.0000
last_2_years_DEBT_COUNSELING	0.0000
received_service_WORK_PRACTICE	0.0000
(Intercept)	0.0000
last_employment_ISCO_code_6	0.0000
education_level_2	0.0000

education_level_unknown	0.0000
last_2_years_SUBSIDY_FOR_MINOR_EMPLOYMENT	0.0000
last_2_years_EURES_COUNSELING	0.0000
last_2_years_EURES_INFORMATION_EVENT	0.0000
last_2_years_ITR_PROCURABLE	0.0000
last_2_years_PROFESSIONAL_EXAM	0.0000
last_2_years_ADDICTION_COUNSELING	0.0000
last_2_years_DEGREE_STUDY_ALLOWANCE	0.0000
last_2_years_JOB_INFO_SESSION	0.0000
last_2_years_EMPLOYER_TRAINING_SUPPORT	0.0000
last_2_years_COMMUTING_SUPPORT	0.0000
last_2_years_SUPPORT_PERSON	0.0000
last_2_years_VOLUNTEER_WORK	0.0000
last_2_years_TRAINING_FOR_WORKING_ABROAD	0.0000
received_service_EURES_COUNSELING	0.0000
received_service_EURES_INFORMATION_EVENT	0.0000
received_service_ITR_PROCURABLE	0.0000
received_service_PROTECTED_WORK	0.0000
received_service_PEER_SUPPORT	0.0000
received_service_ADDICTION_COUNSELING	0.0000
received_service_DEGREE_STUDY_ALLOWANCE	0.0000
received_service_VOLUNTEER_WORK	0.0000

Appendix 4: Best linear projection results

Variable	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept) (*Base group age_group_16-29*)	0.0908	0.0039	23.4923	0.0000
age_group_30-49	-0.0031	0.0048	-0.6587	0.5101
age_group_50-64	0.0497	0.0055	9.0725	0.0000
(Intercept) (*Base group last_employment_ISCO_code_1*)	0.0994	0.0080	12.4874	0.0000
last_employment_ISCO_code_2	-0.0321	0.0099	-3.2434	0.0012
last_employment_ISCO_code_3	-0.0083	0.0097	-0.8506	0.3950
last_employment_ISCO_code_4	-0.0186	0.0113	-1.6479	0.0994
last_employment_ISCO_code_5	-0.0231	0.0092	-2.5089	0.0121
last_employment_ISCO_code_6	0.0197	0.0193	1.0221	0.3067
last_employment_ISCO_code_7	0.0220	0.0093	2.3770	0.0175
last_employment_ISCO_code_8	0.0292	0.0098	2.9823	0.0029
last_employment_ISCO_code_9	0.0140	0.0094	1.4856	0.1374
last_employment_ISCO_code_missing	0.0704	0.0176	3.9963	0.0001
last_employment_ISCO_code_unknown	0.0246	0.0173	1.4250	0.1542
(Intercept)	0.0935	0.0023	40.3679	0.0000
last_employment_duration	0.0000	0.0000	6.3692	0.0000
(Intercept) (*Base group unemployed_3Yr0*)	0.0986	0.0025	40.0692	0.0000
unemployed_3Yr1	0.0100	0.0049	2.0499	0.0404
unemployed_3Yr2	-0.0001	0.0072	-0.0088	0.9930
unemployed_3Yr3	0.0043	0.0099	0.4381	0.6613
unemployed_3Yr4	0.0196	0.0155	1.2633	0.2065
unemployed_3Yr5	-0.0161	0.0229	-0.7026	0.4823
unemployed_3Yr6+	-0.0148	0.0185	-0.8024	0.4223
(Intercept)	0.0958	0.0022	44.0528	0.0000
unemployed_days_3Yr	0.0001	0.0000	5.2245	0.0000
(Intercept)	0.1183	0.0028	42.0909	0.0000
assigned_days_exact	-0.0001	0.0000	-9.9444	0.0000
(Intercept)	0.1148	0.0026	44.1081	0.0000
relative_daily_rate	-0.0416	0.0044	-9.3912	0.0000
(Intercept)	0.1086	0.0022	48.7634	0.0000
assigned_benefits_duration	-0.0001	0.0000	-6.5122	0.0000
(Intercept) (*base group last_employment_end_reason _due_to_contract*)	0.0859	0.0037	22.9049	0.0000
last_employment_end_reason _not_employed	0.0838	0.0161	5.1915	0.0000
last_employment_end_reason _mutual_agreement	0.0221	0.0071	3.1123	0.0019
last_employment_end_reason _unknown	0.0699	0.0075	9.3176	0.0000
last_employment_end_reason _due_to_employer	-0.0071	0.0053	-1.3558	0.1752
last_employment_end_reason _due_to_employee_initiative	0.0097	0.0065	1.4914	0.1359

last_employment_end_reason _due_to_employee	0.0295	0.0067	4.3827	0.0000
(Intercept)	0.1191	0.0030	39.7062	0.0000
average_salary_last_12_months	-0.0000	0.0000	-9.1425	0.0000
(Intercept)	0.1242	0.0028	44.4220	0.0000
genderF	-0.0464	0.0039	-11.8975	0.0000
(Intercept)	0.0954	0.0022	43.8657	0.0000
risk_Estonian_language1	0.0280	0.0049	5.7152	0.0000
(Intercept)	0.0947	0.0021	45.3581	0.0000
disability_risk1	0.0496	0.0059	8.4660	0.0000
(Intercept)	0.0980	0.0022	45.2593	0.0000
village1	0.0164	0.0050	3.2886	0.0010
(Intercept) (*Base group locationEE001*)	0.0800	0.0031	25.8261	0.0000
locationEE004	0.0297	0.0064	4.6465	0.0000
locationEE006	0.0315	0.0071	4.4471	0.0000
locationEE007	0.0488	0.0060	8.1607	0.0000
locationEE008	0.0354	0.0050	7.1228	0.0000
(Intercept) (*Base group education_level_1*)	0.1242	0.0047	26.2462	0.0000
education_level_2	0.0014	0.0164	0.0845	0.9326
education_level_3	-0.0231	0.0058	-4.0078	0.0001
education_level_4	-0.0239	0.0059	-4.0154	0.0001
education_level_5	-0.0500	0.0068	-7.3579	0.0000
education_level_unknown	-0.0060	0.0178	-0.3385	0.7350
(Intercept)	0.1290	0.0052	24.9806	0.0000
computer_skill_EXPERT_LEVEL	-0.0274	0.0202	-1.3573	0.1747
computer_skill_INTERMEDIATE	-0.0262	0.0065	-4.0063	0.0001
computer_skill_MISSING	0.0924	0.0164	5.6448	0.0000
computer_skill_SPECIALIST_LEVEL	-0.0220	0.0109	-2.0134	0.0441
computer_skill_unknown	-0.0396	0.0058	-6.8472	0.0000
(Intercept)	0.1009	0.0020	51.6501	0.0000
last_2_years_SUBSIDY_FOR_MINOR_EMPLOYMENT	0.0786	0.0339	2.3204	0.0203
(Intercept)	0.1011	0.0020	51.6223	0.0000
last_2_years_COMPUTER_TRAINING	0.0018	0.0227	0.0796	0.9366
(Intercept)	0.1012	0.0020	51.5790	0.0000
last_2_years_ESTONIAN_LANGUAGE_TRAINING	-0.0098	0.0179	-0.5463	0.5849
(Intercept)	0.1010	0.0020	51.6977	0.0000
last_2_years_EURES_COUNSELING	0.0389	0.0357	1.0914	0.2751
(Intercept)	0.1011	0.0020	51.4503	0.0000
last_2_years_EURES_INFORMATION	-0.0005	0.0163	-0.0281	0.9776
(Intercept)	0.1011	0.0020	51.8171	0.0000
last_2_years_EURES_INFORMATION_EVENT	-0.0318	0.0633	-0.5033	0.6148

(Intercept)	0.1012	0.0020	51.4255	0.0000
last_2_years_INFO_SESSION	-0.0044	0.0145	-0.3030	0.7619
(Intercept)	0.1010	0.0020	51.7511	0.0000
last_2_years_ITR_PROCURABLE	0.0579	0.0483	1.1987	0.2306
(Intercept)	0.1026	0.0020	50.2983	0.0000
last_2_years_CAREER_COUNSELING	-0.0144	0.0068	-2.1202	0.0340
(Intercept)	0.1009	0.0020	51.7128	0.0000
last_2_years_PROFESSIONAL_EXAM	0.1120	0.0578	1.9356	0.0529
(Intercept)	0.1009	0.0020	51.3842	0.0000
last_2_years_WAGE_SUBSIDY	0.0107	0.0167	0.6428	0.5203
(Intercept)	0.1014	0.0020	51.4390	0.0000
last_2_years_TRIAL_WORK	-0.0177	0.0133	-1.3288	0.1839
(Intercept)	0.1010	0.0020	51.5843	0.0000
last_2_years_PSYCHOLOGICAL_COUNSELING	0.0108	0.0226	0.4785	0.6323
(Intercept)	0.1010	0.0020	51.7272	0.0000
last_2_years_ADDICTION_COUNSELING	0.0767	0.0423	1.8112	0.0701
(Intercept)	0.1010	0.0020	51.7056	0.0000
last_2_years_DEGREE_STUDY_ALLOWANCE	0.0239	0.0344	0.6943	0.4875
(Intercept)	0.1005	0.0020	51.2668	0.0000
last_2_years_JOB_INFORMATION_SESSION	0.0576	0.0184	3.1252	0.0018
(Intercept)	0.1011	0.0020	51.7288	0.0000
last_2_years_JOB_REHABILITATION	-0.0099	0.0275	-0.3604	0.7186
(Intercept)	0.1010	0.0020	51.7105	0.0000
last_2_years_EMPLOYER_TRAINING_SUPPORT	0.0397	0.0386	1.0282	0.3039
(Intercept)	0.1008	0.0020	51.4855	0.0000
last_2_years_WORK_PRACTICE	0.0331	0.0222	1.4927	0.1355
(Intercept)	0.1013	0.0020	51.6600	0.0000
last_2_years_JOB_CLUB	-0.0210	0.0190	-1.1064	0.2685
(Intercept)	0.1012	0.0020	51.8696	0.0000
last_2_years_COMMUTING_SUPPORT	-0.0785	0.0501	-1.5683	0.1168
(Intercept)	0.1013	0.0020	51.5696	0.0000
last_2_years_APPRENTICESHIP	-0.0127	0.0167	-0.7632	0.4453
(Intercept)	0.1012	0.0020	50.5253	0.0000
last_2_years_WORKSHOP	-0.0027	0.0088	-0.3096	0.7569
(Intercept)	0.1012	0.0020	51.8282	0.0000
last_2_years_SUPPORT_PERSON	-0.0640	0.0503	-1.2721	0.2033
(Intercept)	0.1011	0.0020	51.5956	0.0000
last_2_years_GENERAL_SKILL	-0.0031	0.0201	-0.1565	0.8756
(Intercept)	0.1010	0.0020	51.7375	0.0000
last_2_years_VOLUNTARY_WORK	0.0334	0.0422	0.7910	0.4290

(Intercept)	0.1010	0.0020	51.6798	0.0000
last_2_years_TRAINING_FOR_WORKING_ABROAD	0.0272	0.0334	0.8155	0.4148
(Intercept)	0.1012	0.0020	51.7630	0.0000
last_2_years_DEBT_COUNSELING	-0.0136	0.0253	-0.5390	0.5899
(Intercept)	0.1003	0.0020	50.6593	0.0000
received_service_COMPUTER_TRAINING	0.0274	0.0116	2.3698	0.0178
(Intercept)	0.1025	0.0020	51.7497	0.0000
received_service_ESTONIAN_LANGUAGE_TRAINING	-0.0417	0.0112	-3.7218	0.0002
(Intercept)	0.1012	0.0020	51.7705	0.0000
received_service_EURES_COUNSELING	-0.0211	0.0306	-0.6893	0.4906
(Intercept)	0.1016	0.0020	51.3531	0.0000
received_service_EURES_INFORMATION	-0.0165	0.0116	-1.4181	0.1562
(Intercept)	0.1011	0.0020	51.7704	0.0000
received_service_EURES_INFORMATION_EVENT	-0.0051	0.0413	-0.1231	0.9020
(Intercept)	0.1025	0.0020	50.9690	0.0000
received_service_INFO_SESSION	-0.0188	0.0081	-2.3366	0.0195
(Intercept)	0.1011	0.0020	51.8149	0.0000
received_service_ITR_PROCURABLE	-0.0469	0.0488	-0.9607	0.3367
(Intercept)	0.1013	0.0020	51.8930	0.0000
received_service_PROTECTED_WORK	-0.2027	0.0340	-5.9691	0.0000
(Intercept)	0.1052	0.0023	44.8923	0.0000
received_service_CAREER_COUNSELING	-0.0119	0.0042	-2.8195	0.0048
(Intercept)	0.1009	0.0020	51.6962	0.0000
received_service_EXPERIENCE_COUNSELING	0.0686	0.0445	1.5418	0.1231
(Intercept)	0.1014	0.0020	49.8518	0.0000
received_service_WAGE_SUBSIDY	-0.0064	0.0050	-1.2659	0.2055
(Intercept)	0.1037	0.0020	51.8936	0.0000
received_service_TRIAL_WORK	-0.0685	0.0089	-7.6638	0.0000
(Intercept)	0.1017	0.0020	51.4976	0.0000
received_service_PSYCHOLOGICAL_COUNSELING	-0.0191	0.0122	-1.5629	0.1181
(Intercept)	0.1009	0.0020	51.6977	0.0000
received_service_ADDICTION_COUNSELING	0.0796	0.0428	1.8616	0.0627
(Intercept)	0.1010	0.0020	51.7907	0.0000
received_service_DEGREE_STUDY_ALLOWANCE	0.0101	0.0414	0.2435	0.8077
(Intercept)	0.1009	0.0020	51.0131	0.0000
received_service_JOB_INFORMATION_SESSION	0.0058	0.0118	0.4955	0.6202
(Intercept)	0.1006	0.0020	51.3826	0.0000
received_service_JOB_REHABILITATION	0.0378	0.0206	1.8380	0.0661
(Intercept)	0.1008	0.0020	51.3242	0.0000
received_service_WORK_PRACTICE	0.0257	0.0171	1.5017	0.1332

(Intercept)	0.1008	0.0020	50.8995	0.0000
received_service_JOB_CLUB	0.0089	0.0114	0.7838	0.4332
(Intercept)	0.1032	0.0020	51.9705	0.0000
received_service_APPRENTICESHIP	-0.0445	0.0101	-4.3939	0.0000
(Intercept)	0.1020	0.0021	47.6371	0.0000
received_service_WORKSHOP	-0.0052	0.0052	-0.9981	0.3182
(Intercept)	0.1040	0.0020	52.0031	0.0000
received_service_GENERAL_SKILL	-0.0465	0.0089	-5.2435	0.0000
(Intercept)	0.1011	0.0020	51.7021	0.0000
received_service_VOLUNTARY_WORK	-0.0031	0.0278	-0.1124	0.9105
(Intercept)	0.1015	0.0020	51.9127	0.0000
received_service_ TRAINING_FOR_WORKING_ABROAD	-0.0512	0.0260	-1.9673	0.0492
(Intercept)	0.1014	0.0020	51.9606	0.0000
received_service_DEBT_COUNSELING	-0.0215	0.0209	-1.0284	0.3038

Appendix 5: Additional figures

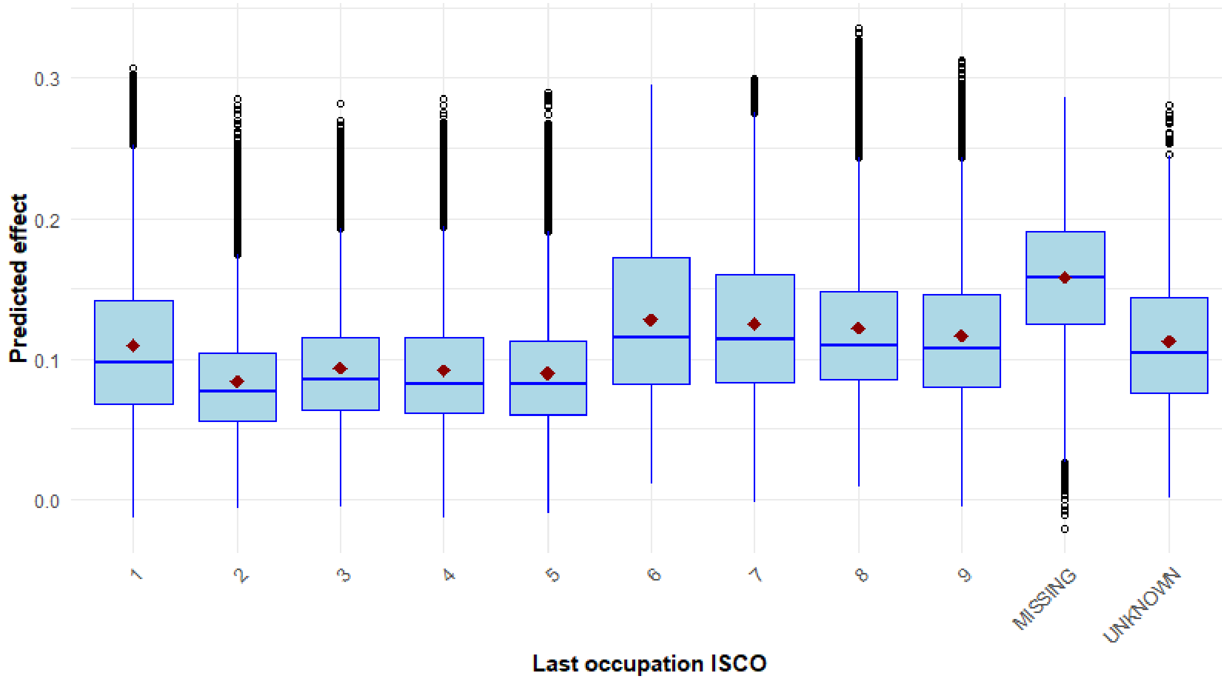

Figure 9: Boxplot showing the predicted effect of vocational training by last occupation

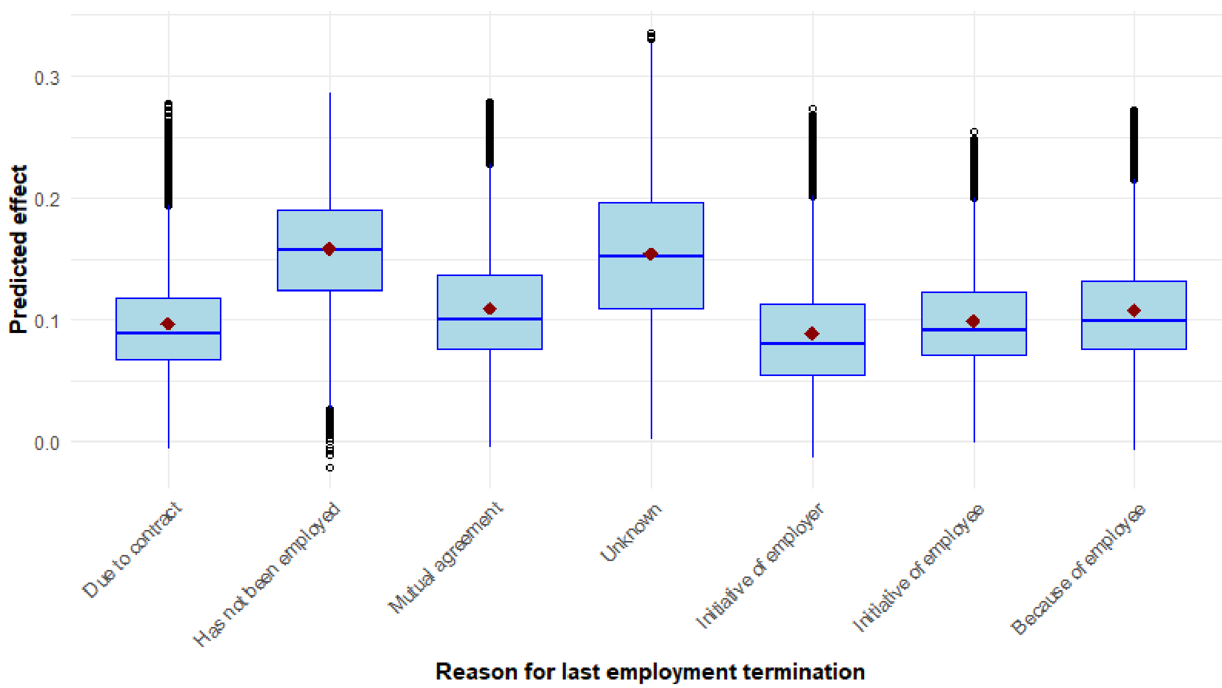

Figure 10: Boxplot showing the predicted effect of vocational training by reason for last employment termination

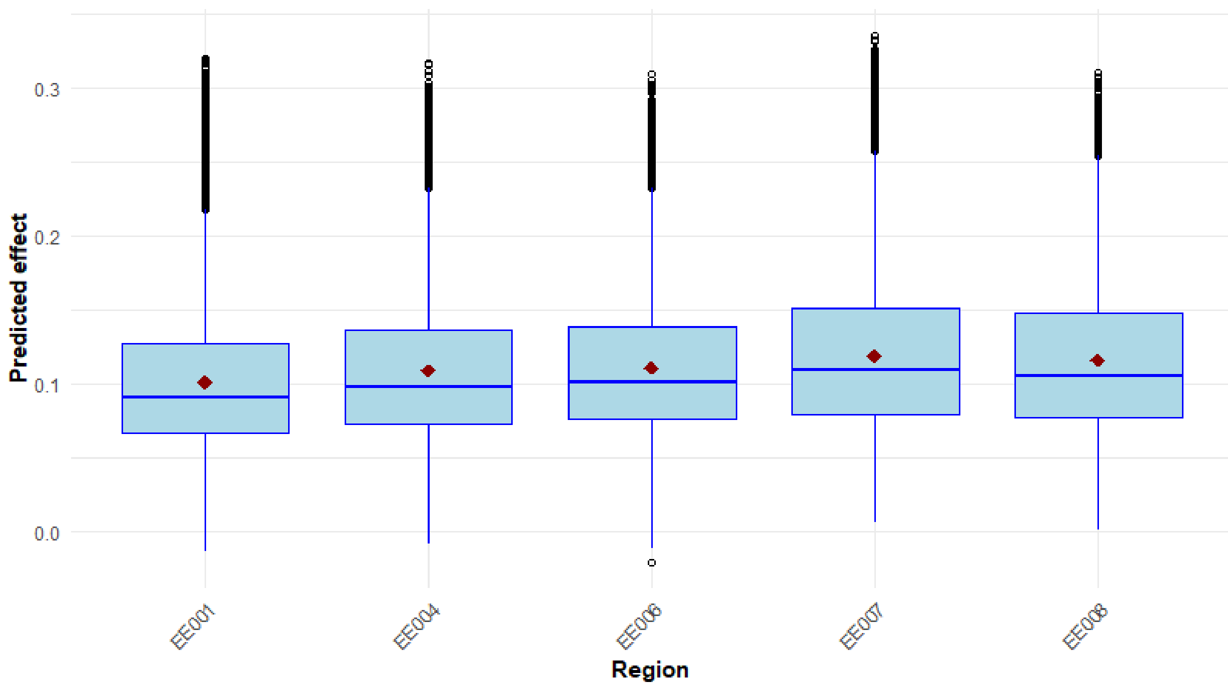

Figure 11: Boxplot showing the predicted effect of vocational training by region

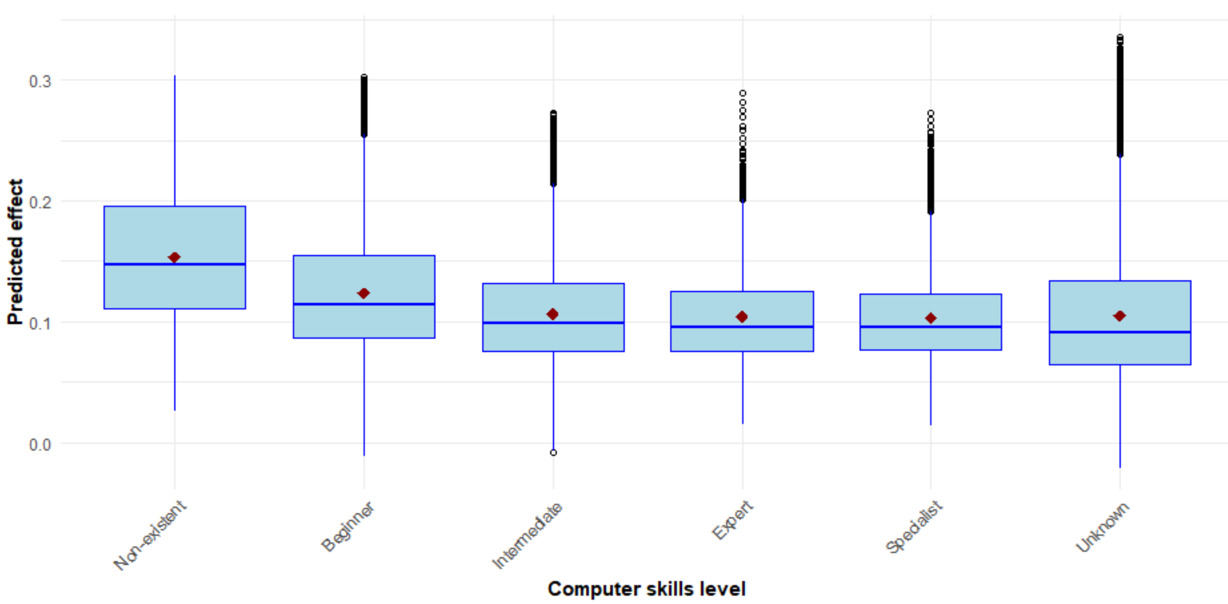

Figure 12: Boxplot showing the predicted effect of vocational training by computer skills level

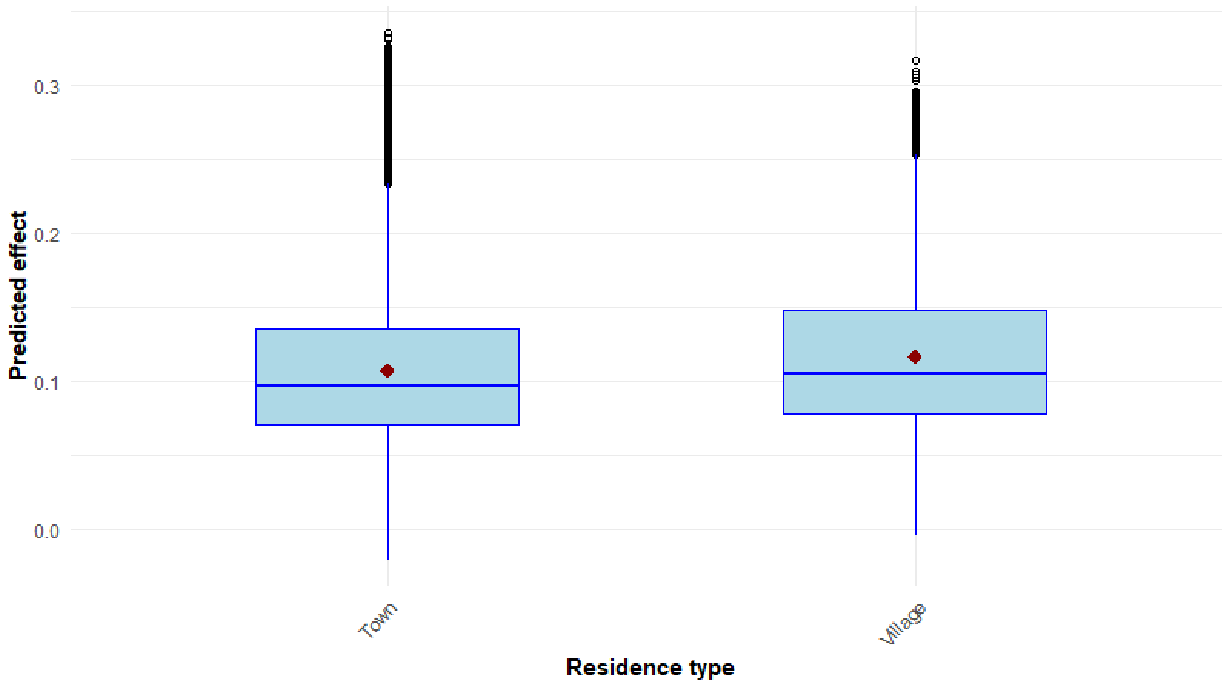

Figure 13: Boxplot showing the predicted effect of vocational training by residence type

Figure 14: Boxplot showing the predicted effect of vocational training by average salary of last 12 months of employment

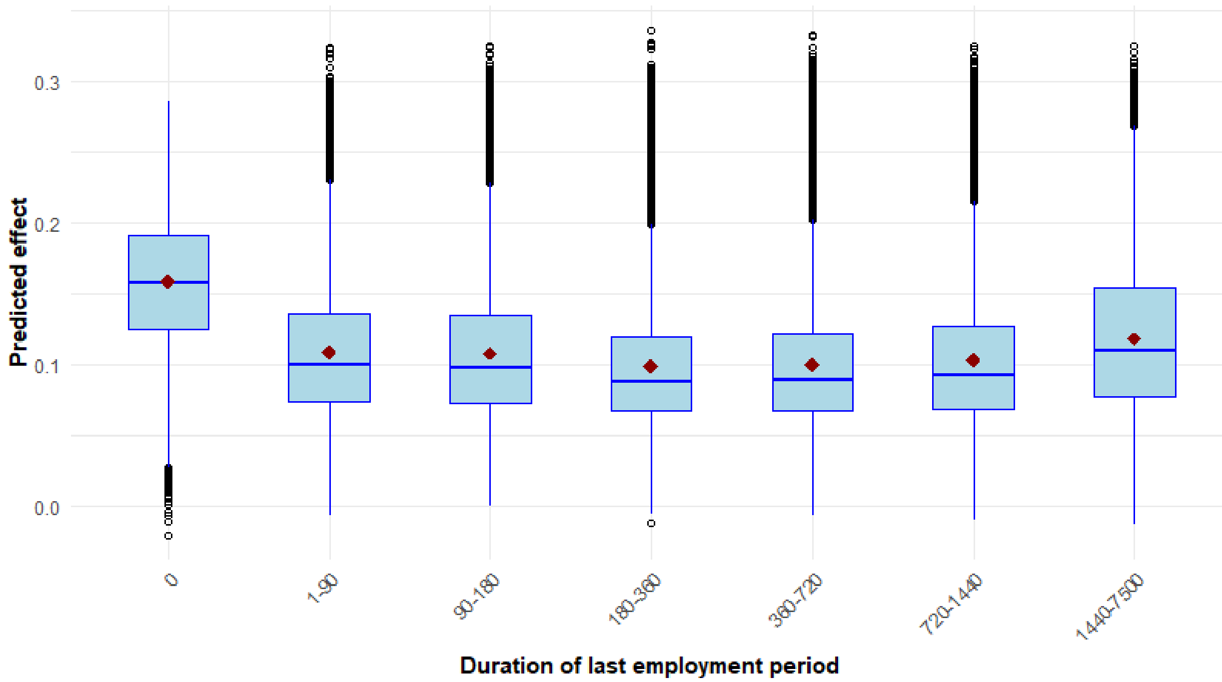

Figure 15: Boxplot showing the predicted effect of vocational training by duration of last employment period

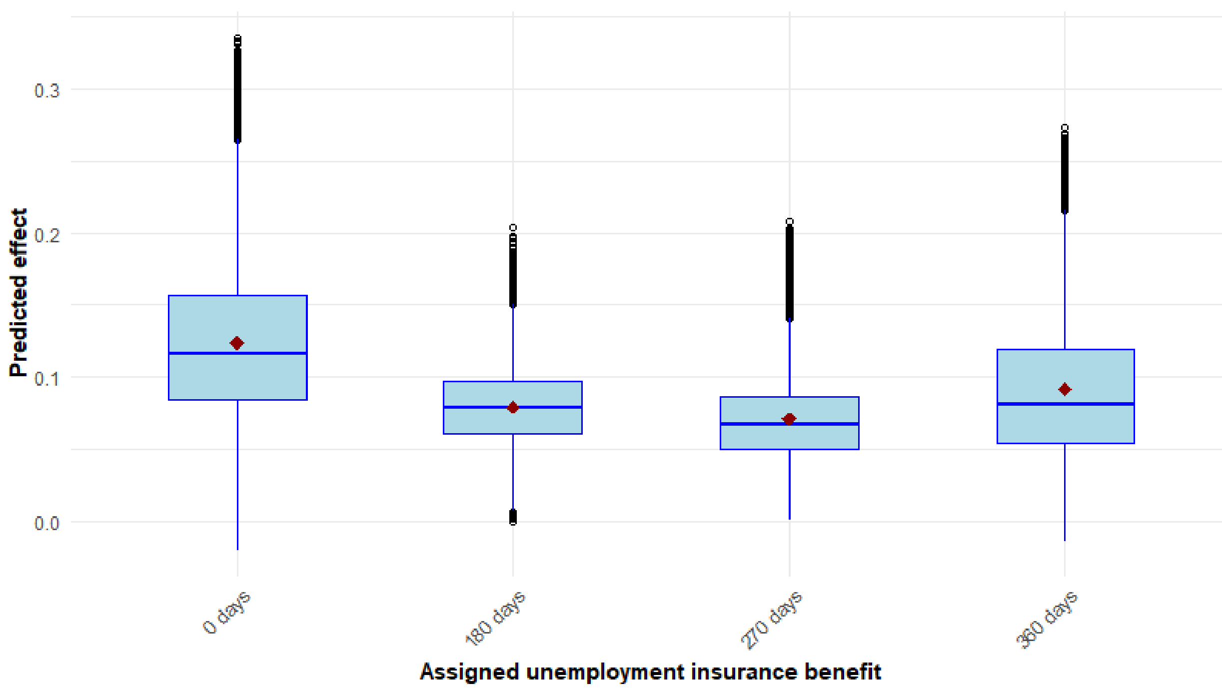

Figure 16: Boxplot showing the predicted effect of vocational training by duration of assigned unemployment insurance benefit

Appendix 6: Propensity Score Matching balance table and sample sizes

<u>Balance measures</u>

Variable	Type	Diff. Adj
distance	Distance	0.0000
age_group_16-29	Binary	-0.0070
age_group_30-49	Binary	0.0016
age_group_50-64	Binary	0.0054
genderF	Binary	0.0046
average_salary_last_12_months	Continuous	0.0068
risk_Estonian_language	Binary	0.0040
disability_risk	Binary	0.0058
last_employment_ISCO_code_1	Binary	0.0012
last_employment_ISCO_code_2	Binary	-0.0009
last_employment_ISCO_code_3	Binary	0.0003
last_employment_ISCO_code_4	Binary	0.0008
last_employment_ISCO_code_5	Binary	-0.0014
last_employment_ISCO_code_6	Binary	0.0001
last_employment_ISCO_code_7	Binary	-0.0004
last_employment_ISCO_code_8	Binary	-0.0042
last_employment_ISCO_code_9	Binary	0.0046
last_employment_ISCO_code_missing	Binary	-0.0003
last_employment_ISCO_code_unknown	Binary	0.0000
relative_daily_rate	Continuous	0.0049
last_employment_duration	Continuous	0.0130
last_employment_end_reason _due_to_contract	Binary	-0.0022
last_employment_end_reason _not_employed	Binary	-0.0003
last_employment_end_reason _mutual_agreement	Binary	-0.0008
last_employment_end_reason _unknown	Binary	0.0017
last_employment_end_reason _due_to_employer	Binary	-0.0005
last_employment_end_reason _due_to_employee_initiative	Binary	0.0008
last_employment_end_reason _due_to_employee	Binary	0.0013
assigned_days_exact	Continuous	0.0018
received_service_APPRENTICESHIP	Binary	0.0049
received_service_GENERAL_SKILL	Binary	0.0047
education_level_1	Binary	0.0045
education_level_2	Binary	0.0008
education_level_3	Binary	-0.0051
education_level_4	Binary	-0.0003
education_level_5	Binary	0.0001
education_level_unknown	Binary	0.0000

Sample sizes

	Control	**Treated**
All	526788	44754
Matched (ESS)	91219.6	44752
Matched (Unweighted)	109053	44752
Unmatched	417023	2
Discarded	212	0

Resümee

Tööturukoolituste mõju hindamine töölesaamise tõenäosusele põhjusliku juhumetsa meetodit kasutades

Antud magistritöös analüüsitakse tööturukoolituste mõju töölesaamise tõenäosusele põhjusliku juhumetsa meetodit kasutades, mis on masinõppealgoritm loodud välise sekkumise mõju tõhusaks hindamiseks. Analüüsi peamiseks eesmärgiks on välja selgitada, missugust mõju tööturukoolitused edasisele töölesaamisele avaldavad ning kuidas tööturukoolituse võimalusi paremini ümber jaotada, et võimalikult palju inimesi tagasi tööle saada.

Töös kasutatakse Eesti Töötukassa anonümiseeritud registriandmeid ajavahemikust 2015 – 2023. Andmestik sisaldab peale sotsiaaldemograafiliste tunnuste infot ka varasemalt saadud tööturuteenuste ja koolituste kohta. Andmestikus on üle poole miljoni töötuse instantsi, nendest osales tööturukoolitusel veidi alla kaheksa protsendi. Andmetest selgub, et tööturukoolitusi saavad suurema tõenäosusega nii kõrgemalt haritud, eesti keelt hästi kõnelevad kui ka puuderiskita inimesed.

Mudel näitab, et tööturukoolitustel on töölesaamisele positiivne mõju, tõstes töölesaamise tõenäosust keskmiselt 9.7 protsendipunkti. Ühtlasi selgub mudelist, et tööturukoolitustel on suurem positiivne mõju töölesaamisele väiksema haridustasemega, eesti keele riskiga ja puuetega inimeste puhul. Seega soovitab töö tulevikus jaotada tööturukoolitusi tihemini nende karakteristikutega inimestele, et tööturukoolituste mõju maksimeerida.

Märksõnad: tööturukoolitus, töötus, töölesaamise tõenäosus, põhjusliku juhumetsa meetod, Eesti

JEL Klassifikatsioon: J24, J64, J65

CERCS Klassifikatsioon: S220, S230

Non-exclusive license to reproduce thesis and make thesis public

I, ***Martin Piir***, born on 22^{nd} of April 1996

1. herewith grant the University of Tartu a free permit (non-exclusive license) to:

1.1. reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and

1.2. make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

Using causal random forest to evaluate vocational training programmes on employment probability in Estonia

supervised by **Andres Võrk** and **Taavi Unt**,

2. I am aware of the fact that the author retains these rights.

3. I certify that granting the non-exclusive license does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu, 21^{st} of May 2024