

Kivi, Laura Helena; Sömer, Marko; Kallaste, Epp

**Article**

## Language training for unemployed non-natives: Who benefits the most?

Baltic Journal of Economics

**Provided in Cooperation with:**

Baltic International Centre for Economic Policy Studies (BICEPS), Riga

*Suggested Citation:* Kivi, Laura Helena; Sömer, Marko; Kallaste, Epp (2020) : Language training for unemployed non-natives: Who benefits the most?, Baltic Journal of Economics, ISSN 2334-4385, Taylor & Francis, London, Vol. 20, Iss. 1, pp. 34-58, <https://doi.org/10.1080/1406099X.2020.1740403>

This Version is available at:

<https://hdl.handle.net/10419/267576>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>

# Language training for unemployed non-natives: who benefits the most?

Laura Helena Kivi, Marko Sömer & Epp Kallaste

To cite this article: Laura Helena Kivi, Marko Sömer & Epp Kallaste (2020) Language training for unemployed non-natives: who benefits the most?, *Baltic Journal of Economics*, 20:1, 34-58, DOI: [10.1080/1406099X.2020.1740403](https://doi.org/10.1080/1406099X.2020.1740403)

To link to this article: <https://doi.org/10.1080/1406099X.2020.1740403>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 19 Mar 2020.



Submit your article to this journal [↗](#)



Article views: 900



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

# Language training for unemployed non-natives: who benefits the most?

Laura Helena Kivi<sup>a,b</sup>, Marko Sömer<sup>a</sup> and Epp Kallaste<sup>a</sup>

<sup>a</sup>Estonian Center for Applied Research (CentAR), Tallinn, Estonia; <sup>b</sup>School of Economics and Business Administration, University of Tartu, Tartu, Estonia

## ABSTRACT

This study evaluates the local language training aimed at the unemployed in Estonia during 2015–2016. The impact of training on employment probability and labour income is estimated by combining propensity score matching with coarsened exact matching. The impact on the probability of being employed is found to be positive after the end of the lock-in effect. Two years after the start of the language training the effect is around 8 pp. The initial lock-in effect is smaller for more flexible and shorter courses, for those with lower initial level of language skills and for those living outside of the capital region. The long-term effect is higher for those with lower level of initial language skills and does not differ by the course type or region. The results indicate that the local language training helps the unemployed non-natives to find employment, but does not give them access to higher-paying positions.

## ARTICLE HISTORY

Received 14 June 2019  
Accepted 2 March 2020

## KEYWORDS

Language skills; immigrant population; unemployment; programme evaluation; matching

## JEL CLASSIFICATION

C21; J24; J61; J64

## 1. Introduction

The analysis of the effectiveness of active labour market programmes (ALMPs) has spread rapidly since the beginning of 2000s. The effectiveness of ALMPs may depend on the economic situation, the socio-demographic composition of the unemployed and the division of resources between labour market programmes and access to these. The recent inflow of migrants to the EU has re-raised the question of effectiveness and efficiency of ALMPs targeted at the immigrant population. Many OECD countries see language training as the key factor of labour market integration (OECD, 2018, pp. 100–101). The current article evaluates the effectiveness of local language training of the unemployed in Estonia during 2015–2016.<sup>1</sup>

In 2018, 32% of active adult population (aged 15–74) in Estonia live in non-Estonian speaking households, while among the unemployed 47% are from non-Estonian speaking households (Statistics Estonia, 2019b, Table TT136). As there is a vast amount of literature (e.g. Aldashev et al., 2009; Budria et al., 2017; Dustmann & Fabbri, 2003) documenting language skills premium of employment probability and income in the labour market,

**CONTACT** Laura Helena Kivi  [laura.kivi@centar.ee](mailto:laura.kivi@centar.ee)

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

the natural choice for increasing labour market outcomes for the non-Estonian speaking population is to offer them state language courses.

However, while language skills are found to hold premium of employment probability and income, the effectiveness of language courses is not uniformly confirmed. The scarce international literature investigating the impact of state language training on labour market outcomes mostly finds that the impact on employment probability is positive (e.g. Clausen et al., 2009; Prey, 2000).

While the literature generally finds positive impact of the language training on the employment probability, this effect occurs only in the long-term. The studies that investigate the impact of language training over a wider length of observation period find considerable lock-in effect during the first months after the start of the language training (Delander et al., 2005; Gerfin & Lechner, 2002; Prey, 2000). The positive effect generally occurs 6–12 months after the start of the training (Delander et al., 2005; Prey, 2000). The lock-in effect occurs while the individuals participating in the training lower their job search activity, which results in smaller possibility of entering to employment in the first months from the start of training. The additional months spent in unemployment are costly both for the state and the individual.

To increase the effectiveness of language courses different options to shorten the lock-in effect and increase the long-term gain could be considered. First possibility is to evaluate, whether participants who attend a certain type of language course yield more favourable results. Language courses might differ from their overall length, flexibility of organization, content and quality of course providers.

Shorter courses might lead to speedier transfers from unemployment to employment, however, they might hold smaller long-run effects. The heterogeneity of effects of shorter labour market training programmes vs. longer training programmes has been investigated in different studies, which mostly find that shorter training programmes yield better short-term outcomes (e.g. Fitzenberger et al., 2008; Osikominu, 2013; Stephan & Pahnke, 2011). In the long term the results are mixed. Some studies find that longer programmes result in greater long-term impact (Fitzenberger et al., 2010; Osikominu, 2013; Stephan & Pahnke, 2011), while others show no difference in long-term effect (Fitzenberger et al., 2008) or find participants of shorter trainings to perform better in the long-run (McGuinness et al., 2014). However, most of these studies concentrate on comparing training programmes with very different structure and content, e.g. long retraining programmes compared to short firm-internal training (Stephan & Pahnke, 2011). Kluve et al. (2012) focus on the effect of different levels of duration for the similar types of training programme and for the continuous levels of duration. They show promising results in the favour of shorter courses. Using a dose–response function it is found that after the time period of 150 days the additional hours of training do not yield additional treatment effects on employment probability. Thus, it is possible that shorter training courses do not result in less beneficial long-run effects as longer training.

Besides the differences in the type of training, some groups might benefit more from the language training than others due to differences in individual characteristics of the participants or of external environment.

In the case of language training two factors are of particular interest. Firstly, the effect of language training is likely to depend on the initial level of language skills. Heterogenous effects of the local language training by initial levels of skills have been estimated by

Clausen et al. (2009). However, the authors argue that these effects are not causal as they are unable to control for the level of education. The level of language skills is additionally important because of the institutional context, e.g. certain position requires the state language skills of some specific levels (B1, B2). Secondly, the economic and linguistic environment might affect the impact of language courses. It has been shown that labour market training courses are more effective at times of recession, when unemployment is high and the number of vacancies is low (see e.g. Lechner & Wunsch, 2009). Along the same lines, the impact of training in regions with high labour demand might differ from the impact in regions with less favourable labour market characteristics. In addition to the labour market environment, the linguistic environment of the region is of importance in case of language training. In the language enclaves, the working language can differ from the state language and thus the language courses might yield less beneficial results. At the same time, state language skills provide comparative advantage over the other residents of the region to compete for the positions, which do require state language skills.

Overall, to increase the effectiveness of language course provision it is necessary to determine, whether participants of some types of training courses, some participant groups with specific individual characteristics or external environment benefit more from the language training.

The rich administrative data set used in this study allows us to shed further light to the issue. First of all, Estonian Unemployment Insurance Fund (UIF) offers language training courses through two different channels, which result in courses of different lengths and levels of flexibility. Thus, we are able to observe the differences in the impact of language courses of different types affecting the labour market outcomes. We contribute to the results of language training effectiveness in two additional ways. Firstly, we allow for heterogeneous impact of training for participants with different prior language skills. Secondly, the effects of language training courses are estimated for different regions to account for differences in language and labour market environment. The dataset used allows us to differ between the capital region, a region with high share of non-natives and other regions. The region with high share of non-natives is an industrial area also characterized by high unemployment rate and lower level of vacancies.

We estimate the effect of Estonian language courses on employment probability and wages by using matched treatment and control groups. The data used in this analysis is the individual level data from the UIF's registry, which is linked to the income data from the Estonian Tax and Customs Board (ETCB). The sample includes all the persons who over the period of 2015–2016 were at least once registered as unemployed and participated in an Estonian language course provided by the UIF (the treatment group) and those who did not participate in a course, but whose main language of communication was not Estonian (the control group). The counterfactuals for the treatment group are found by combining the propensity score matching with coarsened exact matching (Rubin & Thomas, 2000). The average treatment effect on the treated is evaluated by using the logit and linear regression models on the matched sample.

The results show that the impact of Estonian language training on the probability to be employed is significant and positive after the first 11 months from the start of the course. Two years after the start of the course the effect is around 8 pp. We found no statistically significant effect on wage income. Our results indicate that while language training may

help the unemployed non-natives find employment it does not give them access to higher-paying positions. In addition, we estimate the effect on employment by two different types of courses, by the initial level of language proficiency and by the residence region of the participant. The initial lock-in effect is found to be shorter for more flexible and shorter courses, while the long-term effect does not differ by the course type. The results for different prior language proficiency levels indicate that participants with the lowest prior language skills benefit the most from the courses. The results by regions show that the lock-in effect is the largest in the capital region.

The paper is structured as follows: the next section summarizes previous studies on the topic, Section 3 describes the institutional and demographic background, Section 4 explains the methodological approach, Section 5 presents the results of the analysis, and lastly, Section 6 provides the conclusion.

## 2. Earlier research on ALMP language training

Extensive literature has assessed the impact of different active labour market programmes on labour market outcomes of an individual, such as employment probability and earnings. Kluve (2010) provides meta-analysis of an impact of 137 programmes and concludes that the effect of training programmes on employment probability is modestly positive.

A few meta-analyses focus on the evaluation of active labour market programmes targeted at migrants. Overall, the types of ALMPs that are found to be effective in decreasing the duration of unemployment for the unemployed in general are also most effective for unemployed migrants (see for an overview Butschek & Walter, 2014; Rinne, 2012).

Studies focusing on the impact of language courses generally find that participation in a language course increases the probability of employment after the initial lock-in period. Prey (2000) evaluates the impact of language training during a relatively short time period, up to 6 months after the end of the programme, and shows that there is a significant positive effect on the employment probability starting from the 3rd month after the end of a German language course. While the language courses last between 10 and 12 weeks, the positive effect occurs about 6 months after the end of the training. Delander et al. (2005) show based on Swedish data that participation in a pilot programme that combines work-oriented language teaching and practical workplace training results in a faster transfer from open unemployment to employment, training and education. When considering only the transfers to employment, the authors find that programme participants have lower probability to remain unemployed starting from about one year after entering unemployment. Clausen et al. (2009) use the time-of-events-duration model on Danish data of newly-arrived immigrants and find that the increase in language proficiency of language course participants has a positive significant effect on the hazard rate to employment. Gerfin and Lechner (2002) use Swiss data and show that language courses reduce the chance of employment compared to non-participation. The negative effect might be explained by the length of observation period. While the effect remains negative up until the end of the observation period of one year, the absolute value of the negative effect decreases starting from the 8th month after the start of the training. Thus, while the authors find short-term effect of language courses to be negative, the length of

observation period does not allow to determine whether the initial lock-in effect would remain or disappear in the long run.

While the positive effect of language courses on the probability of employment is recorded in several studies, the effect on earnings is more ambiguous and investigated only in a few studies. Hayfron (2001) shows that participation in language courses has an effect on language proficiency, but the latter does not significantly affect earnings. However, they do not control for pre-course language proficiency, which may lead to biased estimates. Sarvimäki and Hämäläinen (2016) exploit the discontinuity in the provision of active labour market programmes to estimate the long-term effect of ALMPs on migrants in Finland. A reform introducing compulsory integration plans increased the time spent on language courses and on other training specifically aimed at migrants while reducing the proportion of traditional ALMPs. The total earnings of compliers over the 10-year follow-up period are found to be 47% higher compared to the expected outcome of compliers in case of no treatment.

Theoretical considerations of the impact of language skills can firstly be based on human capital theory (Becker, 1975). According to human capital theory, investments in human capital transmit into more favourable labour market outcomes, such as higher earnings or employment. Language skills are a vital part of human capital and important for the success in the host country's labour market for two main reasons. Firstly, language skill in itself is a productive trait: it allows for communication and social interaction to obtain relevant information (Hayfron, 2001). Secondly, language skills enable to transfer the pre-immigration human capital to the host country's labour market. Orlov (2017) estimates the effect of attending an English language course in Canada and finds that over half of the impact that language skills have on wage growth is driven by the transfer of pre-immigration cognitive skills into the host country's labour market. The second theoretical consideration explaining the costs of not speaking the language of majority relates to the theory of language discrimination, started by Lang (1986). According to the theory learning the second language is costly and the competitive market will tend to minimize communication through segregating the speakers of different languages. If interaction is required, then the minority will bear the cost.

While research on the effect of language courses is not so common, the impact of language proficiency on labour market outcomes has been widely investigated. The importance of fluency and literacy in the host country language for higher employment probability and earnings has been shown for various countries, e.g. the UK (Dustmann & Fabbri, 2003), the US (Bleakley & Chin, 2004), Germany (Aldashev et al., 2009; Beyer, 2016; Dustmann & van Soest, 2002), Australia (Chiswick et al., 2005), Spain (Budria et al., 2017), and Israel (Cohen-Goldner & Eckstein, 2008).

Although numerous studies have found that local language skills hold reasonable wage premium, Toomet (2011) argues that this is not the case for all immigrant groups in Estonia and Latvia. Using Labour Force Survey (LFS) data for ethnic Russian men in Estonia and Latvia and by controlling a wide variety of characteristics, he finds that local language skills have an effect on earnings only for the lower end of wage distribution and for the ones holding public administration positions.

Earlier studies based on the Estonian data indicate that labour market training has a positive effect on the probability of employment after the initial lock-in period of 3–5



months (Anspal et al., 2012; Leetmaa et al., 2003). The same studies found no significant effect on wages during the observation period of two years.

Lauringson et al. (2011) is, to the best of our knowledge, the only study evaluating the impact of Estonian language training separately from other labour market training courses in Estonia. Using propensity score matching, authors find that the impact of Estonian language courses varies by the end year of the course. Language courses had a positive impact on employment probability for those who finished in 2010, while the impact for those who finished in 2009 is not statistically significant. The authors claim that the difference results from a policy change. In 2010–2011 only work-related Estonian language courses were offered, whereas a year before the rules for course offerings were less restrictive. At the same time, their matched sample sizes are also very small (58 for courses that ended in 2009 and 82 for courses that ended in 2010), the insignificant results for 2009 might be due to an insufficient sample size. Our study differs from the work of Lauringson et al. (2011) in three distinctive ways. Firstly, we are able to use a much larger sample (2383 matches in the final sample). Secondly, we investigate the impact of language training on income separately from its impact on employment probability. Lauringson et al. (2011) evaluate the impact on income, but do not distinguish between the effect that comes from increased employment and the effect that results from accessing higher-paying positions. Lastly, our study evaluates the impact of language training for different types of courses, initial language levels and regions allowing us to draw policy indications for increasing the effectiveness of language training provision.

### 3. Background

Over half (52%) of UIF Estonian language course participants were born in Estonia (Kallaste et al., 2018, p. 24). Around 23% have moved to Estonia since 2005. Somewhat fewer course participants (20%) moved to Estonia before 1990 and only 5% during the period 1991–2004. Thus, around a quarter of the sample are newly-arrived immigrants while the other three quarters are either second-generation immigrants or moved to Estonia during the Soviet era.<sup>2</sup> The term non-natives is used throughout the paper to refer to all of the three groups together.

The largest ethnic minority group in Estonia is ethnic Russians. Before World War II, Estonian population was relatively homogeneous. According to the 1934 population census, 88% of the inhabitants of the Republic of Estonia were ethnic Estonians, while ethnic Russians constituted 8% and other ethnic groups 4% of the population. During the Soviet occupation, ethnic Russians were incentivized to move to Estonia. By the year 1989, the share of ethnic Estonians had dropped to 62%, while ethnic Russians constituted 30% and other groups 8% of the inhabitants (Statistical Office of Estonia, 1995, p. 56). In 2017, the shares were 69%, 25% and 6% accordingly (Statistics Estonia, 2019a, Table RV0222).

During the Soviet time, two parallel school systems were established – one with Estonian as language of instruction, another with Russian. Although there have been considerable discussions since the restoration of independence in 1991 to uniform the systems, two parallel tracks still remain. The non-Estonian schools are required to teach the Estonian language as a separate subject at the level of basic education (1–3 grade 6 lessons per week, 4–9 grade 12 lessons per week). After a reform in 2011 60% of the lessons at



the level of secondary education (10–12 grade) must be taught in the Estonian language. While the younger generation of ethnic Russians are slowly increasing their Estonian language skills, the older generation completed their education with limited number and quality of state language lessons. Consequently, the lack of state language skills leaves many (second-generation) migrants at a disadvantage in the labour market.

The Estonian Unemployment Insurance Fund (UIF) offers language training to all those unemployed who are not proficient in the Estonian language. Participation in training is voluntary and agreed upon between the UIF consultant and the unemployed. There are two channels to take part in language training and both are free for unemployed participants. First, there is a possibility to participate in UIF-procured group training courses. These are long courses in which the whole level of language skills is targeted (e.g. level A2 or B1). Second, there is a possibility to take any language course in the open market and the UIF finances it from the training fund earmarked for the unemployed (the so-called training card). The training card gives each unemployed person a fund of 2500 euros to be spent on any type of training courses (including language training) over the period of two years.

The choice between UIF-procured language courses and open market language courses depends on a number of factors. Firstly, procured courses are offered for the language levels of A2, B1 and B2. The unemployed, who wish to participate in a language course for higher or lower level, are obliged to take a course from the open market. Secondly, procured courses target the whole level of language skills and prepare for the Estonian language proficiency examination. Therefore, if specific language training (e.g. legal language) or specific skill development (e.g. communication language course) is needed, one has to choose a language course from the variety of courses offered in the open market. Lastly, the cost of an open market language course is deducted from the training fund of 2500 euros, which is allocated for all of the training courses. If a person wishes to participate in other open market courses besides language training, it might be optimal to choose a procured language course to retain the training fund resources.

It appears that the language courses taken with the training card are essentially shorter than those provided in groups procured by the UIF. The average planned length of a procured training course is 279 academic hours, while the training card courses are more than hundred hours shorter (on average 121 academic hours). Additionally, the difference between the two strands of courses is that the procured courses take place during working hours 3–4 days per week while the training card enables to choose more flexible courses that take place in the evenings and are more suited for combining learning with working. There is no obligation to interrupt studying if the unemployed person is hired, but in practice it is difficult to combine procured courses with working as the job would have to be extremely flexible.

## **4. Methodology**

### **4.1. Data**

The data used for the analysis are the rich individual-level data drawn from the registry of the Estonian Unemployment Insurance Fund (UIF). The sample included all persons who, during the period 01.01.2015–31.12.2016, were at least once registered as unemployed

and participated in an Estonian language course provided by the UIF (the treatment group) and who did not participate in a course, but whose main language of communication was not Estonian (the control group).

In order to derive the outcome measures, the data from the UIF was linked to the data from the Estonian Tax and Custom Board's (ETCB) Register of Taxable Persons, which provided declared labour incomes for all the persons in the sample over the period 01.01.2013–01.08.2017 on a monthly basis.

The data from the UIF included 3224 persons who participated in a language course and 62,933 persons in the control group. A person could have had several unemployment periods during the query period. As active labour market programmes are related to specific unemployment periods, all the periods that were accompanied by a language course were addressed separately in a process of impact evaluation. Therefore, if a person had two unemployment periods and during both of these periods participated in a language course, then in the process of impact evaluation both of these periods were taken into account separately. There is a total of 3629 unique unemployment periods that were accompanied by a language course. If a person took several courses during one unemployment period, the first course would be taken into account. If a person was unemployed before the query period but her language course lasted longer and fit into the query period, and she then entered unemployment again later on during the query period but at that time did not participate in a language course (and by that failed the requirements of the treatment group), she was left out of the sample.

Of all the courses that were started, 21% were not completed. As the goal of the analysis is to assess the impact of the course as a whole (and not, for example, the impact of the motivation underlying the participation), which presumes completing the whole course, then all the courses that were not completed were left out of the analysis.<sup>3</sup> The individuals who started working during the participation of the language course, but still completed the course, were included in the treatment group.

As the outcomes for the analysis are derived from the income and social tax declarations that cover only income tax payments related to employment, and yet the unemployment status can also be exited for the purpose of starting a business, which may not lead to immediate taxable income, all the persons who received a business start-up subsidy were also excluded from the sample. There were 11 such individuals in the treatment group and 155 in the control group.

The final sample size for the treatment group was 2560 unemployment periods (2531 unique individuals) and for the control group 60,289 unemployment periods (43,391 unique individuals).

## **4.2. Matching**

In order to evaluate the direct effect of Estonian language courses on labour market outcomes, the counterfactuals (the control group) were established for the treatment group via a matching process that combined propensity score matching with coarsened exact matching (Rubin & Thomas, 2000). This is similar to the randomized block design in experimental study, where some key variables are used to divide the population into subgroups (blocks), after which the treatment conditions are randomly assigned to each block. In the case of matching, propensity scores were found within those subgroups. Logistic

regression was used for estimating the propensity scores<sup>4</sup> and 1:1 nearest neighbour algorithm with a specified calliper was used as the distance measure. Rosenbaum and Rubin (1985) recommended that the maximum size of the calliper should be at least 0.25 standard deviations of the propensity score. Lunt (2014) found, based on a simulation study, that although a tighter calliper may lead to a smaller matched sample size, as it is more difficult to find matches for all the treated, it also significantly reduces the selection bias. As the dataset in the current study was sufficiently large, the calliper size of 0.1 standard deviations of the propensity score was used.

The crucial aspect in using matching as an identification strategy for causal inference is the assumption of the ignorable treatment assignment, that is fulfilled only in a case when all covariates that influence treatment assignment are controlled for. This means that all those covariates must be observed and used in a matching process. We are able to control for a broad range of background variables. The variables used in the matching process are presented in Table 1. We acknowledge the risk of selection bias resulting from the inability to control for unobservable characteristics such as search intensity, motivation and learning ability. However, as search intensity is declining continuously within the unemployment spell (Faberman & Kudlyak, 2019), we argue that by matching on the months since unemployment we are able to account for at least part of the effect of the search intensity. Similarly, we try to capture the effect of learning ability by including the levels of education.

**Table 1.** The definition of variables included in the matching process.

Beginning of the course	Enrolment on courses was recorded on a monthly basis. The observation period was 01.01.2015 – 31.12.2016, but as some individuals who participated in a course during this period had enrolled on the course before that period, some starting dates were outside of that period. The earliest starting date was 07.2014.
Months of unemployment prior to the course	Months prior to the course were discretized to periods of 0–3 months, 4–6 months, 7–9 months, 10–12 months, 13–18 months, 19–24 months and 25 or more months.
Male	Dummy variable
Education	Basic education, secondary education, vocational education and higher education.
Region	Most of the migrants in Estonia are concentrated in two regions – Harju County and Ida-Viru County. As the migrant population outside of those regions is quite sparse, all the other regions were combined. Therefore, the variable distinguishes Harju County, Ida-Viru County and other regions.
Other labour market services	A dummy variable indicating whether a person received some kind of a labour market service not directly related to entering the labour market (career counselling, career information cabinets, job search workshop, psychological counselling, debt counselling, addiction counselling, work trial, community work, work practice, job club, voluntary work, work placement).
Estonian language proficiency	At the time of registering as an unemployed, the UIF consultant evaluates together with the unemployed person the Estonian language proficiency of the registrant. Possible levels are discretized into three groups: none, basic, intermediate or advanced.
Mean income prior to the course	Mean monthly income in euros during the 12 months before the language course. Only the months during which the person received income are taken into account. Income is discretized into categories: 0 euros, 0–399 euros, 400–700 euros and 800 or more euros.
Age	Continuous variable
ISCO classification	Job prior to the unemployment according to ISCO-08 major groups classification: managers; professionals; technicians and associate professionals; clerical support workers; service and sales workers; skilled agricultural, forestry and fishery workers; craft and related trades workers; plant and machine operators, and assemblers; elementary occupations and no prior work experience
Labour market training prior to the language course	A dummy variable indicating whether a person received some kind of a labour market training prior to participating in the language course.

The variables used for the exact matching were *course starting date (with a month's accuracy)*, *the length of the unemployment period before the course*, *gender* and *Estonian language proficiency*. The month of the beginning of the course was observed only for the treatment group. For the control group, this variable covered all the months that fell into their unemployment period. Therefore, the potential matches for a treatment group member whose language course started on 10.07.2015 would be all the members of the control group whose unemployment period overlapped with 07.2015 and whose prior employment period was identical to that of the specified member of the treatment group. As every individual in the control group could represent as many potential matches as how many months she was unemployed, all her other time-varying variables (age, length of prior employment period) were recalculated accordingly for each of her unemployed months. Therefore, although there were 60,289 unemployment periods (43,391 unique individuals) within the unmatched control group sample, in the matching process they accounted for 335,459 separate cases (the number of persons in the control group times the number of unemployment months for that person).

As a result of the matching, the balanced sample was established, combining only those members of the treatment group who had a sufficiently precise match from the control group and those members of the control group that were those matches. Thus, all the observations for which sufficiently close match were not found, remain outside of the scope of further analysis. The total of 2383 matches was found, which means that 93% of the original treatment sample is included in the matched sample.

The comparison of unmatched samples from the treatment group and the control group reveals significant differences across multiple attributes (Table 2). It must be noted that these distributions do not reflect the proportions of the population but the samples created for matching. The individuals in the control group are counted several times, separately for each month. One of the biggest differences between the treatment and control group is gender distribution. The treatment group contains only 20% of men, while in the control group the proportion of men is 50%. Thus, although the proportion of the unemployed among men is equal to that of women, or is somewhat higher, more women participate in Estonian language training. The treatment and control group are also observed to have different distributions of people who did not receive a salary during the 12 months preceding the unemployment period (difference of approx. 4%), who have higher education (difference of proportions ca 20%), who are living elsewhere than Harju or Ida-Viru County (about 8% difference), who were blue-collar workers prior to unemployment, and who are proficient in Estonian (difference of approx. 18%). Considering the differences in the levels of language proficiency, it is remarkable that compared to the control group, there are significantly fewer people in the treatment group with no Estonian language skills at all.

The balance within matched pairs and between matched and unmatched pairs was assessed using the standardized mean differences (SMD) (Rosenbaum, 2010). Small values of SMD (<0.1) support the assumption of balance between groups (Cohen, 1988). The values of SMD for most of the covariates in the matched sample were less than 0.1 (Table 2). Additionally, the balance of covariates and their interactions in relation to the treatment status was examined on the matched sample via a logistic regression model, and the propensity score distribution over the values of covariates between the treatment and control group was evaluated graphically. As the additional tests also indicated a

**Table 2.** Proportions of treatment and control groups in total and matched samples over the covariates.

	Total			Matched		
	Control	Treated	SMD	Control	Treated	SMD
<i>Observations</i>	335,459	2560		2383	2383	
<i>Months of unemployment prior to the course</i>			0.265			<0.001
0–3	(54.3%)	(44.6%)		(45.3%)	(45.3%)	
4–6	(19.3%)	(26.8%)		(26.9%)	(26.9%)	
7–9	(10.7%)	(14.5%)		(14.4%)	(14.4%)	
10–12	(5.9%)	(6.6%)		(6.3%)	(6.3%)	
13–18	(4.6%)	(4.4%)		(4.1%)	(4.1%)	
19–24	(1.7%)	(1.5%)		(1.3%)	(1.3%)	
25–	(3.5%)	(1.7%)		(1.7%)	(1.7%)	
<i>Male</i>	(49.7%)	(20.4%)	0.644	(19.8%)	(19.8%)	<0.001
<i>Education</i>			0.540			0.041
Secondary education	(23.4%)	(17.1%)		(17.8%)	(18.1%)	
Higher education	(16.2%)	(36.8%)		(31.5%)	(33.0%)	
Vocational education	(44.7%)	(40.3%)		(43.8%)	(42.6%)	
Basic education	(15.6%)	(5.9%)		(6.9%)	(6.3%)	
<i>Region</i>			0.256			0.080
Harju County	(41.4%)	(43.3%)		(45.9%)	(44.3%)	
Ida-Viru County	(46.6%)	(36.7%)		(40.0%)	(38.7%)	
Other regions	(12.0%)	(20.0%)		(14.1%)	(17.0%)	
<i>Other labour market services</i>	(42.2%)	(55.7%)	0.273	(54.6%)	(54.4%)	0.003
<i>Estonian language proficiency</i>			0.533			<0.001
None	(29.9%)	(10.1%)		(8.5%)	(8.5%)	
Basic	(40.7%)	(59.2%)		(59.4%)	(59.4%)	
Intermediate or advanced	(29.4%)	(30.7%)		(32.1%)	(32.1%)	
<i>Mean income prior to the course</i>			0.115			0.134
0	(32.7%)	(36.8%)		(36.5%)	(33.9%)	
0–400	(26.2%)	(21.8%)		(24.2%)	(22.8%)	
400–799	(25.6%)	(25.9%)		(27.7%)	(27.2%)	
800–	(15.6%)	(15.6%)		(11.6%)	(16.2%)	
<i>Age (mean(sd))</i>	(12.61%)	(11.61%)	0.053	(12.66%)	(11.65%)	0.040
<i>ISCO classification</i>			0.576			0.165
No prior working experience	(5.7%)	(5.4%)		(6.1%)	(5.4%)	
Managers	(3.6%)	(7.2%)		(5.4%)	(6.6%)	
Clerical support workers	(5.0%)	(8.6%)		(7.7%)	(8.5%)	
Elementary occupations	(22.3%)	(14.1%)		(16.7%)	(15.1%)	
Craft and related trades workers	(21.7%)	(8.3%)		(6.4%)	(8.9%)	
Skilled agricultural, forestry and fishery workers	(0.7%)	(0.4%)		(0.6%)	(0.4%)	
Plant and machine operators and assemblers	(13.1%)	(10.7%)		(9.2%)	(11.3%)	
Service and sales workers	(15.7%)	(20.7%)		(25.4%)	(21.4%)	
Technicians and associate professionals	(7.4%)	(13.2%)		(12.7%)	(12.8%)	
Professionals	(4.8%)	(11.6%)		(9.8%)	(9.6%)	
<i>Labour market training prior to the language course</i>	(14.1%)	(11.8%)	0.069	(10.8%)	(12.2%)	0.043

balance between the matched samples, it can be assumed that the compositional differences between the original samples were smoothed out by the matching and the samples were comparable.

### 4.3. Outcome variables

Outcomes, for which the impact of Estonian language training is assessed, are the probability of entering employment, employment sustainability and the potential income related to the employment. Estimates for both employment and labour income and their sustainability are based on the salary declared to the Estonian Tax and Customs Board (ETCB) following the entry to the programme.

The impact of treatment is assessed from the time of entry

- to the time of the first pay (the effect of the treatment on entry into employment);
- to the likelihood of salary being paid within 24 months after the treatment (sustainability of employment);
- to the amount of remuneration during the 24 months following the treatment (the economic benefits of the treatment).

The outcome variables are therefore the salary information for 24 months following the start of the language course. Since the assessment period starts from the moment the training begins, the analysis also takes into account the possibility that if the person does not receive the training (a member of the control group), she could be more likely to exit unemployment earlier. In other words, this logic of analysis allows to take into account the so-called lock-in effect, which means that while participating in the training the job searching activity (and in some cases the possibility of entering to employment) is smaller.

Although the maximum assessment period is 24 months, the effective assessment period is much shorter for a large portion of the sample. The reference period, i.e. the period in which the training had to take place, was 01.01.2015–31.12.2016. However, the impact of the training could only be assessed until 31.07.2017 (there is outcome data from the ETCB up to this point only). Therefore, for those who started the course in December 2016 the effective assessment period was only 7 months long and only the individuals whose training started before 01.08.2015 had a maximum assessment period of 24 months. As a result, the uncertainty contained in longer-term assessments is greater and the impact assessment of the treatment is more difficult.

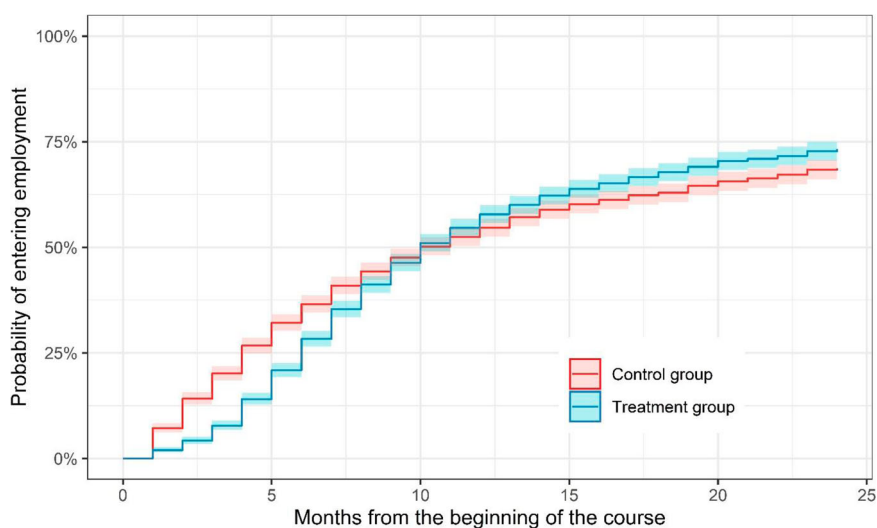
The ATT approach is used to assess the sustainability of the impact of Estonian language training. In the months following the start of the training, the impact of training on the likelihood of being employed is assessed. A separate logit model is used for each of the assessment period months for evaluating the difference between the treatment and control group. The impact of Estonian language training on labour income and the sustainability of its impact is again assessed via ATT. The linear regression model is used to estimate the impact of training on the labour income in subsequent months.

## 5. Results

In this section, the differences between the treatment and control group and the average treatment effect on the treated will be assessed. At first, entry into employment and the probability of staying in employment over the two-year observation period will be investigated. Employment probability will also be investigated by the course type, the level of language skills prior to the course and the residence region. Lastly, the attention is turned to the impact of language training on the income from wages.

### *5.1. The impact of language training on employment probability*

To evaluate the probability of entering employment, the Kaplan-Meier survival curves for the treatment and control group are calculated (see [Figure 1](#)). The probability of entering employment appears to be lower for those participating in training compared to those not participating for up to 10 months from the start of the language course. 10 months after



**Figure 1.** The Kaplan-Meier survival curves. Note. The log-log confidence intervals have been found at 95% confidence level.

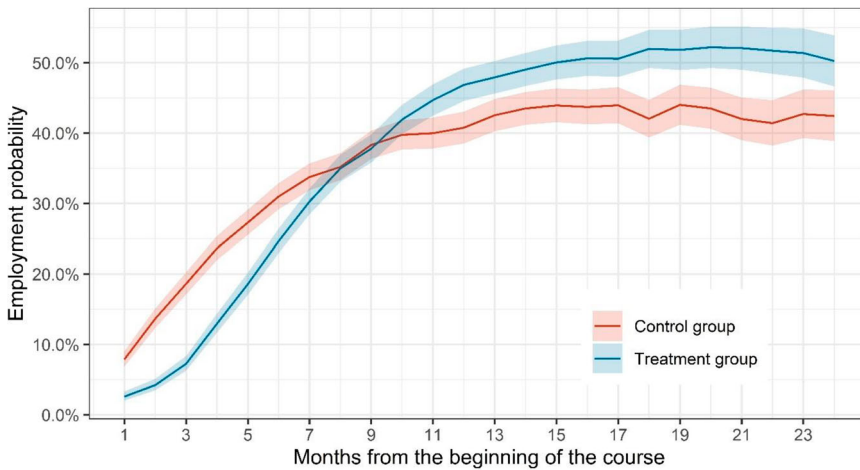
the start of the language course, around half of the treatment and control group have entered employment. By the end of the observation period, 29–34% of the control group and 25–29% of the treatment group have not entered employment.

The Kaplan-Meier survival curves are useful for evaluating the probability of the first entry to employment, but do not contain information on the sustainability of employment, i.e. whether and how long the treated stay in employment after the first entry. Therefore, the probability of being in employment over the observation period is estimated for the treatment and control group (Figure 2). Similarly, to the probability of entering employment, the probability of being employed is higher for the control group during the first few months of the course by around 10 pp. The difference is due to the lock-in effect. After 11 months, around 5 pp. more of those who participated in Estonian language training were in employment. The employment probability remains higher for the treatment group until the end of the two-year observation period.

The average treatment effect on the treated (ATT), estimated with a logit model based on the matched sample, is presented in Figure 3 (see also Table A1). Participation in Estonian language training significantly reduces the probability of being in employment for the first few months. During the first 3 months, the negative effect increases. After 3 months, the treatment effect reaches its lowest value as those participating in training are 9–13 pp. less likely to be employed than non-participants. From 8 to 10 months, the treatment effect is insignificant, but it turns positive from the 11th month after the start of the training. One year after the start of language training, the effect is around 6 pp. Two years after the start of the language training, the treatment effect is around 8 pp.

Although we find that there is a considerable lock-in effect, the long-term effect of language courses is shown to be positive and significant. Overall, our results on the probability of employment are in line with earlier findings on the impact of labour market training in Estonia. The effect of training courses was estimated to be 5.9 pp. after 6 months from the start of the training and 6.3 pp. after 12 months from the start of the training

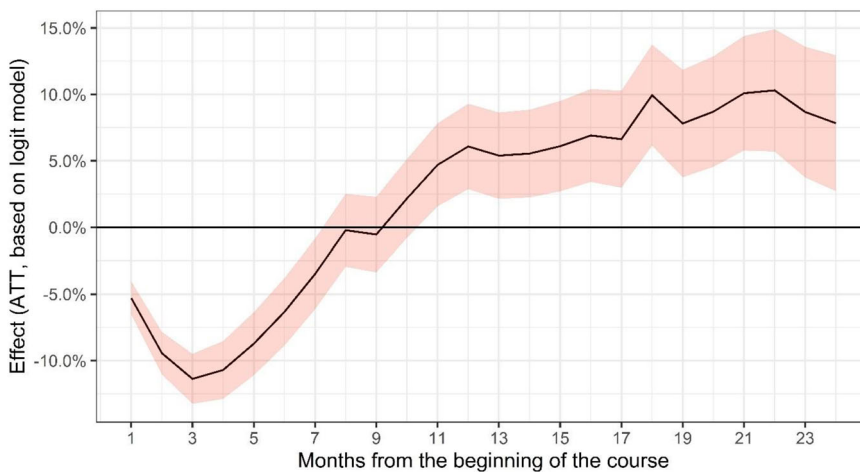




**Figure 2.** Employment probability of treatment and control group (matched sample). Note. The confidence intervals have been found at 95% confidence level.

by Anspal et al. (2012, p. 187). Leetmaa et al. (2003) found the effect of labour market training to be around 6–7 pp. 4–5 months after the end of the training. Lauringson et al. (2011) evaluated the effect of Estonian language courses ended in 2010 to be 13–18 pp. 4–12 months after the end of the training. While this last estimate is somewhat greater than other point estimates, the wide confidence intervals resulting from a small sample size still make the results comparable to earlier estimates.

Our results are also in line with the international evidence on the impact of state language training on employment (Clausen et al., 2009; Delander et al., 2005; Prey, 2000) and with findings on the effect of language proficiency on employment (e.g. Aldashev et al., 2009; Dustmann & Fabbri, 2003).



**Figure 3.** Treatment effect (ATT) on employment probability. Note. The confidence intervals have been found at 95% confidence level.

### 5.2. The impact of language training by the course type

The findings above indicate that language courses are an effective measure for helping unemployed non-natives find employment. However, while the long-term effect is positive, there is also a considerable lock-in effect at the beginning of the observation period. We are interested to see whether shorter and more flexible courses have smaller lock-in effect and consequently higher effectiveness of language training. Thus, the effect of language courses on employment is analysed by two different types of language courses.

As explained above, the procured courses ordered by the UIF are on average twice as long as the training card courses. In addition, the procured courses take place in the daytime and must be 3–6 academic hours long, making it arguably more difficult for participants to combine course participation with employment. Therefore, it would be expected that the lock-in effect is greater and longer-lasting for the procured courses due to inflexibility and longer hours. At the same time, the procured courses might have a greater long-term effect as the overall number of training hours is greater and, presumably, the increase in language knowledge is also greater.

The average treatment effect on the treated, estimated separately for the procured and training card courses, is displayed on [Figure 4](#) (see also [Table A2](#) and [Table A3](#)). The lock-in effect of procured courses is indeed of greater negative value during the first few months of the course. However, no significant long-term employment premium can be seen for the participants of procured courses. The last result supports the findings of [Kluve et al. \(2012\)](#), where additional training hours were not found to yield additional treatment effects after some treatment period.

### 5.3. The impact of language training by the level of language proficiency

The findings so far show that the language course participants enter and stay in employment with a higher probability compared to the control group after the initial lock-in



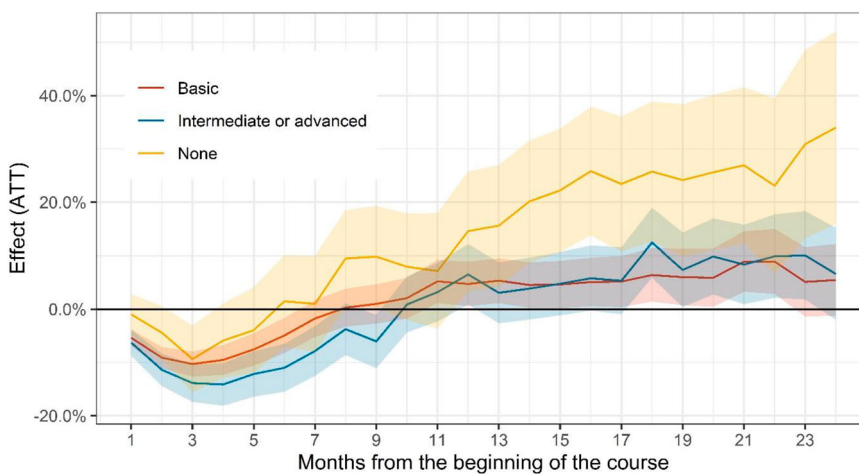
**Figure 4.** Treatment effect (ATT) on employment probability by the course type. Note. The confidence intervals have been found based on a logit model at 95% confidence level.

period. We wish to investigate, whether the impact of language training varies by the initial level of language proficiency measured at the beginning of the unemployment spell.

The average treatment effect on the treated is estimated separately for different levels of initial language proficiency. The results are shown in Figure 5.<sup>5</sup> For those with no prior level of language skills, the lock-in effect is not significant and the positive effect on employment probability increases to over 30 pp. by the end of the two-year observation period. The participants with basic skills of state language also experience positive impact of around 5 pp. on employment probability in the long-term. The largest lock-in effect is estimated for the language course participants with the highest prior level of language proficiency. Although the effect of language training turns positive from the 10<sup>th</sup> month after the start of the training, it stays only marginally significant until the end of the observation period. Overall, the results for different levels of language proficiency indicate that language training is most effective for the unemployed with the lowest level of language skills.

#### 5.4. The impact of language training by the region

As the last aspect of language training effectiveness on employment probability the effects by regions are investigated. One would expect heterogenous effects of language training by the regions for two main reasons. Firstly, the impact of language training may depend on the overall situation of the labour market in the regions. In the regions where the number of vacancies is higher, the alternative cost of staying in training (the lock-in effect) is expected to be greater. Thus, labour market training can have larger lock-in effects in capital regions or other regions with larger labour markets. The second explanation is related to the language environment of the regions. On one hand, the unemployed non-natives living in language enclaves might find employment possibilities with companies where the working language is not the state language. On the other hand, the knowledge of state language gives them comparative advantage over the other residents of the region.



**Figure 5.** Treatment effect (ATT) on employment probability by the initial level of language proficiency. Note. The confidence intervals have been found based on a logit model at 95% confidence level.

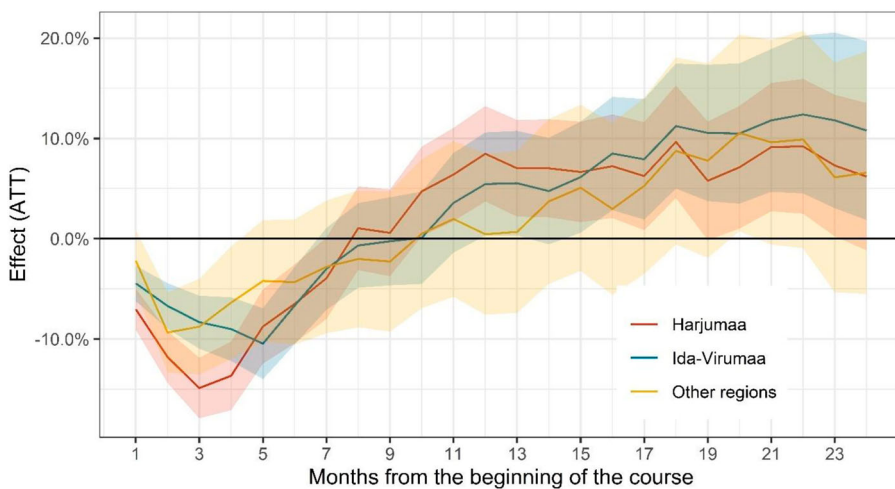
The estimates of the average treatment effect on the treated for different regions of residence of the participants are displayed in [Figure 6](#).<sup>6</sup> Ida-Virumaa, predominantly the Russian-speaking region, has smaller lock-in effect for the first 4 months than capital region Harjumaa. As explained above this might result from the differences in the number of vacancies in two regions. Participants from other regions also exhibit smaller lock-in effect than those in the capital region. In the long term, we see no significant difference in results for Ida-Virumaa and capital region Harjumaa. The long-term results of other regions are not statistically significant, most likely due to the small sample size of participants from those regions. Overall, despite the differences in the lock-on effect, the state language training seems to be beneficial in the two main regions it is offered.

### 5.5. The impact of language training on labour income

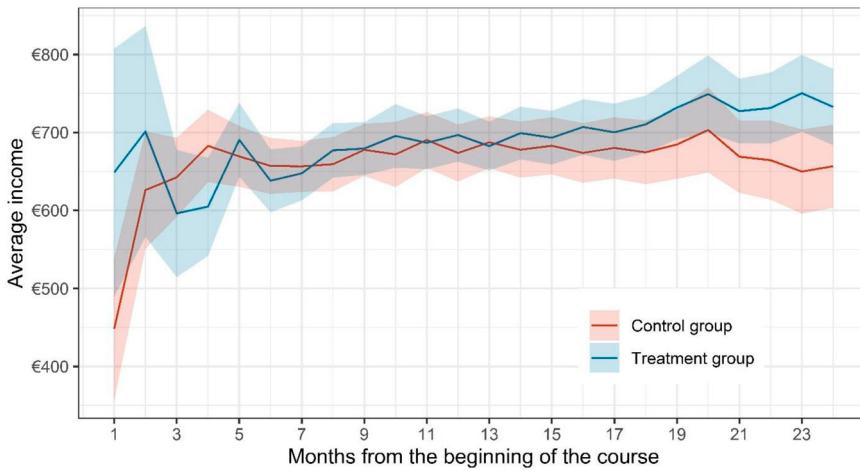
The results so far confirm that the treated have a higher probability to enter and stay in employment compared to the non-treated after the initial lock-in period. In addition to the impact on employment, the effect of Estonian language training on the labour income of the treated is investigated. For every month we only include those treated and non-treated who received a labour income. The results shown on [Figure 7](#) imply that the average labour income of the treatment group is not significantly different from the average labour income of the control group.

The effect of treatment is estimated by using the linear regression model on the matched sample. A separate regression model is evaluated for each of the 24 months. As above, only observations receiving a labour income for a particular month were included in the analysis. The ATT estimates indicate that language training does not have a significant impact on the labour income of the treated (see [Figure 8](#) and [Table A4](#)). For most months no significant effect is seen.

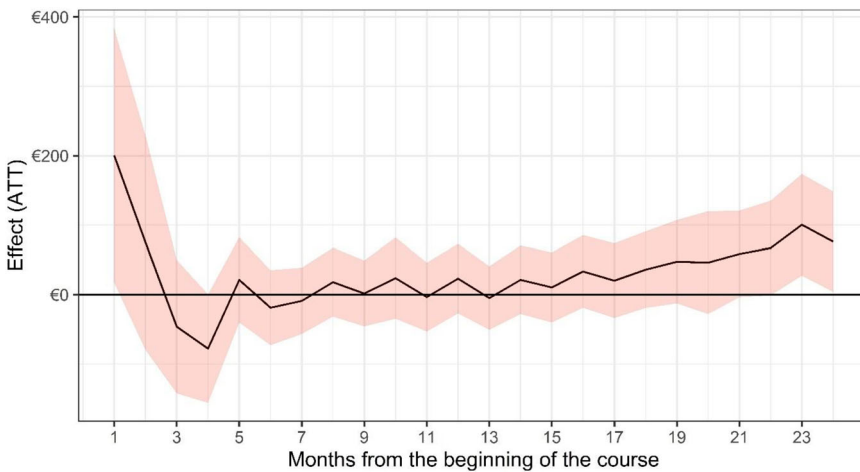
We found that language training had no significant effect on labour income. This result might seem puzzling at first. Based on the vast international evidence on the impact of



**Figure 6.** Treatment effect (ATT) on employment probability by the region. Note. The confidence intervals have been found based on a logit model at 95% confidence level.



**Figure 7.** Average labour income of treatment and control group (matched sample). Note. The confidence intervals have been found at 95% confidence level.



**Figure 8.** Treatment effect (ATT) on labour income. Note. The confidence intervals have been found at 95% confidence level.

language proficiency on earnings (e.g. Beyer, 2016; Bleakley & Chin, 2004), one would predict that language training that increases language proficiency would consequently increase earnings. The results of Toomet (2011) show that for ethnic Russians in Estonia, the language proficiency has an effect only on specific groups. However, those groups include low-wage earners, which leads to believe that language training that is aimed at unemployed would lead to positive effects on earnings.

However, when looking on the studies investigating labour market training effects based on Estonian data (Anspal et al., 2012; Leetmaa et al., 2003), we see our results in line with their findings. Furthermore, the international evidence on language training has found both insignificant and positive significant results on the impact on earnings

(Hayfron, 2001; Sarvimäki & Hämäläinen, 2016). Thus, we see that although language training does help non-natives to find positions, it does not lead them to access higher-paying positions.

## 6. Conclusion

By using propensity score matching combined with coarsened exact matching, the study evaluates the impact of state language training on labour market outcomes of the unemployed in Estonia. This study finds that state language training has a positive significant effect on the probability of being employed starting from the 11th month after the start of the course. The effect stays significant and is around 8 pp. at the end of the two-year observation period. No statistically significant effect is found on wages.

Our study shows a considerable lock-in effect of participants during the first few months of the course. Thus, despite long-term effectiveness it is worth to consider, whether language training should be targeted to participants with the lowest lock-in effect due to their individual characteristics, used more extensively when and where the external environment would predict more beneficial results or is it possible to alter the language course content and organization in some way. To provide useful information in this aspect, the study estimated the effects of training for different types of language courses, for participants with different level of prior language skills and for those living in regions with different language environment and labour market characteristics.

The lock-in effect for language training differs by course type. The more flexible and shorter training card courses have a smaller lock-in effect than the procured courses ordered by the Unemployment Insurance Fund. While there is no difference between the longer-term employment effect of two strands of language training courses and the participants of procured courses exhibit larger lock-in effect and also considerably larger dropout rates (25% vs. 13%), it is worth considering how to introduce additional flexibility into procured courses. One option is to divide the courses that aim to achieve the whole level of language skills into smaller standardized sublevels, which could be flexibly combined. While considering these changes it is important to note that an option to achieve the whole level of language skills should still be possible for those who wish to apply for a position, for which they are required to pass a state language proficiency examination (e.g. teachers, service sector workers).

The results for different levels of language proficiency indicate that language training is most effective for the unemployed with the lowest level of language skills. Although, the language training should be available for all levels of language skills (especially considering the official requirements of high level of state language skills for some positions), it is worth to consider, how to target language training to include more of those unemployed with no prior state language skills.

The results show some differences in the lock-in effect between the capital region and the other regions included. State language training appears to be slightly less beneficial in the capital region, which is characterized by higher number of vacancies and consequently larger lock-in effect of participating in the training. However, the positive long-term effects support the provision of language training in all of the regions.

In conclusion, the results of this study support the use of language training as an effective measure for helping the unemployed non-natives to find employment and

increase labour market integration. At the same time, it must be kept in mind that language training might not help non-natives find positions with better quality, i.e. with higher wages. Whether the insignificant impact on wages results from the insufficient increase in the level of language skills or from other reasons remains a question for future research.

## Notes

1. The estimation forms a part of a wider assessment of Estonian language training provision and needs in Estonia, commissioned by the Ministry of Social Affairs and the Ministry of Culture, and financed with the aid from the European Regional Development Fund (program “Strengthening of sectoral R&D (RITA)” under activity 2 “Support for knowledge-based policy formulation” [project number RITA2/030]).
2. Among the control group the share of those born in Estonia is around 61%. Around 29% moved to Estonia since 2005, 7% before 1990 and only 2% during the period 1991–2004 (Kallaste et al., 2018, p. 24).
3. The exclusion of the participants, who did not finish the course, does not alter the main conclusions of this study. The results including all of the course participants are available from the authors on request.
4. As the coarsened exact matching was performed prior to the propensity score estimation, the total of 30x7 logistic regression models were computed. Therefore, the regression models are not presented in the paper and are available from the authors on request.
5. The tables with the results are available from the authors on request.
6. The tables with the results are available from the authors on request.

## Acknowledgements

The authors are grateful for the three anonymous reviewers for their helpful comments and advice. We would like to thank Tiiu Paas for her comments and suggestions on earlier versions of this paper, Lenno Uusküla and Raul Eamets for their advice and comments during a presentation of this article. We are grateful for the thoughtful comments of Maryna Tverdostup and an anonymous pre-reviewer of the PhD summer school of Doctoral School in Economics and Innovation in Laulasmaa 2019. The authors take sole responsibility for all errors and omissions.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Notes on contributors

**Laura Helena Kivi** works as a researcher at the Estonian Center for Applied Research (CentAR). She is working towards a PhD degree in Economics at the University of Tartu. Her research areas cover the issues of labour market integration of migrants and regional labour markets.

**Marko Sõmer** works as a researcher at the Estonian Center for Applied Research (CentAR). His main research areas are labour market, social protection and education.

**Epp Kallaste** is a CEO and a senior researcher at the Estonian Center for Applied Research (CentAR). She received her PhD from the University of Tartu in 2010. Her main areas of expertise are labour relations, labour market and social protection.



## References

- Aldashev, A., Gernandt, J., & Thomsen, S. L. (2009). Language usage, participation, employment and earnings: Evidence for foreigners in West Germany with multiple sources of selection. *Labour Economics*, 16(3), 330–341. <https://doi.org/10.1016/j.labeco.2008.11.004>
- Anspal, S., Järve, J., Jääts, K., Kallaste, E., Maar, K., Mägi, A., & Toots, A. (2012). Programmi “Kvalifitseeritud tööjõu pakkumise suurendamine 2007(2013)” teenuste hindamine. *Palgatoetuse ja tööturukoolituse teenuste vahehindamine*. Estonian Centre for Applied Research (CentAR) and Interact Projektid & Koolitus, 199 p. [https://centar.ee/wp-content/uploads/2012/01/2012\\_06\\_07-%C3%B6%C3%B6turumeetmete-hindamise-l%C3%B5ppraport-CENTAR-InterAct-final.pdf](https://centar.ee/wp-content/uploads/2012/01/2012_06_07-%C3%B6%C3%B6turumeetmete-hindamise-l%C3%B5ppraport-CENTAR-InterAct-final.pdf).
- Becker, G. S. (1975). *Human capital: A theoretical and empirical analysis, with special reference to education* (2nd ed). Columbia University Press for NBER, 264 p.
- Beyer, R. C. M. (2016). *The labor market performance of immigrants in Germany* (IMF Working Papers, 16(6)), 38 p. <https://doi.org/10.5089/9781498376112.001>.
- Bleakley, H., & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and Statistics*, 86(2), 481–496. <https://doi.org/10.1162/003465304323031067>
- Budria, S., de Ibarreta, M., & Swedberg, C. (2017). The impact of host language proficiency across the immigrants’ earning distribution in Spain. *IZA Journal of Migration and Development*, 7(12), 1–27. <https://doi.org/10.1186/s40176-017-0094-2>
- Butschek, S., & Walter, T. (2014). What active labour market programmes work for immigrants in Europe? A meta-analysis of the evaluation literature. *IZA Journal of Migration*, 3(48), 1–18. <https://doi.org/10.1186/s40176-014-0023-6>
- Chiswick, B. R., Lee, Y. L., & Miller, P. W. (2005). Immigrant earnings: A longitudinal analysis. *Review of Income and Wealth*, 51(4), 485–503. <https://doi.org/10.1111/1467-9361.00039>
- Clausen, J., Heinesen, E., Hummelgaard, H., Husted, L., & Rosholm, M. (2009). The effect of integration policies on the time until regular employment of newly arrived immigrants: Evidence from Denmark. *Labour Economics*, 16(4), 409–417. <https://doi.org/10.1016/j.labeco.2008.12.006>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). L. Erlbaum Associates. 567 p.
- Cohen-Goldner, S., & Eckstein, Z. (2008). Labor mobility of immigrants: Training, experience, language and opportunities. *International Economic Review*, 49(3), 837–872. <https://doi.org/10.1111/j.1468-2354.2008.00499.x>
- Delander, L., Hammarstedt, M., Månsson, J., & Nyberg, E. (2005). Integration of immigrants. The role of language proficiency and experience. *Evaluation Review*, 29(1), 24–41. <https://doi.org/10.1177/0193841X04270230>
- Dustmann, C., & Fabbri, F. (2003). Language proficiency and labour market performance of immigrants in the UK. *The Economic Journal*, 113(489), 695–717. <https://doi.org/10.1111/1468-0297.t01-1-00151>
- Dustmann, C., & van Soest, A. (2002). Language and the earnings of immigrants. *Industrial and Labor Relations Review*, 55(3), 473–492. <https://doi.org/10.2307/2696052>
- Faberman, R. J., & Kudlyak, M. (2019). The intensity of job search and search duration. *American Economic Journal: Macroeconomics*, 11(3), 327–357. <https://doi.org/10.1257/mac.20170315>
- Fitzenberger, B., Osikominu, A., & Paul, M. (2010). *The heterogeneous effects of training incidence and duration on labor market transitions* (IZA Discussion Paper No. 5269). The Institute for the Study of Labor, 58 p.
- Fitzenberger, B., Osikominu, A., & Völter, R. (2008). Get training or wait? Long-run employment effects of training programs for the unemployed in West Germany. *Annales D’Économie Et De Statistique*, 91/92, 321–355. <https://doi.org/10.2307/27917250>
- Gerfin, M., & Lechner, M. (2002). A Microeconometric evaluation of the active labour market policy in Switzerland. *The Economic Journal*, 112(482), 854–893. <https://doi.org/10.1111/1468-0297.00072>
- Hayfron, J. (2001). Language training, language proficiency and earnings of immigrants in Norway. *Applied Economics*, 33(15), 1971–1979. <https://doi.org/10.1080/00036840010018630>
- Kallaste, E., Anspal, S., Kivi, L. H., Sömer, M., Lang, A., Sandre, S.-L., & Järve, J. (2018). *Eesti keelest erineva emakeelega täiskasvanute eesti keele õpe lõimumis- ja tööhõivepoliitikas: kvaliteet, mõju ja*

- korraldus. 4. Tööturuteenusena osutatav eesti keele õpe. Estonian Centre for Applied Research (CentAR), 118 p. <https://centar.ee/wp-content/uploads/2018/07/4.-osa-eesti-keele-%C3%B5pe-t%C3%B6%C3%B6turuteenusena.pdf>
- Kluge, J. (2010). The effectiveness of European active labor market programs. *Labour Economics*, 17(6), 904–918. <https://doi.org/10.1016/j.labeco.2010.02.004>
- Kluge, J., Schneider, H., Uhlenhorff, A., & Zhao, Z. (2012). Evaluating continuous training programmes by using the generalized propensity score. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 175(2), 587–617. <https://doi.org/10.1111/j.1467-985X.2011.01000.x>
- Lang, K. (1986). A Language Theory of Discrimination. *The Quarterly Journal of Economics*, 101(2), 363–382. <http://doi.org/10.2307/1891120>
- Lauringson, A., Villsaar, K., Tammik, L., & Luhavee, T. (2011). *Impact evaluation of labour market training*. Estonian Unemployment Insurance Fund, 60 p. [https://www.tootukassa.ee/sites/tootukassa.ee/files/Impact\\_Evaluation\\_of\\_Labour\\_Market\\_Training.pdf](https://www.tootukassa.ee/sites/tootukassa.ee/files/Impact_Evaluation_of_Labour_Market_Training.pdf)
- Lechner, M., & Wunsch, C. (2009). Are training programs more effective when unemployment is high? *Journal of Labor Economics*, 27(4), 653–692. <https://doi.org/10.1086/644976>
- Leetmaa, R., Võrk, A., Eamets, R., & Sõstra, K. (2003). *Aktiivse tööpoliitika tulemuslikkuse analüüs Eestis*. Tallinn: Praxis Centre for Policy Studies, 108 p. <http://www.praxis.ee/wp-content/uploads/2014/03/2003-Aktiivse-toopoliitikatulemuslikkus.pdf>
- Lunt, M. (2014). Selecting an appropriate caliper can be essential for achieving good balance with propensity score matching. *American Journal of Epidemiology*, 179(2), 226–235. <https://doi.org/10.1093/aje/kwt212>
- McGuinness, S., O'Connell, P. J., & Kelly, E. (2014). The impact of training programme type and duration on the employment chances of the unemployed in Ireland. *The Economic and Social Review*, 45(3), 425–450.
- OECD. (2018). *International migration outlook 2018*. [https://doi.org/10.1787/migr\\_outlook-2018-en](https://doi.org/10.1787/migr_outlook-2018-en)
- Orlov, G. (2017). *The impact of language training on the transfer of pre-immigration skills and the wages of immigrants*. Department of Economics, The University of Western Ontario, Job Market Paper, 45 p. [http://economics.uwo.ca/PhD\\_job\\_candidates\\_docs/Orlov\\_Job%20Market%20Paper.pdf](http://economics.uwo.ca/PhD_job_candidates_docs/Orlov_Job%20Market%20Paper.pdf)
- Osikominu, A. (2013). Quick job entry or long-term human capital development? The dynamic effects of alternative training schemes. *The Review of Economic Studies*, 80(1), 313–342. <https://doi.org/10.1093/restud/rds022>
- Prey, H. (2000). Evaluation of training programs in St. Gallen, Switzerland. *The Swiss Journal of Economics and Statistics*, 136(3), 417–432.
- Rinne, U. (2012). *The evaluation of immigration policies* (IZA Discussion Paper No. 6369). The Institute for the Study of Labor, 37 p.
- Rosenbaum, P. R. (2010). *Design of observational studies*. Springer series in statistics. Springer.
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1), 33–38. <https://doi.org/10.2307/2683903>
- Rubin, D. B., & Thomas, N. (2000). Combining propensity score matching with additional adjustments for prognostic covariates. *Journal of the American Statistical Association*, 95(450), 573–585. <https://doi.org/10.1080/01621459.2000.10474233>
- Sarvimäki, M., & Hämäläinen, K. (2016). Integrating immigrants: The impact of restructuring active labor market programs. *Journal of Labor Economics*, 34(2), 479–508. <https://doi.org/10.1086/683667>
- Statistical Office of Estonia. (1995). *Eesti rahvastik rahvaloenduste andmetel. I*. [Population of Estonia by population censuses]. Tallinn, 263 p.
- Statistics Estonia. (2019a). *Table RV0222: Rahvastik soo, rahvuse ja maakonna järgi, 1. Jaanuar*. Retrieved February 12, 2019 from [http://pub.stat.ee/px-web.2001/Dialog/varval.asp?ma=RV0222&ti=RAHVASTIK+SOO%2C+RAHVUSE+JA+MAAKONNA+J%C4RGI%2C+1%2E+JAANUAR&path=../Database/Rahvastik/01Rahvastikunaitajad\\_ja\\_kooseis/04Rahvaarv\\_ja\\_rahvastiku\\_kooseis/&lang=2](http://pub.stat.ee/px-web.2001/Dialog/varval.asp?ma=RV0222&ti=RAHVASTIK+SOO%2C+RAHVUSE+JA+MAAKONNA+J%C4RGI%2C+1%2E+JAANUAR&path=../Database/Rahvastik/01Rahvastikunaitajad_ja_kooseis/04Rahvaarv_ja_rahvastiku_kooseis/&lang=2)
- Statistics Estonia. (2019b). *Table TT136: 15-17-aastased majanduslikult aktiivsed (tööjõud) hõiveseisundi, haridustaseme ja isikute rühma järgi, 2*. Retrieved February 3, 2019 from <http://pub.stat.ee/px-web.2001/Dialog/varval.asp?ma=TT136&ti=15-17-aastased%20majanduslikult%20aktiivsed%20t%C3%B6%C3%B6j%C3%B6ud%20h%C3%B5iveseisundi%20haridustaseme%20ja%20isikute%20r%C3%BChma%20j%C3%A4rgi%2C%202>

pub.stat.ee/px-web.2001/dialog/varval.asp?ma=TT136&ti=15%2D74%2DAASTASED+MAJANDUS LIKULT+AKTIIVSED+%28T%D6%D6J%D5UD%29+H%D5IVESEISUNDI%2C+HARIDUSTASEME+JA+ ISIKUTE+R%DCHMA+J%C4RGI&path=../database/eri\_valdkondade\_statistika/loimumine/&search =TT136&lang=2

Stephan, G., & Pahnke, A. (2011). The relative effectiveness of selected active labor market programs: An empirical investigation for Germany. *The Manchester School*, 79(6), 1262–1293. <https://doi.org/10.1111/j.1467-9957.2010.02221.x>

Toomet, O. (2011). Learn English, not the local language! Ethnic Russians in the Baltic States. *American Economic Review: Papers and Proceedings*, 101(3), 526–531. <https://doi.org/10.1257/aer.101.3.526>

## Appendix

**Table A1.** Language course treatment effect on employment probability (logit model).

Month from the beginning of the training	Marginal effect	SE	Lower conf. int.	Upper conf. int.	p-value
1	-5.29%	0.006	-6.54%	-4.03%	0.000
2	-9.44%	0.008	-11.04%	-7.84%	0.000
3	-11.37%	0.010	-13.25%	-9.49%	0.000
4	-10.70%	0.011	-12.88%	-8.53%	0.000
5	-8.73%	0.012	-11.10%	-6.35%	0.000
6	-6.34%	0.013	-8.88%	-3.80%	0.000
7	-3.48%	0.014	-6.13%	-0.84%	0.010
8	-0.21%	0.014	-2.95%	2.52%	0.878
9	-0.53%	0.014	-3.37%	2.30%	0.713
10	2.17%	0.015	-0.79%	5.14%	0.150
11	4.70%	0.016	1.57%	7.83%	0.003
12	6.09%	0.016	2.87%	9.30%	0.000
13	5.39%	0.017	2.15%	8.64%	0.001
14	5.55%	0.017	2.25%	8.85%	0.001
15	6.11%	0.017	2.71%	9.50%	0.000
16	6.91%	0.018	3.42%	10.41%	0.000
17	6.63%	0.019	2.98%	10.28%	0.000
18	9.94%	0.019	6.13%	13.75%	0.000
19	7.81%	0.021	3.77%	11.86%	0.000
20	8.70%	0.021	4.55%	12.86%	0.000
21	10.09%	0.022	5.78%	14.40%	0.000
22	10.30%	0.023	5.70%	14.90%	0.000
23	8.67%	0.025	3.75%	13.60%	0.001
24	7.83%	0.026	2.72%	12.94%	0.003

**Table A2.** The treatment effect of an UIF-ordered language course on employment probability (logit model).

Month from the beginning of the training	Marginal effect	SE	Lower conf. int.	Upper conf. int.	p-value
1	-6.61%	0.007	-8.03%	-5.18%	0.000
2	-11.88%	0.009	-13.65%	-10.11%	0.000
3	-14.36%	0.011	-16.48%	-12.24%	0.000
4	-13.33%	0.013	-15.80%	-10.86%	0.000
5	-12.00%	0.014	-14.76%	-9.24%	0.000
6	-8.30%	0.015	-11.30%	-5.30%	0.000
7	-5.88%	0.016	-9.02%	-2.74%	0.000
8	-1.82%	0.016	-5.05%	1.41%	0.268
9	-1.58%	0.017	-4.93%	1.78%	0.357
10	1.28%	0.018	-2.23%	4.80%	0.474
11	4.56%	0.019	0.82%	8.31%	0.017
12	6.62%	0.020	2.77%	10.47%	0.001
13	5.34%	0.020	1.44%	9.25%	0.007

(Continued)

**Table A2.** Continued.

Month from the beginning of the training	Marginal effect	SE	Lower conf. int.	Upper conf. int.	p-value
14	5.26%	0.020	1.32%	9.20%	0.009
15	4.99%	0.021	0.92%	9.06%	0.016
16	5.89%	0.021	1.71%	10.07%	0.006
17	6.27%	0.022	1.88%	10.66%	0.005
18	8.77%	0.023	4.18%	13.35%	0.000
19	8.43%	0.025	3.50%	13.35%	0.001
20	9.84%	0.026	4.78%	14.89%	0.000
21	9.24%	0.027	3.96%	14.51%	0.001
22	9.88%	0.029	4.21%	15.56%	0.001
23	8.23%	0.031	2.06%	14.40%	0.009
24	7.41%	0.033	1.00%	13.82%	0.024

**Table A3.** The treatment effect of a training card language course on employment probability (logit model).

Month from the beginning of the training	Marginal effect	SE	Lower conf. int.	Upper conf. int.	p-value
1	-2.32%	0.013	-4.84%	0.20%	0.071
2	-3.96%	0.017	-7.25%	-0.66%	0.019
3	-4.64%	0.019	-8.41%	-0.87%	0.016
4	-4.77%	0.022	-9.08%	-0.47%	0.030
5	-1.36%	0.023	-5.90%	3.17%	0.556
6	-1.91%	0.024	-6.64%	2.82%	0.428
7	1.91%	0.025	-2.98%	6.80%	0.444
8	3.59%	0.026	-1.52%	8.69%	0.169
9	1.94%	0.027	-3.33%	7.21%	0.470
10	4.25%	0.028	-1.22%	9.71%	0.128
11	5.00%	0.029	-0.64%	10.64%	0.082
12	4.91%	0.030	-0.88%	10.71%	0.097
13	5.51%	0.030	-0.32%	11.34%	0.064
14	6.20%	0.031	0.21%	12.20%	0.043
15	8.59%	0.031	2.50%	14.69%	0.006
16	9.26%	0.032	2.93%	15.59%	0.004
17	7.43%	0.033	0.87%	13.99%	0.026
18	12.53%	0.035	5.72%	19.34%	0.000
19	6.54%	0.036	-0.53%	13.62%	0.070
20	6.37%	0.037	-0.91%	13.65%	0.086
21	11.80%	0.038	4.33%	19.27%	0.002
22	11.11%	0.040	3.25%	18.97%	0.006
23	9.44%	0.042	1.29%	17.59%	0.023
24	8.55%	0.043	0.13%	16.97%	0.047

**Table A4.** Language course treatment effect on labour income (linear regression model).

Month from the beginning of the training	Coef.	SE	p-value
1	200.4139	93.56925	0.033
2	75.20548	78.97118	0.341
3	-46.3158	48.99697	0.345
4	-77.8927	39.92849	0.051
5	21.21535	31.33217	0.498
6	-18.9378	27.55665	0.492
7	-9.10329	24.3435	0.708
8	17.87752	25.25048	0.479
9	1.581843	24.22295	0.948
10	23.70324	29.83483	0.427
11	-3.55838	25.28079	0.888
12	22.95786	25.52377	0.369
13	-4.99703	23.36619	0.831
14	21.24296	25.15457	0.399
15	10.32525	25.69916	0.688

(Continued)

**Table A4.** Continued.

Month from the beginning of the training	Coef.	SE	<i>p</i> -value
16	33.24338	26.67365	0.213
17	20.07186	27.43188	0.464
18	35.83313	28.08724	0.202
19	47.3544	30.62306	0.122
20	45.95474	37.70232	0.223
21	58.48029	31.80425	0.066
22	66.97887	34.74926	0.054
23	100.505	37.34997	0.007
24	76.08873	36.89266	0.04