



Make IT Work: The Labor Market Effects of Information Technology Retraining in the Netherlands

Melvin Vooren^{1,2}  · Carla Haelermans³ · Wim Groot³ · Henriette Maassen van den Brink²

Accepted: 4 June 2022 / Published online: 6 July 2022
© The Author(s) 2022

Abstract

The aim of this paper is to evaluate the effects of an active labor market program (ALMP) for higher educated workers in the Netherlands. The one-year program is characterized by six months of full-time IT retraining followed by a six-month internship. We estimate the effects of participating on earnings per month and working days per month. The results show significant lock-in effects during the program, lasting up to five months after program start. After this lock-in period, we find significant positive effects on earnings and working days. These positive effects remain significant until the end of the 36-month evaluation period. A conservative cost–benefit analysis based on the effects on earnings shows a return of 2.53% (95% CI 2.02–3.04%), which is low compared to the returns to education in the Netherlands of around 8%. We conclude that IT retraining has positive effects on the labor market outcomes of the participants yet relatively low returns.

Keywords Active labor market policy · Retraining · Information technology · Labor market out-comes · Internal rate of return

JEL Classification I26 · J68

✉ Melvin Vooren
M.Vooren@vu.nl

¹ Faculty of Behavioral and Movement Sciences, Amsterdam Center for Learning Analytics, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

² Amsterdam School of Economics and Top Institute for Evidence Based Education Research, University of Amsterdam, Amsterdam, Netherlands

³ School of Business and Economics, Maastricht University, Maastricht, Netherlands

1 Introduction

Active labor market programs (ALMPs) are policy instruments that are frequently used to increase re-employment rates among the unemployed. The programs often vary in duration, target group, and the exact treatment (see Vooren et al. (2019a) for an overview and meta-analysis of recent evaluation studies). Many have been evaluated in the past, focusing on their short-, medium-, and long-term results, and exposing short-term negative lock-in effects as an important challenge. This is especially true for training and retraining programs. These lock-in effects produce a negative impact on labor market outcomes during the period that the participants are in training and during the first months after finishing their training. For an ALMP to be effective, it is important that these negative lock-in effects are compensated by positive long-term effects, as described in Vooren et al. (2019a) and Card et al. (2017). By calculating the internal rate of return (IRR) of a training program, the lock-in effects are taken into account. The results presented in this paper point at positive long-term effects of IT retraining, and a positive IRR of 2.53%.

The main research question answered in this paper is whether IT retraining is an effective tool for improving the labor market outcomes of retrainees. We do this by estimating the labor market effects of an IT retraining ALMP in the Netherlands that matches participant skill sets with vacancies in the information technology (IT) sector. The program, called *Make IT Work*, consists of a combination of classroom training and on-the-job training. It was designed as a regional ALMP for the Amsterdam metropolitan area during the economic crisis of 2014 and begun in September 2015 with a government subsidy from the Dutch Ministry of Social Affairs and Employment. An interesting feature of this particular retraining program is that it is open for everyone who has at least a bachelor's degree, instead of just for unemployed individuals on welfare or unemployment benefits, who are the usual target group for ALMPs. In the Netherlands, young university and college graduates without prior employment history are not entitled to unemployment benefits. Therefore, they do not get exposed to the regular palette of ALMPs implemented through the public unemployment insurance system. In the aftermath of the 2007–2008 financial crisis, concern has been escalating over rising unemployment rates; and since the European sovereign debt crisis of 2010, growing youth unemployment rates have demanded particular attention. As a result, political interest in ALMPs has developed commensurately in the Netherlands and in the Eurozone as whole.

The literature on ALMPs mainly focuses on the effects of compulsory programs for workers with low levels of education. For instance, Kruppe and Lang (2018) evaluate the effects of retraining on low-skilled workers in Germany between 2004 and 2007. In the German program, the participants eventually acquired a formal vocational degree. Wapler et al. (2018) look at the effects of AMLP on regional German labor markets, taking into account general equilibrium effects. Wunsch and Lechner (2008) study the long-term effects of training programs in Germany. Cavaco et al. (2013) employ matching techniques to estimate the effects of ALMPs in France on the re-employment rates of displaced workers, finding positive effects in the medium run. The broad conclusion in recent literature is that training programs

are relatively ineffective in the long-term. This is also predicted by the theory presented in the *Handbook of Labor Economics* by Heckman et al. (1999) that ALMP is usually ineffective. Heckman et al. (1999) argue that ALMP usually fails to be effective, because it mainly targets, in Heckman's words, 'less able' individuals, for whom the rate of return to schooling is lower than for the average population.

However, the Dutch *Make IT Work* training program that is evaluated in this paper is distinct from other programs in two aspects. First, it is voluntary. Completing the program requires a contribution of time and effort by the participant. In addition, the Dutch program requires a down payment of EUR 1000 from each participant. The level of financial commitment may benefit from a comparison: the tuition fee for a full-time master's program at a Dutch university is around EUR 2000. Given the investment of time and money required for participation in *Make IT Work*, it is probably safe to assume that the program attracts motivated participants. Second, the target group of this particular training program is different from most programs previously evaluated. The participants are highly educated: a formal university bachelor's degree is prerequisite to enrollment. The combination of increased personal commitment and higher levels of education may suggest a reason why *Make IT Work* is found to be so much more successful than other ALMPs described in the meta-analyses in Card et al. (2017) and Vooren et al. (2019a).

Another reason that Heckman et al. (1999) give for the relative ineffectiveness of ALMPs is that they are usually administered by agencies in the public sector. They argue that for ALMPs to be associated with large gains, the rate of return associated with these public training programs would have to be higher than the rate of return to private sector training, which is almost never the case. In contrast to the evaluation literature, which focuses on public sector training programs, this paper assesses the labor market effects of a private sector IT retraining ALMP in the Netherlands that has been designed and implemented in cooperation with employers that are active in the IT labor market. In the Dutch IT sector, companies struggle to find qualified personnel (CBS, 2019). Given the labor market shortages for highly educated IT workers in the Netherlands, the results presented in this paper also offer transferable knowledge concerning the potential for training programs in other sectors with shortages of highly educated workers.

The treatment group in our sample consists of the 250 people who started the program between September 2015 and November 2018. In order to evaluate the labor market effects of the retraining program, we apply matching techniques to construct a control group of non-participants from the Statistics Netherlands register data. The results indicate a substantial lock-in period on the number of working days per month of up to six months. The lock-in impact on gross earnings per month are slightly longer lasting: up to nine months after the start of the program. After the lock-in period, there is a positive and sustainable effect on earnings and on the number of working days. Based on the results, we conclude that the program is successful in improving the labor market outcomes of the participants. The results are stable across different matching techniques. A conservative cost-benefit assessment points at an annual IRR of 2.53%, which is statistically significant at the 95% level. Compared to the average return on one year of schooling, the returns from *Make IT Work* are low (Psacharopoulos & Patrinos, 2018) and confirm the observation by

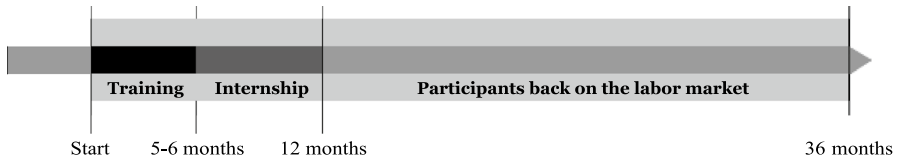


Fig. 1 Timeline in months after program start. *Notes:* The 36-month evaluation period of this paper covers the period from the month of the program start-up to three years thereafter

Heckman et al. (1999) that ALMP training programs are not very productive. However, our results point to a small effect, which is in line with the hypothesis that private sector training (and training for highly educated workers in particular) is likely to generate positive returns.

This paper is organized as follows. In the next section, we briefly describe the *Make IT Work* program. In Section 3, we delineate the data collected during the retraining program and the register data from Statistics Netherlands that we used to enrich input from the program. In Section 4, we address the evaluation design and the methodology before turning to the results in Section 5. We finish with a conclusion and a discussion in Section 6.

2 The *Make IT Work* Program

The *Make IT Work* retraining program has been implemented in the Amsterdam area in the Netherlands in September 2015. The program has been designed in consultation with IT firms from the same geographical area. The program has two goals: (i) to fill the growing number of IT vacancies in the Netherlands in general, and in the Amsterdam metropolitan area in particular; and (ii) to enhance the labor market outcomes of the participants. By design, the program consists of two semesters. In the first semester, the participants attend a full-time vocational IT training program at the Amsterdam University of Applied Sciences (AUAS). The training program is focused on computer programming and software engineering. In the second semester, the participants attend the training for one day per week and follow a paid internship at a firm for the remaining four days. After successful completion of the program, the participants receive a certificate, but no formal university of applied sciences degree.

Applicants are first required to take an on-line personality and cognitive aptitude test. Upon successful completion of the cognitive aptitude test (i.e. a 50% score or higher), the candidate is invited to an organized event called the *employers' market*. The AUAS is responsible for organization of the employers' market, recruitment of the candidates, and identifying firms that might potentially offer an internship. At the employers' market, which lasts a few hours, candidates meet with potential IT employers. At this phase, candidates are required to find an employer for the internship stage of the program. Only candidates who find a prospective employer are admitted to the training program, which starts two to three months after the employers' market. Due to lack of spaces, not all participants who passed the test are able

to find an employer and therefore not every applicant can start the program. Figure 1 gives an overview of the different stages of the *Make IT Work* program, and how the program fits into the 36-month evaluation period of this paper.

The internship is a requirement to both start and successfully finish the program. Vice versa, the training is also required to start the internship and to be considered a suitable candidate by any IT employer. The employer gives a written commitment to offer a temporary, 6-month employment contract for the internship phase. However, this is conditional on the prospect that the candidate accumulates the required skills before starting the internship, by successfully completing the training phase. It is a possibility that the interns are offered a job at the same firm after completing the program at the end of the internship, but this is no guarantee *ex ante*. In line with this, it is unlikely that the participants are offered a job at the same or any IT firm without successfully completing the training and therefore lacking the necessary skills.

3 Data

3.1 Application Procedure

Potential candidates for the *Make IT Work* program are required to submit their application online, at the program's website www.it-omscholing.nl. Anyone with at least a formal bachelor's degree is eligible to apply. Apart from this minimum level of education, there are no further entry requirements. The online selection platform is based on a widely-used recruitment pre-selection algorithm called *Harver Talent-Pitch*. An interesting feature of the application process is that every applicant takes a cognitive ability test. Only applicants who score more than 50 percent are permitted to continue their application to the next step, a personality test. The outcome of the personality test is not binding for completion of the application, and it is only available for review by the applicant who passes the cognitive aptitude test. Once the applicants complete all steps, they are invited by a staff member to visit a 'match-making' event where they meet with potential IT employers. The goal of this event is to give applicants an opportunity to acquire a letter of intent from an employer to start the training program. With this letter of intent, the candidate is assured of an internship with a firm during the second semester of the program. The letter of intent is a prerequisite to starting the training program.

Using data from the application process, we can make an exploratory analysis of the factors related to the probability of program participation. Table 1 presents descriptive statistics of the data that the program collects during the application procedure. Column 1 shows the means and standard deviation of all people who pass the cognitive ability test. Column 2 gives these statistics for those who start the training program, and column 3, for those who do not start the training. Column 4 shows the differences in the means of the two groups and the corresponding p-values. Based on the raw numbers, many more people pass the cognitive ability test than eventually start the training. Fewer than 20 percent of the people who pass the

Table 1 Descriptive statistics applicants, before matching

	(1)		(2)		(3)		(4)	
	Total passed test		In training		Not in training		Mean difference	
	Mean	SD	Mean	SD	Mean	SD	Diff	p-value
Ability score	56.55	4.78	58.35	5.20	56.15	4.59	2.20	***
Problem analysis	62.20	15.51	64.81	15.09	61.62	15.55	3.19	***
Verbal comm	43.90	18.29	45.19	16.32	43.69	18.69	1.50	
Sensitivity	47.99	15.66	49.48	15.47	47.66	15.69	1.82	*
Persuasiveness	49.60	14.69	50.25	13.52	49.45	14.94	0.80	
Team working	44.82	15.21	47.60	15.25	44.20	15.14	3.40	***
initiative	52.52	12.92	52.21	12.56	52.59	13.00	- 0.38	
Age (31 Dec. 2017)	37.13	9.19	33.63	6.58	37.89	9.50	- 4.26	***
<i>N</i>	1391		250		1141			

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

cognitive test decide to participate. Inspecting the differences in means in columns 2 and 3, it looks like a selection on ability for those who do proceed. The internship firms do not observe this ability score, however. The group that enters into training is, on average, four years younger than the group that does not. The average participant is in his or her mid-30's at the time of enrollment. This is approximately ten years older than the age of the average bachelor's graduate, so the *Make IT Work* program attracts not only new graduates, but also older workers.

Table 2 gives the results of three ordinary least squares models for the probability of participation in the program. The underlying data set is the same as the one for Table 1, but with standardized ability and personality scores. Based on column 2, it seems that although the selection is based primarily on ability scores, it takes into account personality scores, including problem analysis and team-working skills. This was also found in a discrete choice experiment described in Vooren et al. (2019b), investigating preferences for IT retraining candidates. The results of the experiment also show strong preferences for team-working skills among employers. If we add age to the specification in column 3, the coefficient of the ability score becomes smaller. This suggests that there is also some selection based on age. However, the coefficients are not very large, and the R^2 is relatively low. It might be that the actual selection is based on other variables correlated with age, such as personality scores, and résumé indicators such as education level of education.

3.2 Evaluation and Matching Data

To evaluate the retraining program, labor market outcome variables are needed in addition to data from the program. To obtain the labor market variables, we utilize longitudinal micro data from Statistics Netherlands. These data allow us to track each inhabitant of the Netherlands in terms of employment status and earnings. The source of these data is the administration of the employee insurance agency

Table 2 Linear probability model for training acceptance

	(1)	(2)	(3)
	Treat	Treat	Treat
Ability score	0.068*** (0.010)	0.062*** (0.011)	0.053*** (0.011)
Problem analysis		0.034*** (0.013)	0.032** (0.013)
Verbal communication		0.006 (0.015)	0.009 (0.015)
Sensitivity		- 0.012 (0.014)	- 0.010 (0.014)
Persuasiveness		- 0.007 (0.018)	- 0.006 (0.018)
Team working		0.036** (0.015)	0.027* (0.015)
Initiative		- 0.030* (0.018)	- 0.024 (0.017)
Age			- 0.058*** (0.011)
Female			- 0.010 (0.023)
Constant	0.180*** (0.010)	0.183*** (0.010)	0.186*** (0.013)
<i>N</i>	1391	1352	1352
<i>R</i> ²	0.031	0.042	0.063

Standard errors in parentheses. Continuous variables standardized to mean zero and standard deviation one

p* < 0.10, ** *p* < 0.05, * *p* < 0.01

in the Netherlands. This administrative data set (in Dutch: *Polisadministratie*) provides monthly data on gross earnings and working days that we include as outcome variables.

Not all participants start the training program at the same time. The treatment is distributed over fourteen cohorts, with each cohort starting in September, November, February, or April. The first group of participants started the program in September 2015, and the last group that is included in our analysis began in November 2018. As a first step, the data set containing information on the participants is matched to the Statistics Netherlands register data. For privacy reasons, this cannot be done one-to-one on the basis of social security numbers. Instead, Statistics Netherlands generates an anonymized and randomized identifier (RIN) on the basis of the address, date of birth, and gender of the participants in our data set. Consequently, we can only observe the outcome variables for 242 out of the 250 participants. In other words, 3% have not been matched to an RIN and therefore cannot be included in the effect evaluation. However, the impact of this small percentage on the analysis

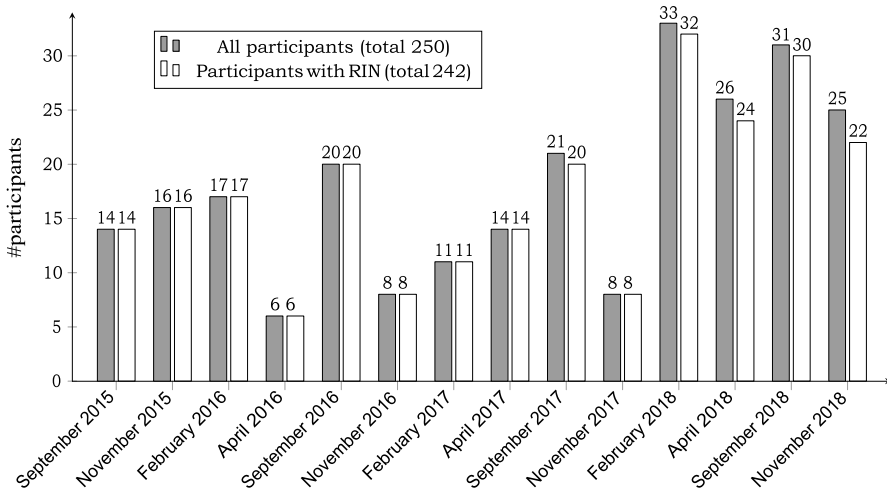


Fig. 2 Number of participants by treatment cohort

results is presumably negligible. Figure 2 shows the distribution of the participants over the various cohorts.

For the matching population, we use the applicants who passed the test but decided not to participate in the program, which corresponds to the group in column (3) of Table 1. Hence, the matching procedure captures all of the three selection steps, which are: (i) applying online, (ii) passing the cognitive test, (iii) and finding an employer for the internship in the second semester. In the original data set, this amounts to 1141 observations. However, 18% of the non-participating applicants who passed the test cannot be matched to the register data. This percentage is larger than in the treatment group, because we do not have the exact address for the non-participants. The group of non-participants has been matched to the register data on the basis of birth date, gender, and place of residency, which is less precise than birth date, gender, and exact address. This yields 936 observations for the matching population.

3.3 Outcome Variables

To evaluate the effectiveness of the program, we use gross earnings per month and the number of working days per month as outcome variables. The first outcome variable, gross earnings per month, is the most important from the participants' point of view. It also allows us to measure the loss of income during the training phase, which serves as a monetary measure of the lock-in effect. The reason we use monthly earnings as a measure, instead of daily or hourly wages, is because this is the way that earnings are registered in the data. Using daily or hourly wages would require assumptions about the number of actual working days or hours per month, which might lead to skewed results. It is also difficult to make assumptions in the

case of flexible hours and part-time work, so total monthly earnings is the most accurate measure.

The second outcome variable, workings days per month, provides information about the participants' employment status. Again, we focus on working days per month, instead of working hours per week, because this is the way it is stated in the data. For the social security system in the Netherlands, it does not matter whether a working day (*SV-dag*) lasts one hour or eight: it counts as a day, just as a maximum of five working days counts as a week and more cannot be registered. In addition to monthly earnings, the number of working days per month gives us information about the participants' employment status during and after the program. While we cannot observe whether the participants work in the IT sector, with this outcome variable, we can evaluate whether the goals of the program are met.

The outcome variables are extracted from the aforementioned register data set from the Dutch Employee Insurance Agency (*Uitvoeringsinstituut Werknemersverzekeringen* or UWV). The variables in which we are interested contain the raw unprocessed information that employers submit to the tax authorities each month. These raw monthly data are not corrected for incidental payments, such as holiday allowances or end-of-year bonuses. For instance, some workers receive an end-of-year bonus while others do not. Also, the statutory holiday allowance can be paid out in a single specific month, or spread out over twelve months. In order to eliminate these irregularities in the data, we divide the matching population into 14 groups, corresponding to the 14 cohorts with different start dates, based on the time of application. In this way, we are able to observe the outcome variable for different months relative to the program start. We always compare treatment units with control units within the same year and calendar month.

4 Methodology

4.1 Matching Procedure

To address potential selection bias in the treatment group, we apply a matching strategy to account for selection based on observable characteristics. The identifying assumption relies on common support, which entails observing non-participants with the same characteristics as the participants. To achieve common support, it is necessary to have good data on observable characteristics of the participants and non-participants (Imbens, 2004). As described in Section 3, we use the rich data from Statistics Netherlands for this. As matching algorithms (Leuven & Sianesi, 2003), we apply Mahalanobis, 1st-nearest neighbor (without caliper and with replacement), and Gaussian kernel matching and compare the results among these three methods. Mahalanobis matching relies on a distance metric (Rubin, 1980), whereas the nearest neighbor and Gaussian kernel matching rely on the estimation of the propensity scores. In the evaluation literature, matching is frequently applied to estimate the effect of training methods (Heckman et al., 1997, 1998) due to its ability to produce inferences very close to randomized experiments.

We use the nonparticipants as a matching population instead of the entire Dutch population for the following reasons. First, the data we have available on the personality and ability scores allow for a match on characteristics that would otherwise be unobservable. Second, the participants of the Make IT Work program are more comparable to those who applied but did not participate than to the general Dutch population. For instance, the labor market positions and levels of education are more similar in these two groups. Most nonparticipants were not able to participate due to lack of space in the program and limited internship opportunities. We are aware of the fact that both positive and negative selection into the Make IT Work program may be at play here. This could be that they found a better job opportunity in the meantime, whereas training participants did not. Such positive selection out of the program may give rise to an underestimate of the returns to the training. Alternatively, nonparticipants may have moved out because they failed to find an employer. Such negative selection out of the program would likely give rise to an overestimation. However, the employers had the personality and ability scores at their disposal at the applicant level when hiring candidates. We use these variables in the matching procedure, assuring conditional mean independence with respect to these personality and ability differences.

The observable characteristics that we include are (i) a dummy variable for gender, (ii) dummy variables for migration status (i.e., native Dutch, first generation migrant, or second generation migrant), (iii) a dummy variable for whether the individual is receiving a disability insurance benefit, and (iv) the year of birth. The inclusion of most of these variables is based on the exploratory investigation of the selection procedure as provided in Section 3.1. Since the application procedure includes consideration of ability and personality scores by employers, we include these variables as well. Ability and personality are not normally observable in the register data, but the unique data from the *Make IT Work* program allows us to include them in the matching procedure. Table 3a shows the means of the treated and control groups after Mahalanobis matching, as well as the corresponding bias and t-tests. Tables 3b and 3c show the same for nearest neighbor and Gaussian kernel matching, respectively.

Comparing Tables 3a, b, and c, it appears that the nearest neighbor matching produces the most comparable control groups based on observable characteristics. This is also reflected in the mean and median bias of the three matching algorithms, which is shown in Table 3d. Taking the mean and median biases of these three matching algorithms into consideration, we prefer the nearest neighbor algorithm over the Mahalanobis and Gaussian kernel neighbor algorithms for the effect evaluation. After matching, all observations rest on common support.

Table 3e shows the average pre-treatment labor market outcomes of the treatment and control groups, for each of the three different matching methods. These variables are observed seven months before the start of the program, which is well before the moment when the participants apply for the program. We do not observe

Table 3 Descriptive statistics, after matching

(a) Mahalanobis matching					
	Treated (N = 242)	Control (N = 242)	Bias (%)	t-test	
				t	p-value
Ability score	58.39	57.71	13.6	1.14	0.16
Problem analysis	64.57	64.17	2.6	0.29	0.77
Verbal communication	44.95	44.19	4.3	0.48	0.63
Sensitivity	49.42	48.84	3.7	0.39	0.70
Persuasiveness	50.17	50.51	- 2.5	- 0.28	0.78
Team working	47.61	46.27	8.8	0.96	0.34
Initiative	51.81	52.47	- 5.1	- 0.58	0.57
Female	0.28	0.28	0.0	- 0.00	1.00
Migrant	0.21	0.21	0.0	0.00	1.00
Disabled benefit recipient	0.02	0.02	0.0	- 0.00	1.00
Year of birth	1984.1	1983.2	11.2	1.35	0.18
(b) 1st-nearest neighbor matching					
	Treated (N = 242)	Control (N = 242)	Bias (%)	t-test	
				t	p-value
Ability score	58.39	58.24	3.0	0.30	0.77
Problem analysis	64.57	64.48	0.5	0.06	0.96
Verbal communication	44.95	44.46	2.8	0.28	0.78
Sensitivity	49.42	49.49	- 0.5	- 0.05	0.96
Persuasiveness	50.17	49.77	2.8	0.28	0.78
Team working	47.61	47.11	3.3	0.34	0.74
Initiative	51.81	51.65	1.2	0.12	0.90
Female	0.28	0.29	- 2.4	- 0.25	0.80
Migrant	0.21	0.20	2.4	0.25	0.80
Disabled benefit recipient	0.02	0.02	1.3	0.15	0.88
Year of birth	1984.1	1984.0	1.4	0.16	0.88
(c) Gaussian kernel matching					
	Treated (N = 242)	Control (N = 242)	Bias (%)	t-test	
				t	p-value
Ability score	58.39	58.33	1.3	0.13	0.90
Problem analysis	64.57	63.39	7.7	0.79	0.43
Verbal communication	44.95	44.07	5.0	0.51	0.61
Sensitivity	49.42	47.99	9.1	0.96	0.34
Persuasiveness	50.17	49.75	3.0	0.30	0.77
Team working	47.61	46.16	9.5	1.01	0.31
Initiative	51.81	51.07	5.7	0.56	0.58
Female	0.28	0.28	- 1.0	- 0.11	0.91
Migrant	0.21	0.19	5.5	0.61	0.55

Table 3 (continued)

(c) Gaussian kernel matching					
	Treated (N = 242)	Control (N = 242)	Bias (%)	t-test	
				t	p-value
Disabled benefit recipient	0.02	0.02	– 3.2	– 0.34	0.74
Year of birth	1984.1	1984.2	– 2.0	– 0.24	0.81
(d) Matching quality					
	Mean bias			Median bias	
Mahalanobis	4.7			3.7	
Nearest neighbor	2.0			2.4	
Gaussian kernel	4.8			5.0	
(e) Pre-treatment labor market outcomes, 7 months before the treatment					
	Treated (N = 242)	Control (N = 242)	Bias (%)	t-test	
				t	p-value
<i>Mahalanobis matching</i> earnings per month	1233.0	1395.9	– 9.7	– 1.08	0.28
Working days per month	13.60	14.91	– 9.7	– 0.97	0.33
<i>1st-nearest neighbor matching</i> earnings per month	1233.0	1322.4	– 5.3	– 0.59	0.56
Working days per month	13.60	14.56	– 7.1	– 0.70	0.48
<i>Gaussian kernel matching</i> earnings per month	1233.0	1385.8	– 9.1	– 0.59	0.56
Working days per month	13.60	15.68	– 15.5	– 0.70	0.48

any statistically significant differences between the treatment and control groups in terms of monthly earnings and working days, providing evidence for conditional mean independence on observable characteristics.

In addition, these statistics give some numerical support for the hypothesis about the composition of the target group of the Make IT Work program stated in the Introduction section. When looking at the ratio of average monthly earnings and working days, we can conclude that this is just slightly above the daily minimum wage in the Netherlands.¹ Since a considerable proportion target group of this program is likely not eligible for unemployment benefits due to lack of employment history when graduating from university—it is required to have a minimum employment duration of 26 weeks in order to be eligible for unemployment benefits in the Netherlands—it looks like that they predominantly hold on to minimum-wage jobs when applying for the Make IT Work program, even though they possess a bachelor's degree.

¹ As of July 2022, the minimum wage in the Netherlands is EUR 81.06 per day (see <https://www.government.nl/topics/minimum-wage/amount-of-the-minimum-wage>).

4.2 Empirical Model

To calculate the average treatment effect (ATE), we estimate the following two model specifications for the two outcome variables after matching, as shown in Equations 1 and 2. In these models the ATE is represented by β_{1j} , which is the parameter of interest. For each outcome variable, we estimate a total amount of $j \in [0, 35]$ models, where j represents the month relative to the program start.

$$EARNINGS_{ij} = \beta_0 + \beta_{1j} TREATMENT_i + \mu_{ij} \quad (1)$$

$$WORKING\ DAYS_{ij} = \beta_0 + \beta_{1j} TREATMENT_{i+j} \quad (2)$$

We estimate both models using weighted least squares (WLS) with clustered standard errors at the cohort level (see Figure 2), because the cohorts are sampled at 14 different moments (Abadie et al., 2020). This provides more conservative standard errors than the heteroskedasticity-robust standard errors provided by the *psmatch2* package (Leuven & Sianesi, 2003). In the next section, we report the results of estimating the empirical models.

5 Results

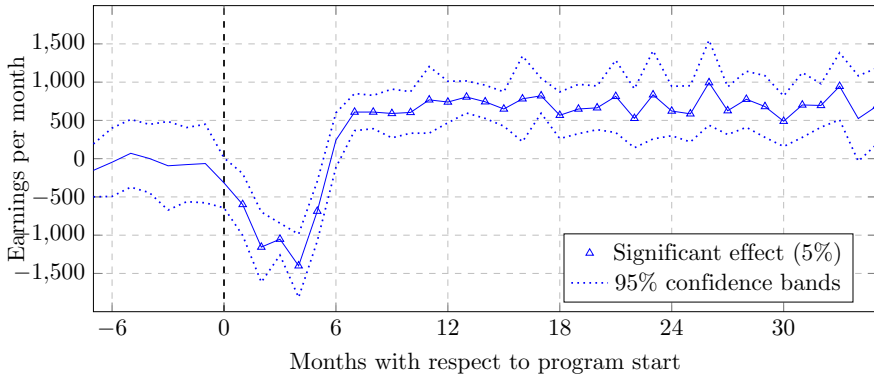
5.1 Effect on Earnings

Figure 3a shows the effects of the training program on gross earnings per month after nearest neighbor matching. In the first five months following the start of the training, there is a substantial negative effect on participants' earnings. This negative effect gradually becomes smaller between five to six months after the start of the training. This moment of six months after the start of the training coincides with the start of the second semester of the program, in which the participants start their internships. After this period, the participants start earning more than the non-participants. This positive effect on earnings amounts around EUR 500 gross per month and is statistically significant at the 5% level. This effect stays statistically significant until the end of the evaluation window of our analysis, which is 36 months after the program start.

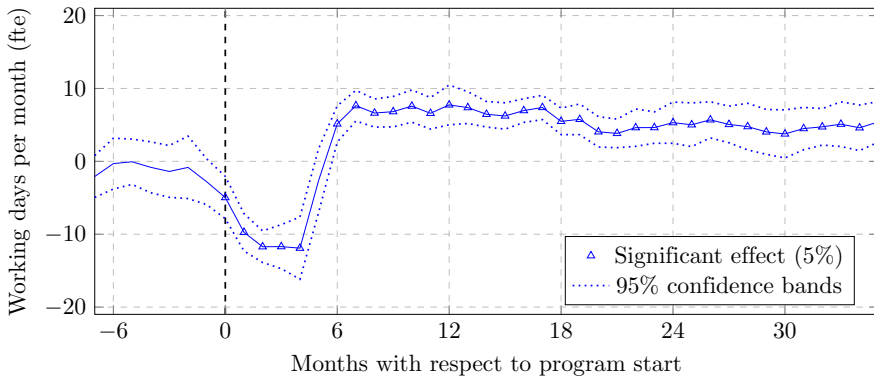
5.2 Effect on Working Days

Figure 3b shows the effects on working days per month. As can be seen, there is a negative effect of around 10 working days per month less until four months after the start of the training. As with the effect on gross earnings per month, this negative effect on working days per month diminishes at five to six months after the start of the training. The negative effect on working days per month turns into a positive effect at six months after the program start. This coincides with the month in which

(a) Gross earnings per month



(b) Working days per month

**Fig. 3** Results after 1st-nearest neighbor matching

the participants start the internship, which explains this jump in working days per month. This positive effect amounts to roughly eight working days per month. This gradually abates after 12 to 18 months after program start, but remains statistically significant over the 36-month evaluation window. The decrease in working days after 18 months might be due to some of the participants reducing their hours after the initial contract during the internship phase. As an additional robustness check, we present the results based on Gaussian kernel and nearest neighbor matching in Figs. 4 and 5 in the Appendix. These methods yield comparable results.

5.3 Cost-benefit analysis

Once we estimate the ATEs of program participation on earnings, we can use those estimates to calculate the program's IRR. The ATE measures the average effect on earnings when participating versus not participating. Therefore, the ATEs that we use in the cost-benefit analysis take the lock-in period into account. For the participants, the out-of-pocket costs of the retraining program consist of a EUR 1000 down

payment, EUR 300 of which has to be paid up front and the remaining EUR 700 in the second month of the program. The employers pay a EUR 6000 fee per participant after the end of the first semester, but, as it is not paid by the participants, we do not include the employer contribution in the cost-benefit analysis. This leads to the following formula for the net present value (NPV):

$$NPV = \sum_{t=0}^{35} \frac{ATE_{earnings,t}}{(1+r)^t} - \frac{300}{(1+r)^0} - \frac{700}{(1+r)^2} \quad (3)$$

The IRR is the discount rate r for which the NPV equals zero. When we set equation 3 to zero and solve for r , this gives an estimate of the internal rate of return (IRR). Using the ATE estimates and 95% confidence bands in Figure 3a, this procedure leads to an estimated IRR of 7.80% for the three-year evaluation period. In order to compare this IRR with an annual rate of return, such as the return to schooling, we annualize this IRR to a compound annual growth rate. In doing so, the annual IRR equals 2.53% per year (95% CI 2.02–3.04%). This implies that for any interest rate smaller than 2.53%, the NPV of participating in the retraining program is positive for the first three years after the program start. This is a conservative estimate, because it is possible that the IRR will increase when we expand our 36-month evaluation period and the positive effects on earnings are still positive and significant in the period thereafter. To construct the 95% confidence interval, we have bootstrapped the 36 wage coefficients with 10,000 replications. This gives us 10,000 approximations of the IRR estimate, giving the sampling distribution and standard error of the IRR estimate used for the confidence interval.

To get an idea of what the returns to the training would look like on a longer time horizon, we extrapolate the ATE during the two years after the 12-month training and internship period.

We believe it is safe to assume that these effects will continue for more years, looking at the steady effect on earnings displayed in Figure 3a. When we extrapolate the ATE with another 24 months by taking the mean of the ATE in month 12 to 35, the annual IRR for this 5-year horizon amounts to 2.75%. Repeating this exercise for a 10-year horizon results in an annual IRR of 2.78%. For a 15-year horizon, the IRR remains at 2.78%.

To put an IRR of 2.53% for retraining program into context, we can compare it to the rate of return to a year of education. Recent overviews conclude that the annual return to one additional year of schooling (Psacharopoulos & Patrinos, 2018) is about 9%. For the Netherlands, recent estimates of the return to one additional year of schooling are around 8% (Gerritsen & Hartog, 2016). Compared to these returns, an annual IRR of 2.53% is low, even though it is statistically significant based on a 95% confidence interval.

6 Discussion and conclusion

The results of this evaluation study answer the main research question of whether IT retraining is an effective tool for improving the labor market outcomes of participants. Better labor market outcomes is the goal of the *Make IT Work* retraining

program. In order to assess the effects of the retraining program, we examine (i) monthly earnings and (ii) working days per month. Considering monthly earnings, this study shows that IT retraining increases earnings over a period from 6 until 36 months after program start, which is the end of our evaluation time frame. Consequently, we conclude that the program is effective at increasing the earnings of the participants in the long run. With respect to the number of working days per month, there is a positive effect from six months after program start onward. This is a persistent effect of around four to five additional working days per month. Therefore, we conclude that the IT retraining program is effective at increasing the number of working days per month in the long run. It is possible that, if we were able to extend the evaluation time frame, the positive effects on earnings and working days would be shown to persist. However, we are unable to extend the time span at this point, because no further data is available. Extending the evaluation timeframe up to 15 years under the assumption that the effect on earnings remains stable does not increase the returns to the training by a significant amount.

In accordance with the recent ALMP evaluation literature, we find substantial, albeit relatively short-lived, negative effects in the months following the program start for both outcome variables. These negative effects last up to five months after the program start. However, most of the ALMPs that revolve around training and retraining are associated with relatively long-lasting lock-in effects (Card et al., 2017; Vooren et al., 2019a). In this respect, the retraining program evaluated in this paper performs relatively well. In general, we conclude that the program goals of *Make IT Work* are met and that, as an ALMP, it demonstrates that IT retraining is effective at increasing both earnings and the number of working days per month. In order to assess whether the positive long-term effects on earnings justify the costs as well as the initial negative effects on earnings, we perform a conservative cost-benefit analysis that includes the costs of both the lock-in effects and the tuition fee. Our cost-benefit analysis points to a 2.53% annual internal rate of return, which leads to the conclusion that this IT retraining program for highly educated workers is associated with positive returns. However, these returns are lower than the returns to schooling estimated for the Netherlands (Gerritsen & Hartog, 2016; Psacharopoulos & Patrinos, 2018). Therefore, we conclude that the program is effective, although the effects are small.

In the introduction to this paper, a couple of reasons were given as to why this particular training program could be potentially more effective than other training and retraining ALMPs. The first reason is that, in contrast to most ALMPs, *Make IT Work* is voluntary. It calls for a willing commitment of both time and money, since participants must make a down payment of EUR 1000. Given these considerations, participants are expected to be more motivated and, accordingly, the program is anticipated to be more successful than other ALMPs. In addition, the participants in *Make IT Work* are highly educated, lending support to the contention of Heckman et al. (1999) that ALMPs are usually ineffectual, because they target 'less able' individuals. Indeed, most programs evaluated in the literature focus on poorly educated workers (Cavaco et al., 2013; Kruppe & Lang, 2018; Wapler et al., 2018; Wunsch & Lechner, 2008). The participants of the *Make IT Work* program must have at least a bachelor's degree to be admitted.

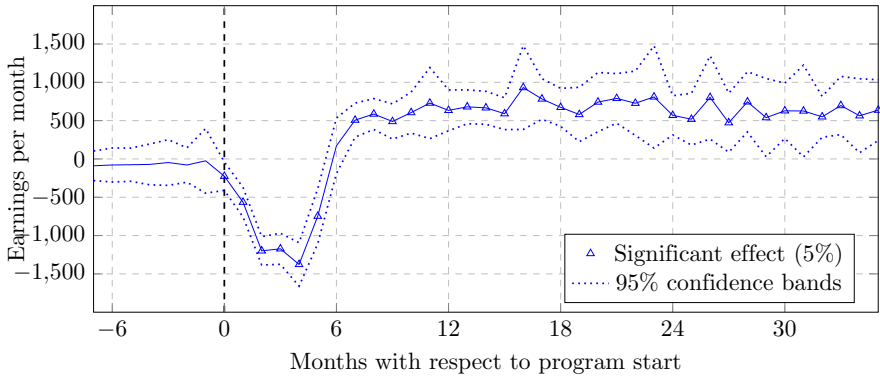
A final argument why ALMPs are usually ineffective is that in most cases, they are administered by public sector parties. Heckman et al. (1999) contend that the internal rate of return associated with public training programs is usually lower than that of private sector training. However, *Make IT Work* is a hybrid: it was designed in co-operation with private sector employers, but implemented by a public sector institution. The results from our study indicate a positive albeit comparatively modest internal rate of return. This outcome serves to demonstrate the positive contribution that the private sector can make to training and retraining programs, even when they are established in cooperation with the public sector. *Make IT Work*'s positive internal rate of return, however limited, nevertheless suggests that it is not necessary to entrust ALMP retraining programs entirely to the private sector to enjoy some success.

The positive return was further strengthened by the combination of classroom training with on-the-job training. On-the-job training is usually associated with shorter lock-in effects than classroom training programs (Card et al., 2017; Vooren et al., 2019a). Last but not least, measures of success such as increased earnings and working days, may also be attributed to the participants' education level and the voluntary nature of the program. Indeed, there might be positive effects on other labor market outcomes related to job stability that we cannot observe in the register data, such as the probability of having a fixed-term contract. Altogether, the results from this paper indicate that IT retraining is effective at improving the labor market outcomes of the participants, although is associated with relatively small returns when compared to the returns to schooling.

Appendix

See Figs. 4 and 5.

(a) Gaussian kernel matching



(b) Mahalanobis matching

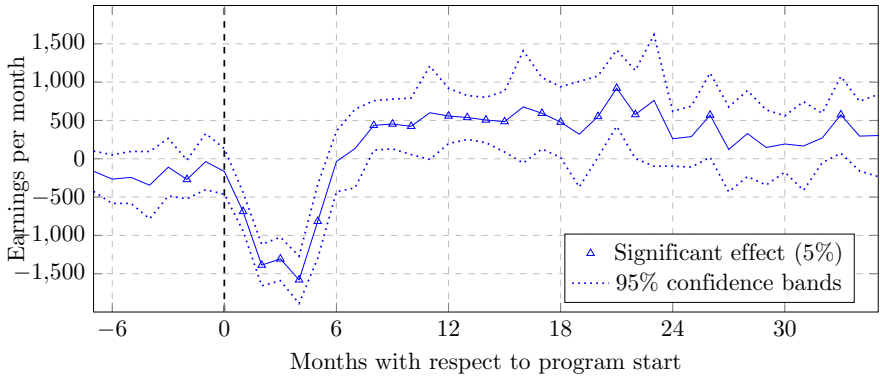
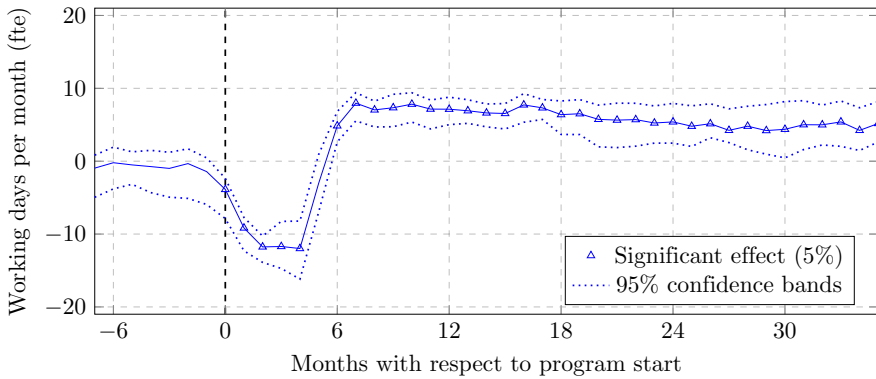


Fig. 4 Results alternative matching algorithms for gross earnings per month

(a) Gaussian kernel matching



(b) Mahalanobis matching

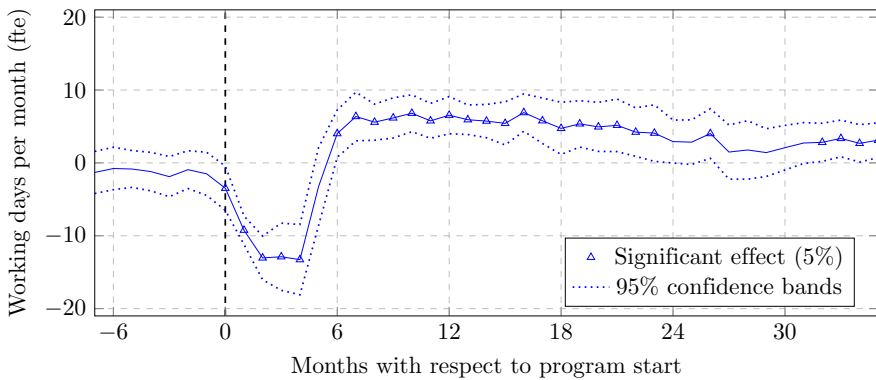


Fig. 5 Results alternative matching algorithms for working days per month

Funding Funding was provided by European Regional Development Fund (Grant No. KVW-00147).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2020). Sampling-based versus design-based uncertainty in regression analysis. *Econometrica*, 88(1), 265–296.
 Card, D., Kluve, J., & Weber, A. (2017). What works? A meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, 16(3), 894–931.

- Cavaco, S., Fougere, D., & Pouget, J. (2013). Estimating the effect of a retraining program on the re-employment rate of displaced workers. *Empirical Economics*, *44*(1), 261–287.
- CBS. (2019). ICT, kennis en economie.
- Gerritsen, S., & Hartog, J. (2016). De ontwikkeling van het rendement op scholing 1962–2012. *Economisch Statistische Berichten*, *100*(4711), 340–343.
- Heckman, J. J., Ichimura, H., & Todd, P. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, *64*(4), 605–654.
- Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *Review of Economic Studies*, *65*(2), 261–294.
- Heckman, J. J., LaLonde, R. J., & Smith, J. A. (1999). *The economics and econometrics of active labor market programs, volume 3A of handbook of labor economics*. Amsterdam: North-Holland.
- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *The Review of Economics and Statistics*, *86*(1), 4–29.
- Kruppe, T., & Lang, J. (2018). Labour market effects of retraining for the unemployed: The role of occupations. *Applied Economics*, *50*(14), 1578–1600.
- Leuven, E., & Sianesi, B. (2003). Psmatch2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.
- Psacharopoulos, G., & Patrinos, H. A. (2018). Returns to investment in education: A decennial review of the global literature. *Education Economics*, *26*(5), 445–458.
- Rubin, D. B. (1980). Bias reduction using Mahalanobis-metric matching. *Biometrics*, *36*(2), 293–298.
- Vooren, M., Haelermans, C., Groot, W., & Maassen van den Brink, H. (2019a). The effectiveness of active labor market policies: A meta-analysis. *Journal of Economic Surveys*, *33*(1), 125–149.
- Vooren, M., Haelermans, C., Groot, W., & Maassen van den Brink, H. (2019b). Employers' preferences for IT-retrainees: Evidence from a discrete choice experiment. *International Journal of Manpower*, *40*(7), 1273–1287.
- Wapler, R., Werner, D., & Wolf, K. (2018). Active labour market policies in Germany: Do regional labour markets benefit? *Applied Economics*, *50*(51), 5561–5578.
- Wunsch, C., & Lechner, M. (2008). What did all the money do? On the general ineffectiveness of recent West German labour market programmes. *Kyklos*, *61*(1), 134–174.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.