Long-term effects of job search training:

Evidence from two randomized field experiments[1]

Kari Hämäläinen*

Roope Uusitalo**

Jukka Vuori***

* Government Institute for Economic Research

** Helsinki School of Economics and Labour Institute for Economic Research

*** Finnish Institute of Occupational Health

Abstract

The Finnish Institute of Occupational Health implemented two randomized field experiments to evaluate the effects of job search training in the late 1990s. We link data from these experiments to administrative register data on the participants to evaluate the long-term effects of job search training. We also compare the experimental results to various matching estimators calculated using program participants and external comparison groups.

Our experimental results indicate that job search training has no significant effects on average employment rates. We also find that commonly used non-experimental matching estimators tend to over-estimate the program effects and that the bias differs between programs due to different procedures used in selecting the program participants. The bias tends to be larger when participation is voluntary. Finally, selective non-response in the follow-up surveys leads to higher estimates than those calculated based on administrative data.

[1] We would like to thank the Finnish Ministry of Labor for financial support and Tiina Salokangas and Elina Nykyri for their excellent research assistance.

1 Introduction

Job search training is a common element of labor market programs in many European countries. These training courses are aimed at improving job search skills and provide help in writing a job application and cv, contacting employers and preparing for a job interview. The courses also attempt to help unemployed to recognize their skills and act as support groups that prepare for setbacks in the job search process. The policy is attractive because the training courses are rather short, typically between one and two weeks, and job search training, therefore, potentially cost-effective way of getting the unemployed back to work. Still there are few rigorous evaluations of their effects on re-employment rates.

In 1996, the Finnish Institute of Occupational Health carried out a demonstration project for evaluating the employment effects of job search training programs. The program was based on the JOBS-program originally developed at the Michigan Prevention Research Centre. The program participants were selected by randomly allocating a half of the applicants to a training course and a half to a control group that did not participate in training. The effects of job search training were evaluated using two follow-up surveys conducted six months and two years after participation.

Three years after the first experiment job search training courses were added as standard element to more extensive labor market training programs. In connection to this reform, another randomized field experiment was carried out in nineteen employment offices across the country. Again participants were randomly chosen, this time by assigning two thirds of the eligible applicants to a job search training course and leaving one third in the control group. Impacts of training on employment rates were evaluated based on a follow-up survey six months after participation.

Both courses were based on group intervention lasting for one week. In both cases over one thousand unemployed took part in the experiment either as treatment or control group members. The main difference in content was that the demonstration project followed closely the intervention design with the researchers monitoring recruitment and implementation, while the latter courses were job search training courses that applied various methods and that were organized by local employment offices. Another key difference had to do with the selection process. Participating in the demonstration

project was entirely voluntary and the participants were recruited by the program administrators. In the second experiments the participants were selected by employment offices using their usual procedures.

The results based on the follow-up surveys have been reported previously by Vuori, Silvonen, Vinokur & Price; 2002 and Malmberg-Heimonen and Vuori; 2005). In this paper we combine data from the two experiments to administrative register data on participants. This allows us to trace the effects of job search training on a monthly basis from the date of randomization up to four or six years after participation. We also collect data from a random sample of persons that were unemployed during the experiments and use these data to calculate various non-experimental matching estimators on the program effects.

Our paper contributes to the literature that compares non-experimental estimates experimental results, a tradition started by LaLonde (1986) analysis of National Supported Work Programme. The LaLonde study motivated another large scale randomized experiment evaluating the effects of Job Training Partnership Act. The results from this experiment have been evaluated in several papers by Heckman and various co-authors. (eg. Heckman, Ichimura, Smith and Todd, 1998).

However, so far the most of the evidence comes from the U.S. Even though European countries rely more heavily on active labor market programs, very few experimental studies have evaluated the effects of European programs. European labor markets are characterized by, on average, much higher unemployment rates, higher fraction of long-term unemployed and more generous unemployment benefits. It would be important to know whether the same principles of intervention that have been shown to promote reemployment in the U.S. context also work in European countries. Furthermore, the ability of non-experimental methods to replicate experimental results may depend on the selection process into the programs. These processes vary across countries and across different programs. By comparing the results based on two different experiments, we will be able to evaluate how different selectivity influences the bias in non-experimental procedures.

The rest of this paper is organized as follows. First in Section 2 we describe the contents of the intervention and the design of randomized field experiment. Next in Section 3 we

report how the participants of the experiment were traced from the register data and compare the self-reported outcomes to those deduced based on administrative registers. In Section 4 we first report long-term outcomes of job search training programs and compare these results to various non-experimental matching estimators. Section 5 concludes with some final comments.

2. Intervention

The MPRC Job Search Program was designed at the Michigan Prevention Research Center as a preventive intervention for recently unemployed job seekers to facilitate their return to labor market and to prevent the negative mental health consequences of unemployment. The program aimed at supporting the role change from passive unemployed individuals to active job seekers. The intervention was designed to maximize active learning processes through active engagement in group problem solving, discussions, and role plays. The active learning processes were also designed to increase participants' job-search self-efficacy and their inoculation against setbacks. (Vuori and Vinokur 2005) The Finnish version of the program, named "Työhön"-program, was guided by same principles with minor procedural changes. For example, in contrast to the U.S. version, the Finnish program also recruited long-term unemployed to job search training.

In practice the Työhön-program consisted of five half-day sessions organized in groups with 6 to 17 unemployed participants in each group. The training was designed to enhance job search skills by a) recognizing one's marketable skills, b) identifying and using one's social networks to find job openings, c) contacting promising employers, d) drawing up a job application and resume, and e) preparing for successful job interviews. Each group had two trainers, one male and one female. These trainers themselves were recruited from unemployed job seekers and trained for two months.

Recruitment and randomization

The participants were recruited from four employment offices in southwestern Finland. To become participants, respondents had to agree on the randomization procedure and to turn in a baseline questionnaire. About 5000 unemployed workers were contacted and

1471 agreed to participate in the experiment. Almost 90 percent of the participants were unemployed at the time when they were recruited. Median duration of elapsed unemployment was five months. Others had received a termination notice or were searching for a new job for some other reason. Only unemployed were eventually allowed to participate in the experiment.

The 1261 eligible respondents who completed the pre-test questionnaire were randomized into treatment (n = 629) and control groups (n = 632). Those randomized into the treatment group were invited to participate in the program that typically started about three weeks after randomization. Of these 70 percent started in the program. Most common reason of dropping out was finding a job before program started.

Job search training experiment 1999

The second experiment differed from the demonstration program in a number of ways. First, the job search training programs were part of the usual re-employment services of the employment offices. The participants took part in one of the 137 job search training groups organized by the participating 19 local employment offices. In most cases the groups included also other participants than those taking part in the randomized experiment. Including these other participants the group size was between 8 and 20. Program length varied from 3 to 10 days (median 5) and hours per day from 3 to 7 (median 5). The trainers were career advisors employed at the employment offices and job search training was part of their usual duties. There was also variation in the content and the quality of training. Some courses relied on the MPRC Job Search Program while others used other methods or could not specify the method used. Variable quality was also reflected in participant evaluations. Variation in quality assessments across courses was much larger than in the strictly structured MPRC Job Search Program implemented three years earlier.

Recruitment and randomization

The participants to job search training were selected by caseworkers at employment offices. Only those who had no previous job search training but who, according to

caseworker assessment, would benefit from job search training were selected to the program. Since job search training groups were part of the usual services of the employment offices, refusing to participate could lead to sanctions and reductions of unemployment benefits.

After being screened by the caseworkers, the unemployed were approached by mail or personally in connection to their regular visits at employment offices and asked if they were willing to participate in an experiment evaluating the effects of job search training. They were also told that program participants would be randomly assigned to treatment and control groups. Those who refused were entitled to usual re-employment services. Those who wanted to attend job search training immediately but did not want to participate in the experiment were placed to training programs. Only those 1017 unemployed who agreed to randomization procedure and returned the baseline questionnaire were then randomly allocated to treatment (677) and control (338) groups. The treatment group was invited to a job search training program starting immediately. Those assigned to control group were invited to participate to job search training programs that would start later, scheduled at least seven months after randomization.

3. Data collection

In order to find those who took part in the experiment from administrative registers we first had to attach a personal id-number to each person. This was never asked directly from the participants. However, Finnish Institute of Occupational Health kept records of names and addresses so that the follow-up surveys could be sent to the right person. Using this information together with answers to question on the date of birth we were able to find almost all participants from the Population Register. Only three persons participating in the first and one person participating in the second experiment were not found from the Population Register. Also for almost everyone, we could be certain that we found a right person. Only in five cases in the first experiment and three cases in the second experiment we could not find exact matches of current or previous addresses.

This data that now only included person-id and limited amount of basic information (treatment status, date of participation, key outcomes) was then sent to Statistics Finland. There the experimental data was linked to Employment Statistics, an

administrative register database containing information on all persons living in Finland. Employment Statistics database is based on more than twenty different registers. For our purposes the most important information sources were the Job Seekers register of the Ministry of Labor containing all unemployment spells, the Tax register that reports annual earnings and the Pension Register reporting all employment spells that count for future pensions. Currently individual level data is available on an annual basis from 1987 to 2003.

Statistics Finland attached information on unemployment and employment spells, annual earnings and various background variables (education, age, occupation, region of residence, number and ages of children, marital status, unemployment benefits) to each participant's records. Data was collected from 1993 to 2003 allowing us to track the labor market outcomes both before experiment and up to four or six years after the experiment. We also collected a random sample of 5000 persons who were unemployed during at least one of the experiments from the same register to be used as a non-experimental comparison group. Finally, to comply with the data protection requirements, Statistics Finland removed all information that would allow identifying individual participants and returned this anonymized data to us.

We used data on employment and unemployment spells and information on the labor market status in the end of each year to create an indicator of labor market status on a monthly basis. In most of the analysis we classify the individuals into four categories: employed, unemployed, in labor market training and outside of the labor force. Knowing the date when the participants were randomized to the treatment and the control groups we could then re-define the time so that we can measure labor market outcomes during each month after randomization date up to four years in the first and up to six years in the second experiment.

Differences between register and survey data

In the follow-up surveys conducted six months after the training program, both the treatment and comparison group members were asked about their current labor market status. Thirteen alternatives were given. We re-classified the answers into the following categories: 1) employed, 2) unemployed, 3) in labor market training and 4) outside the

labor force. In order to be comparable with the register data, the first category also included those in subsidized jobs. The second and third categories are easily defined both from register and from the survey data. The last category included all others, according to the survey responses these were most often mothers of small children or full-time students.

The results based on survey responses may well be different from the results based on register data. The program participants have typically rather unstable job market attachment and responses may vary according to the exact reference date. In some cases it is unclear how a labor market status should be defined, for example, if a person is employed but works only a few hours per week, and does not necessarily work at all in the reference week even if s/he happens to have a valid employment contract. To assess these differences we first cross-tabulated the labor market status six months after the Job search training program according to the survey and register data in Table 1a. In Table 1b we do the same for the 1999 experiment.

Table 1a Labor market status 6 months after treatment. Työhon-program 1996/7

	Labor market status according to the register data					
	Employed	Unem-ployed	In labor market training	Outside of the LF	Missing	Total
Labor market status in the survey data						
Employed	455 (76.7 %)	34 (7.6 %)	2 (1.4 %)	9 (14.3 %)	1 (25.0 %)	501 (40.0%)
Unemployed	71 (12.0 %)	374 (83.9 %)	9 (6.1 %)	20 (31.8 %)	1 (25.0 %)	475 (37.9%)
In labor market training	26 (4.4 %)	21 (4.7 %)	134 (90.5 %)	9 (14.3 %)	0 (0.0 %)	190 (15.2%)
Outside of the labor force	23 (3.9 %)	9 (2.0 %)	0 (0.0 %)	21 (33.3 %)	0 (0.0 %)	53 (4.2%)
Missing	18 (3.0 %)	8 (1.8 %)	3 (2.0 %)	4 (6.4 %)	2 (50.0)	35 (2.8)
Total	593 (100.0 %)	446 (100.0 %)	148 (100.0 %)	63 (100.0 %)	4 (100.0 %)	1254 (100.0%)

Table 1b Labor market status 6 months after treatment. Job search training 1999

	Labor market status according to the register data					
	Employed	Unem-ployed	In labor market training	Outside of the labor force	Missing	Total
Labor market status in the survey data						
Employed	262 (64.2 %)	10 (2.2 %)	1 (1.5 %)	8 (10.5 %)	0 (0.0 %)	281 (27.9 %)
Unemployed	64 (15.7 %)	376 (83.0 %)	7 (10.3 %)	14 (18.4 %)	2 (66.7 %)	463 (45.9 %)
In labor market training	11 (2.8 %)	4 (0.9 %)	54 (79.4 %)	7 (9.2 %)	0 (0.0 %)	76 (7.5 %)
Outside of the labor force	12 (2.9 %)	3 (0.7 %)	0 (0.0 %)	33 (43.4 %)	1 (33.3 %)	49 (4.9 %)
Missing	59 (14.5 %)	60 (13.3 %)	6 (8.82)	14 (18.4 %)	0 (0.0 %)	139 (13.8 %)
Total	408 (100.0 %)	453 (100.0 %)	68 (100.0 %)	76 (100.0 %)	3 (100.0 %)	1008 (100.0 %)

According to the results in Tables 1a and 1b, the register data and the survey data are in most cases consistent, especially so for those who are unemployed or in labor market training. We also usually classify employment in the same way based on register data and survey responses. Results are less consistent for those who are outside the labor force. For this group the quality of register data is questionable. For example, maternity leave spells are not included in the data and many mothers who are on leave will be coded as employed. Also full-time students may be coded as employed even if they work just a few hours per week. One might also note that non-response rate is very small in the first follow-up survey (2.8%). In the latter experiment non-response is much higher (13.8%) but there is no clear sign that non-response would be strongly correlated with the labor market status.

Table 2 breaks down the differences between the follow-up survey and the register data according to the treatment status. According to the results employment rates are clearly higher and unemployment rates somewhat lower in the register data than in the survey data. Still in the first experiment the difference in employment rates between the treatment and control groups is very similar in both the survey and the register data.

Therefore, conclusions on the employment effects of the program do not depend on whether survey or register data is used. Interestingly this result does not hold in the second experiment. There employment rates are about 3 percent higher in the treatment group according to the survey data but about as much lower according to the register data. This difference is partly due to selective non-response in the survey data. If those who did not respond to the survey are coded according to the information in the register data, the employment rates are almost identical in the treatment and in the control groups.

Table 2 Labor market status in survey and register data by treatment status

Työhön-program 1996/7

	Follow-up survey		Register data	
Labor market status	Treatment	Control	Treatment	Control
Employed	253 (41.9 %)	248 (40.3%)	303 (48.6%)	290 (46.3%)
Unemployed	227 (37.6 %)	248 (40.3%)	214 (34.4%)	232 (37.0%)
In labor market training	99 (16.4 %)	91 (14.8%)	71 (11.4%)	77 (12.3%)
Outside of the labor force	25 (4.1 %)	28 (4.6%)	35 (5.6%)	28 (4.5%)
Missing	21	14	2	2

Job search training 1999

	Follow-up survey		Register data	
Labor market status	Treatment	Control	Treatment	Control
Employed	194 (33.3%)	87 (30.4%)	266 (39.7 %)	142 (42.5%)
Unemployed	302 (51.8%)	161 (56.3%)	307 (45.8 %)	146 (43.6%)
In labor market training	51 (8.8%)	25 (8.7%)	50 (7.5 %)	18 (5.4%)
Outside of the labor force	36 (6.2%)	13 (4.6%)	47 (7.0%)	29 (8.7%)
Missing	88	51	1	2

4. Results

In this section we first report the experimental estimates of the employment effects of the job search training programs based on a comparison of the employment rates between the treatment and the comparison groups in the register data. After that we discard the experimental controls and compare employment rates between the treatment group and a matched non-experimental control group. This section, thus, provides evidence both on the long-term effects of the training programs and on the bias in typical non-experimental approaches.

4.1 Long-term effects on job search training

Since the treatment and control groups were randomly chosen, the analysis of the effects of the program is simple. Random assignment assures that the treatment group and the control group do not differ systematically in any other way except access to the program. Thus subject only to uncertainties of sampling error the resulting differences in outcomes can be attributed to the program. Differences in mean outcomes are therefore valid estimates of the program effects.

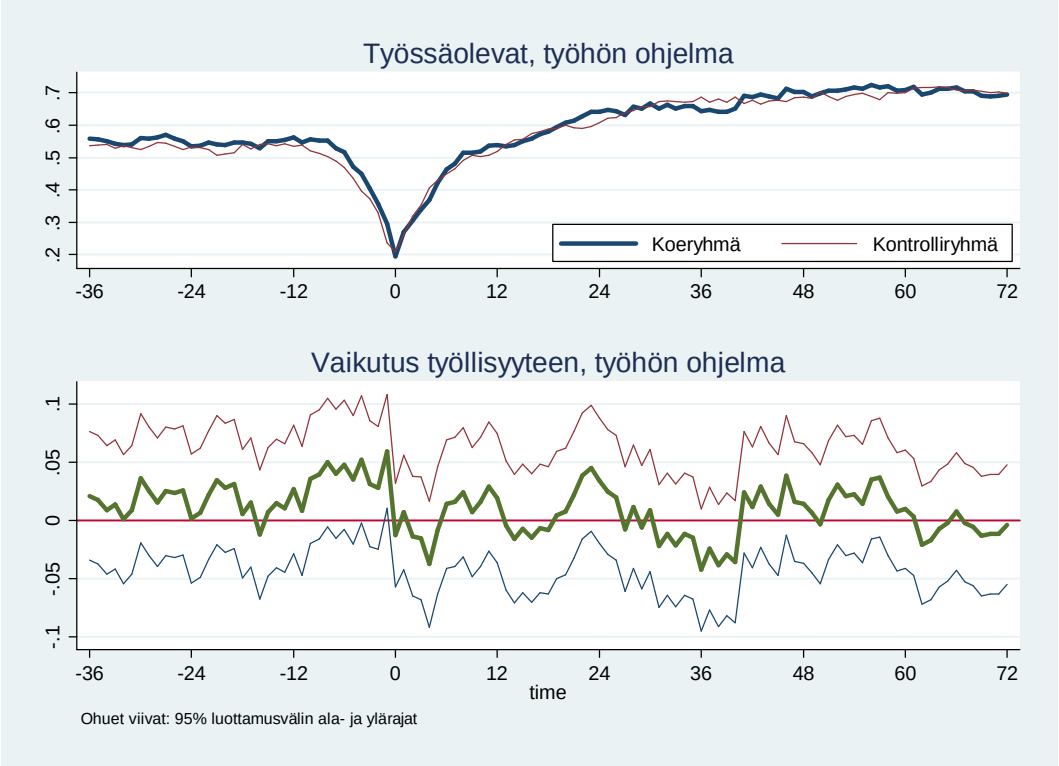
Figure 1 reports the monthly employment rates in the treatment and in the control groups of the Työhön-program starting from three years before the experiment and extending up to six years after experiment. The lower panel of the figure presents the difference between treatment and the control groups together with a 95 percent confidence bands. As can be seen from the figure, employment rates were very similar before the experiment indicating no problems in the randomization procedure. Employment rates decline during the 12 months before program participation. Such a decline is a common feature of the training programs targeted to those who are in a weakest position at the labor market when they participate in a training program.

Employment rates increase rapidly after the experiment but the increase is very similar in the treatment and in the control groups. There is no evidence of statistically significant effects during the entire six year follow-up period. Still, the difference between the treatment and the control groups is almost five percentage points 24 months after randomization, exactly at the time when the second follow-up survey was conducted. In a previous study by Silvonen and Vuori (2005) that was based on this

two-year follow-up survey the difference was very similar but statistically significant in one-tailed test only.[2]

Looking at monthly employment rates calculated from the register data over the entire six-year follow-up period leads to more pessimistic conclusions regarding the program benefits. The aggregate employment effects are close to zero and slightly larger effect 24 months after program participation seems more like random variation.

Figure 1 Employment effects of the Työhön –program

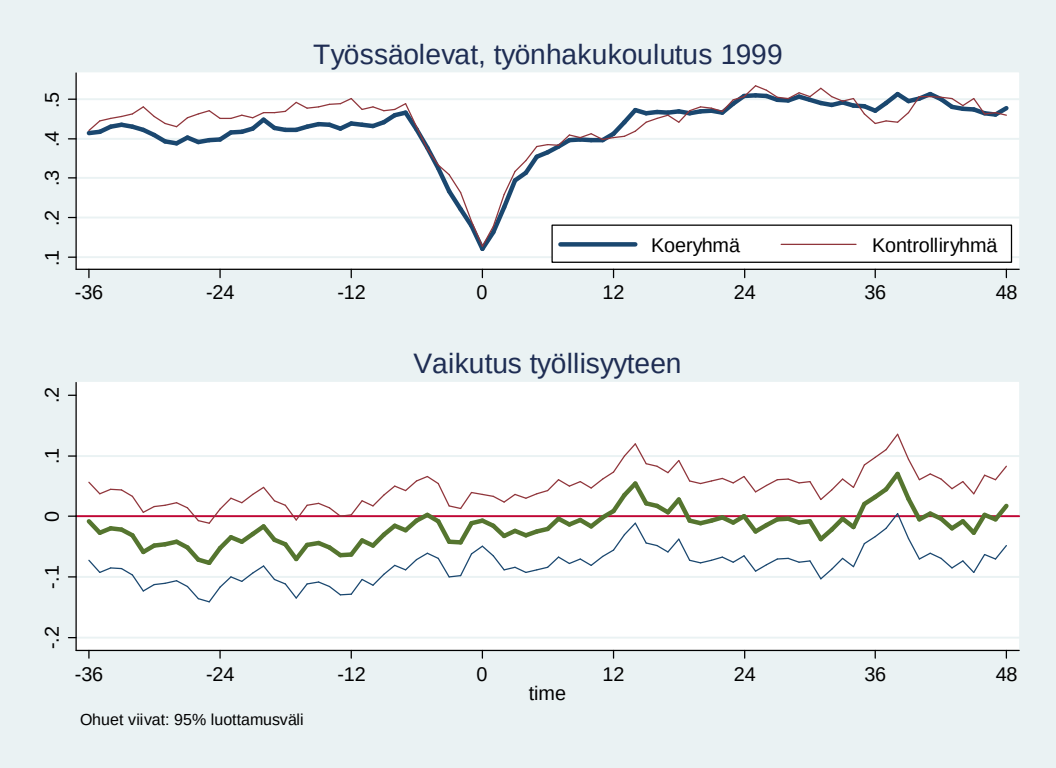


We repeat a similar analysis for the job search training experiment conducted in 1999. Figure 2 plots the employment rates in the treatment and the comparison groups from three years before randomization to four years after the program. The lower panel of the figure again reports the differences in monthly employment rates and the confidence bands of these differences.

[2] Previous research has also identified interesting interaction effects showing that program benefits are significantly larger for the participants at greater risk of depression and for the participants who are at risk of becoming long-term unemployed (Vuori et al 2002). Since our main focus is in the comparison of

The results in Figure 2 are rather similar to those in the first experiment. The difference in employment rates between treatment and the control groups after treatment are close to zero and generally not significant. However, there seems to be a slight difference before the experiment: employment rate is on average higher in the control group. This is puzzling since also here treatment status is randomly assigned. We examined possible reasons for the difference by checking the differences in pre-treatment employment rates by employment office searching for evidence for failures in randomization process but found no clear signs of any systematic differences. Our preliminary conclusion is therefore that also these pre-treatment differences are due to random variation.

Figure 2 Employment effects of job search training 1999



experimental and non-experimental estimates, we will not explore these differences across different groups any further in this paper.

4.2. Comparison of experimental and non-experimental estimates

In most cases the effects of labor market programs are evaluated based on non-experimental methods comparing employment rates between program participants and suitably chosen comparison group. In a typical approach the pre-treatment differences between the participants and the comrarison are controlled for using regression or matching methods. If all relevant differences can be controlled for the non-experimental procedure should produce similar results than randomized experiment. However, even with very rich data, some differences between the participants and non-participants may remain and cause bias in the estimates. Particularly differences in factors such as job-search intensity or motivation are hard to control for.

To compare the results of the experiment to typical non-experimental estimates we discarded the experimental control group and replaced it with a random sample of persons who were unemployed at the time of the experiment. We used propensity score matching for balancing the observable characteristics of the treatment and the comparison group and calculated the effects of program participation using nearest-neighbor estimators.

Our outcome measures are monthly employment rates during the four-year period after treatment. We calculate the effects of the program on employment rate at four discrete points: 6, 12, 24 and 48 months after treatment. In addition we estimate the effect of treatment on the total months in employment during the four-year period.

The matching estimates for the Työhön-program are reported in Table 3a and the results for the 1999 Job search training experiment in Table 3b. For each outcome we first report the raw differences in employment rates between the treatment and the comparison group and then matching results from four different specifications with increasing number of covariates. Below the matching estimates we report the experimental estimate for comparison. First we use just demographic characteristics (sex, age, marital status and number and ages of children). Then we add information on broad occupation (10 categories) and education (4 categories). Next we also control for regional differences and seasonal variation by adding information on regional unemployment rate, whether the person lived in an urban area and the indicators of the month when the randomization took place. Finally we also use information on the

unemployment history (duration of the on-going unemployment spell and total months in unemployment during past three years) and the type unemployment benefits received at the time of randomization (earnings-related unemployment insurance, flat rate unemployment assistance or labor market subsidy). Average values of all these covariates in treatment and matched comparison groups are reported in the appendix. General conclusion based on these averages is that matching successfully balances the distributions of the covariates. Of the 41 covariates included in the model there is a statistically significant difference between the treatment and the comparison group in only one case: treatment group members are less likely to have started the program in October.

The results in Table 3a are not easily summarized. However, some general conclusions can be made. First, the difference between the treatment group and the unmatched control group is small shortly after training but grows over time. The difference in employment rates is 3.9% 12 months, 13.6% 24 months and 16.0 48 months after the program. Second, these differences decline when more covariates are added but not always in a monotonous fashion. Third, even after controlling for a large number of covariates there tends to be a substantial difference between the treatment and the matched comparison group. If the success of the program is evaluated based on the total months in employment during the four year period, the matching estimators indicate significant positive employment effects while the experimental estimate is very close to zero.

Table 3a Matching estimates for employment effects in Työhön-program

Employed 6 months after program				
	Treatment	Comparison	Difference	T-value
No controls	46.5 %	48.1 %	-1.6 %	-0.79
Demographic differences	46.5 %	41.6 %	4.8 %	0.61
+ education and occupation	46.5 %	49.5 %	-3.1 %	-0.76
+ time and region	46.5 %	43.4 %	3.1 %	1.03
+ unemployment history	46.5 %	45.8 %	0.6 %	0.22
Experimental estimate	46.4 %	45.0 %	1.4 %	0.5

Employed 12 months after program				
	Treatment	Comparison	Difference	T-value
No controls	54.0 %	50.2 %	3.9 %	1.91
Demographic differences	54.0 %	51.6 %	2.4 %	0.29
+ education and occupation	54.0 %	51.6 %	2.4 %	0.6
+ time and region	54.0 %	50.3 %	3.7 %	1.24
+ unemployment history	54.0 %	49.8 %	4.2 %	1.43
Experimental estimate	53.8 %	51.9 %	1.9 %	0.69

Employed 24 months after program				
	Treatment	Comparison	Difference	T-value
No controls	64.2 %	50.6 %	13.6 %	6.73
Demographic differences	64.2 %	46.3 %	17.9 %	2.16
+ education and occupation	64.2 %	55.2 %	9.0 %	2.25
+ time and region	64.2 %	55.0 %	9.2 %	3.13
+ unemployment history	64.2 %	58.9 %	5.3 %	1.86
Experimental estimate	64.2 %	60.8 %	3.4 %	1.24

Employed 48 months after program				
	Treatment	Comparison	Difference	T-value
No controls	69.7 %	53.7 %	16.0 %	7.93
Demographic differences	69.7 %	65.3 %	4.4 %	0.53
+ education and occupation	69.7 %	60.0 %	9.7 %	2.44
+ time and region	69.7 %	58.4 %	11.3 %	3.94
+ unemployment history	69.7 %	59.8 %	9.8 %	3.51
Experimental estimate	70.2 %	68.8 %	1.4 %	0.55

Total months in employment during 4 years after experiment				
	Treatment	Comparison	Difference	T-value
No controls	28.3	24.3	4.0	5.69
Demographic differences	28.3	26.5	1.8	0.68
+ education and occupation	28.3	26.3	2.0	1.61
+ time and region	28.3	25.2	3.1	3.30
+ unemployment history	28.3	26.1	2.2	2.45
Experimental estimate	28.3	28.2	0.1	0.10

In Table 3b, we do similar comparison using data from the 1999 job search training programs. The most striking difference is that the employment outcomes of the treatment group are much worse. Only slightly over a third of the treatment group members are employed six months after the program and even after two years the fraction is only about a half. Still the matching procedure removes the differences between the treatment and the comparison groups so that the difference in employment rates between the treatment and the comparison groups are usually not significant. Since the experimental estimate is also zero, there is no evidence on systematic bias in the matching estimates.

There are several potential explanations for the differences in the results between the two experiments. For example, the raw difference between the treatment and unmatched comparison groups is much larger in the second experiments which may be reflected also in the differences between the matched groups. Our favourite explanation still is that it is more difficult to control for the unobserved differences between the treatment and the comparison groups when participation is entirely voluntary. In the second experiment the caseworkers selected the program participants and unobserved differences between the treatment and the comparison groups are smaller, because these differences are likely be unobserved by the caseworkers.

Table 3b Matching estimates for job search training program 1999

Employed 6 months after program				
	Treatment	Comparison	Difference	T-value
No controls	36.8 %	57.6 %	-20.8 %	-11.01
Demographic differences	36.7 %	37.7 %	-1.0 %	-0.13
+ education and occupation	36.7 %	44.2 %	-7.5 %	-2.08
+ time and region	36.8 %	36.5 %	0.3 %	0.10
+ unemployment history	36.6 %	38.0 %	-1.5 %	-0.51
Experimental estimate	36.5 %	38.6 %	-2.1 %	0.65

Employed 12 months after program				
	Treatment	Comparison	Difference	T-value
No controls	42.0 %	58.7 %	-16.7 %	-8.84
Demographic differences	41.9 %	51.3 %	-9.4 %	-1.23
+ education and occupation	41.9 %	40.6 %	1.3 %	0.36
+ time and region	42.0 %	34.6 %	7.4 %	2.58
+ unemployment history	41.9 %	37.7 %	4.2 %	1.48
Experimental estimate	41.3 %	40.4 %	0.9 %	0.26

Employed 24 months after program				
	Treatment	Comparison	Difference	T-value
No controls	51.1 %	59.9 %	-8.8 %	-4.69
Demographic differences	51.2 %	48.1 %	3.0 %	0.40
+ education and occupation	51.2 %	45.2 %	5.9 %	1.62
+ time and region	51.1 %	45.5 %	5.6 %	1.93
+ unemployment history	51.1 %	47.5 %	3.6 %	1.24
Experimental estimate	50.8 %	50.7 %	0.1 %	0.02

Employed 48 months after program				
	Treatment	Comparison	Difference	T-value
No controls	46.3 %	56.6 %	-10.3 %	-5.42
Demographic differences	46.2 %	52.7 %	-6.5 %	-0.87
+ education and occupation	46.2 %	51.2 %	-4.9 %	-1.34
+ time and region	46.3 %	42.7 %	3.6 %	1.24
+ unemployment history	46.4 %	41.7 %	4.8 %	1.65
Experimental estimate	47.7 %	46.0 %	1.7 %	0.52

Total months in employment during 4 years after experiment				
	Treatment	Comparison	Difference	T-value
No controls	21.2	27.5	-6.4	-9.32
Demographic differences	21.2	23.1	-2.0	-0.78
+ education and occupation	21.2	22.2	-1.1	-0.90
+ time and region	21.2	19.1	2.1	2.25
+ unemployment history	21.2	20.3	0.8	0.93
Experimental estimate	21.5	21.6	-0.05	0.10

5 Conclusion

Randomized experiments are often viewed as gold-standard of program evaluation. In addition to providing reliable estimates of the program effects they provide a benchmark against which to evaluate non-experimental evaluation methods. In this paper we used data from two different experiments with different selection processes to the program. Also the results are somewhat different. Our conclusion is that simple matching methods like the one used in this paper yield reasonably good estimates when the selection to the program is based on “objective” criteria as might be the case when caseworkers select participants based on information that they provide in program application. In contrast, even a large number of covariates does not allow getting rid of selectivity bias when program participation is entirely voluntary and largely determined by unobserved motivational or other characteristics of the applicants.

Our study also paints a relatively pessimistic view of the benefits of job search training programs. This should not lead to a conclusion that these programs should be abolished. Previous research has indicated that these programs may have other benefits related especially related to mental health outcomes. The programs are also shown to enhance re-employment rates in specific high risk groups. Therefore, the appropriate policy conclusion might not be abolishing these programs but rather more careful targeting to groups that are most likely to benefit from job search training.

References

Heckman, J., Ichimura, H., Smith. J.; ja Todd, P. (1999): “Characterizing selection bias using experimental data”, *Econometrica* 66, 1017-1098

LaLonde, R. (1986): “Evaluating the Econometric Evaluations of Training Programs with Experimental Data”, *American Economic Review* 76, 604-620

Malmberg-Heimonen, I. and Vuori, J. (2005): “Activation or discouragement - the effect of enforced participation on the success of job-search training”, *European Journal of Social Work* 8, No. 4, December 2005, pp. 451 - 467.

Silvonen, J. and Vuori, J. (2005): “The benefits of a preventive job search program on re-employment and mental health at 2-year follow-up”, *Journal of Occupational and Organizational Psychology* 78, 43–52

Vuori, J., Silvonen, J., Vinokur, A., & Price, R. (2002): “The Työhön Job Search Program in Finland: Benefits for the unemployed with risk of depression or discouragement?”, *Journal of Occupational Health Psychology* 7, 5-19

Vuori. J. and Vinokur, A. (2005): “ Job-search preparedness as a mediator of the effects of the Työhön Job Search Intervention on re-employment and mental health”, *Journal of Organizational Behavior* 26, 275–291

Appendix 1. Average characteristics of the treatment and matched comparison groups in Työhön program

		Treatment	Comparison	T-statistics for equal means
Male	Unmatched	23.2 %	50.0 %	-13.28
	Matched	23.2 %	23.9 %	-0.27
Age 25 - 34	Unmatched	41.0 %	24.9 %	9.14
	Matched	41.0 %	40.3 %	0.23
Age 35 - 44	Unmatched	31.8 %	23.9 %	4.57
	Matched	31.8 %	29.2 %	0.99
Age 45 – 54	Unmatched	15.3 %	20.7 %	-3.29
	Matched	15.3 %	16.9 %	-0.77
Not married	Unmatched	43.1 %	45.1 %	-0.99
	Matched	43.1 %	40.5 %	0.92
Divorced or widow	Unmatched	14.2 %	14.3 %	-0.09
	Matched	14.2 %	14.4 %	-0.08
Children, age 0-3	Unmatched	19.8 %	11.0 %	6.97
	Matched	19.8 %	20.0 %	-0.07
Children, age 4-6	Unmatched	8.1 %	5.9 %	2.27
	Matched	8.1 %	8.5 %	-0.31
Children, age 7 – 17	Unmatched	21.1 %	15.0 %	4.24
	Matched	21.1 %	20.0 %	0.49
Disability	Unmatched	4.4 %	5.5 %	-1.27
	Matched	4.4 %	2.6 %	1.71
Occupation				
Unknown	Unmatched	3.5 %	8.9 %	-4.68
	Matched	3.5 %	3.9 %	-0.3
Technical	Unmatched	11.1 %	5.9 %	5.4
	Matched	11.1 %	10.0 %	0.65
Health care	Unmatched	9.2 %	6.7 %	2.51
	Matched	9.2 %	7.9 %	0.81
Administrative	Unmatched	18.5 %	9.1 %	8.1
	Matched	18.5 %	20.0 %	-0.65
Trade	Unmatched	6.1 %	5.2 %	1
	Matched	6.1 %	5.8 %	0.24
Industrial	Unmatched	6.5 %	16.6 %	-6.76
	Matched	6.5 %	7.3 %	-0.56

Services	Unmatched	10.2 %	8.8 %	1.16
	Matched	10.2 %	11.1 %	-0.55
Education (ref. primary)				
Secondary	Unmatched	62.1 %	54.6 %	3.72
	Matched	62.1 %	57.9 %	1.51
Lower tertiary	Unmatched	10.5 %	5.7 %	5.05
	Matched	10.5 %	13.2 %	-1.49
Higher tertiary	Unmatched	4.8 %	2.3 %	4.24
	Matched	4.8 %	4.4 %	0.41
Month of participation				
March	Unmatched	7.7 %	9.3 %	-1.31
	Matched	7.7 %	7.1 %	0.43
April	Unmatched	11.1 %	9.0 %	1.86
	Matched	11.1 %	10.0 %	0.65
May	Unmatched	7.6 %	9.1 %	-1.32
	Matched	7.6 %	8.1 %	-0.32
June	Unmatched	1.6 %	9.2 %	-6.53
	Matched	1.6 %	1.5 %	0.23
July	Unmatched	8.4 %	8.9 %	-0.46
	Matched	8.4 %	7.9 %	0.31
August	Unmatched	11.1 %	9.0 %	1.86
	Matched	11.1 %	10.3 %	0.46
September	Unmatched	10.5 %	9.0 %	1.25
	Matched	10.5 %	10.8 %	-0.18
October	Unmatched	18.4 %	9.1 %	7.96
	Matched	18.4 %	23.5 %	-2.24
November	Unmatched	11.5 %	9.2 %	1.92
	Matched	11.5 %	11.5 %	0
Lives in a city	Unmatched	80.5 %	59.4 %	10.66
	Matched	80.5 %	79.7 %	0.36
Lives in urban area	Unmatched	6.6 %	15.5 %	-6.12
	Matched	6.6 %	6.8 %	-0.11
Regional unemployment %	Unmatched	17.9	20.6	-15.64
	Matched	17.9	17.9	0
Graduated in 1995	Unmatched	4.2 %	5.2 %	-1.11
	Matched	4.2 %	2.7 %	1.4
Received home care	Unmatched	9.4 %	5.1 %	4.83
allowance in 1995	Matched	9.4 %	8.9 %	0.3

Unemployed at the time of randomization	Unmatched	73.5 %	52.3 %	10.55
	Matched	73.5 %	74.7 %	-0.45
Months unemployed in past 3 years	Unmatched	11.8	15.0	-6.64
	Matched	11.8	11.4	0.6
Duration of current unemployment, months	Unmatched	3.7	6.6	-6.87
	Matched	3.7	4.1	-1.37
Received labor market support in 1995	Unmatched	13.2 %	23.5 %	-6
	Matched	13.2 %	11.8 %	0.77
Received earnings related UI benefits in 1995	Unmatched	41.9 %	41.4 %	0.26
	Matched	41.9 %	41.9 %	0
Received flat rate UA benefits in 1995	Unmatched	11.5 %	17.7 %	-4.05
	Matched	11.5 %	9.5 %	1.11
Does not seek full-time work	Unmatched	12.4 %	10.8 %	1.26
	Matched	12.4 %	13.2 %	-0.42
N		620	620	