WORKING PAPER 2020:21

Labor market integration of low-educated refugees

RCT evidence from an ambitious integration program in Sweden

Matz Dahlberg
Johan Egebark
Ulrika Vikman
Gülay Özcan

IFAU
INSTITUTE FOR EVALUATION OF LABOUR MARKET AND EDUCATION POLICY

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala.

IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

More information about IFAU and the institute's publications can be found on the website www.ifau.se

ISSN 1651-1166

Labor market integration of low-educated refugees

RCT evidence from an ambitious integration program in Sweden [a]

by

Matz Dahlberg[b] and Johan Egebark[c] and Gülay Özcan[d] and Ulrika Vikman[e]

October 22, 2020

Abstract

This paper evaluates an ambitious and newly designed program for increased integration in Sweden. The purpose of the program is to help newly arrived, low-educated refugees into employment. The program includes four main components: (1) intensive initial language training, (2) work practice under close supervision, (3) job search assistance, and (4) extended cooperation between the local public sector and firms. An important feature of the program is that the demand side of the labor market, represented by the largest real estate company in Gothenburg, is involved in designing the program. Our evaluation is based on a randomized controlled trial, where potential participants in one of the first waves were randomly assigned to treatment and control groups. The paper presents results from the first two years after randomization. Using inference based on Fisher's exact test, we show that the program has positive effects on employment: around 30 % of the individuals in the treatment group are employed each month during the first year following the end of the program, compared to an average of approximately 15 % in the control group.

Keywords: Refugee immigration; Integration; Randomized experiment; Labor market program
JEL-codes: C93, J08, J15, J23, J61

[a] We thank Matti Sarvimäki, Johan Vikström, and seminar participants at the EALE SOLE AASLE World Conference 2020, IFAU, and UIL Workshop in Uppsala for valuable comments and suggestions. We also thank Ann-Sofie Arvidsson and Asllan Hajdari for helping out implementing the randomization, as well as Sükran Dogan, Peter de Jonge and Ulf Wallin for sharing valuable insights about the program in Gothenburg and for organizing visits on location in Gothenburg. AEA RCT ID: AEARCTR-0004788.
[b] Uppsala University and IFAU. matz.dahlberg@ibf.uu.se.
[c] Swedish PES and IFN. johan.egebark@ifn.se.
[d] Swedish PES. gulay.ozcan@arbetsformedlingen.se.
[e] IFAU. ulrika.vikman@ifau.uu.se.

Table of contents

1 Introduction

It is well-established that labor market integration takes time for immigrants, something that is particularly true when it comes to refugees (see, for instance, Cortes, 2004, and Ruiz and Vargas-Silva, 2018). In recent years, large inflows of refugees to many Western (European) countries has spurred a substantial and frequently intense and heated debate on how to best integrate immigrants in their new countries (see, for instance, the discussion in Hangartner and Sarvimäki, 2017). As a consequence, finding successful integration policies is at the top of the political agenda, as getting this group of immigrants well-integrated is considered crucial from the perspective of both the individuals and the host countries (see, for instance, Dustmann et al., 2017, Hangartner and Sarvimäki, 2017, and Battisti et al., 2019).

In this paper, we use a randomized experiment to evaluate an ambitious integration program developed and implemented in the city of Gothenburg in Sweden. The program rests on four main building blocks: (1) intensive language training (including work-specific language courses), (2) work practice with supervisors (1–3 supervisors per participant), (3) job search assistance performed by professional caseworkers at the Swedish Public Employment Service (PES), and (4) extended cooperation between the local public sector and firms. We study the effect on the probability of being employed using a follow-up horizon of two years. Previous evidence suggests that immigrants making an early transition to employment is key for achieving positive outcomes in the long term.[1] Hence, studying the short-run effects of the program is highly motivated from a policy perspective.

The program was designed as a direct reaction to the poor labor market situation for newly arrived refugees in the Gothenburg region, where unemployment levels for disadvantaged workers in general, and low-skilled refugees in particular, are high. The program came into effect as a result of a local initiative involving three separate actors: the PES in Gothenburg, the City of Gothenburg, and the largest real estate company in Gothenburg, AB Framtiden.[2] One important feature here is that the demand side of the labor market, represented by the real estate company and its subsidiaries, has been deeply involved in designing the program.

The eligibility criterion used in the program is very strict: only newly arrived, low-educated refugees are allowed to participate.[3] Our evaluation is based on a randomized controlled trial (RCT), where 140 potential participants in one of the first waves were randomly assigned to treatment and control groups.[4] While the individuals in the treatment group were invited to participate in the program, the control group joined the baseline services at the PES. Most of the activities offered in the program are similar to those found in the baseline services, such as language train-

[1] See, for example, Fasani et al. (2020) who find large negative medium- to long-term effects on refugee labor market outcomes of temporary employment bans imposed on recently arrived asylum seekers.

[2] The Swedish Public Employment Service belongs to the central government sector. However, each municipality in Sweden, including Gothenburg, has its own autonomous local PES offices. The City of Gothenburg, which belongs to the local public sector, is responsible for providing education at different levels, including adult education. AB Framtiden, founded in 1915, is one of the largest real estate companies in Sweden. The company is owned by the municipality but its organizational structure corresponds to that of a private firm.

[3] All refugees in our target group have recently received a Swedish residence permit and an establishment plan, i.e. a plan for establishment in the new society.

[4] The 140 individuals constituted the universe of eligible individuals in Gothenburg at the time of the start of the program.

ing, work practice, and job search assistance. The main difference is instead found in the level of intensity of the activities during the year that the program takes place. Our comparisons show substantial differences in treatment intensity between treated and non-treated individuals throughout the entire program. Given the high level of ambition, we consider the program providing "a most likely case for a positive labor market outcome." If we do not see a positive effect here, we could start to worry that no policy will have a positive effect on the (short-run) labor market integration of newly arrived low-skilled refugees.

We present compelling evidence of the program having positive effects on employment. Around 30% of the individuals in the treatment group are employed each month during the second year after the start (i.e., during the first year after the end of the program), compared to an average of approximately 15% in the control group. Inference based on Fisher's exact test shows that the difference in employment between the treated and control groups is most likely not given by chance. We also follow up on the individuals in the control group and show that they to a larger extent participate in job creation schemes.

We add to the existing literature in three ways. First, we study effects for a specific and understudied subgroup of immigrants having become increasingly relevant for policymakers: *low-skilled newly arrived refugees*. This group comprises the vast majority of the influx of immigrants to Western countries in the past years, and many host countries struggle to find ways of integrating this group due to its weak initial position in the labor market. The only other study we know of with a specific focus on refugees (and which uses a credible research design) is Battisti et al. (2019), who perform an RCT where the treatment group received job search support from an NGO. For the treatment group, the NGO identifies potentially suitable employers and, upon the agreement of the job seeker, sends a CV to the employers. The study measures outcomes using three surveys: a baseline survey, a follow-up survey after six months, and a second follow-up survey after twelve months. The authors find no effect on employment after 6 months, but weak evidence of a positive effect after 12 months.

Second, we estimate the *causal effects* of an integration program. Evaluations of integration policies often encounter difficulties due to the selection of individuals into programs. For example, many studies show important associations between language skills and employment outcomes,[5] but only a few utilize exogenous variation. Exceptions include Lochmann et al. (2019), Sarvimäki and Hämäläinen (2016) and Arendt et al. (2020). Lochmann et al. (2019) study the effects of language training in France and find positive effects on labor force participation, where the discontinuity originates from test scores on an initial language exam. They find larger effects for highly educated immigrants but no differential effect by type of migration, gender, or age. Using a discontinuity originating from a reform, Sarvimäki and Hämäläinen (2016) find increased earnings for unemployed immigrants receiving more language training. Arendt et al. (2020) use exogenous variation created by a Danish reform to study effects of improved and more intensive language training. While their point estimates indicate positive effects on earnings over an eighteen year period for treated individuals, their estimated standard errors are relatively large. Our study manages selection issues by using a randomized control trial.

Third, in contrast to previous studies on integration policies, we consider a program that in-

[5] See, for example, Chiswick (1991), Berman et al. (2003), and Dustmann and Soest (2001).

cludes skill acquisition (language skills, work skills, and social skills) in combination with addressing labor market frictions through professional job search assistance and through close cooperation with firms. Previous works have focused on either language training (Lochmann et al., 2019 and Sarvimäki and Hämäläinen, 2016), intensive coaching (Andersson Joona and Nekby, 2012 and Åslund and Johansson, 2011), contacts with employers (Battisti et al., 2019), or wage subsidies (Butschek and Walter, 2014 and Clausen et al., 2009). Given that immigrants in general, and refugees in particular, face many different obstacles in the labor market, it is justified studying the effects of programs that combine several key elements.

The rest of the paper is organized as follows. In the next section, we describe the integration program and examine to what extent the objectives of the program show up in the data. In section 3, we present the experimental design and the estimation method, and in section 4 we present the data and provide some summary statistics. Section 5 presents the results and section 6 concludes the paper.

2 The integration program

The integration program *Välkommen till Framtiden* was launched in 2016 when the first cohort of newly arrived refugees was enrolled. In total, four waves have started the program, one per year during 2016–20. In this paper, we follow the second wave, which started in 2017. The overall aim of the program is that all participants should learn a profession and have a full-time job or to start regular education after the end of the establishment period, which in Sweden is said to last for two years.[6] All participants are expected to participate in the program full time.

Figure 1 illustrates the different stages of the program. The randomization of potential participants was conducted in late April 2017. All individuals who were randomized to belong to the treatment group were summoned to an information meeting, which was held at the local PES office in Gothenburg at the end of April 2017 (less than one week after the randomization).[7] In the analysis, we denote April 2017 as month zero. At the initial information meeting, the individuals in the treatment group received an offer to participate in the program. Those who accepted the offer were summoned to a second meeting, which was held at the end of May 2017, just before the program activities started. In the analysis, we denote May 2017 as the first month of the program, since the individuals in the treatment group may react already based on the information about the program.

The program activities started in June 2017 when the participants started their initial language training. The language training lasted for three months, months 2–4 after the randomization (June to August 2017). The initial language training was followed by six months of supervised work practice, months 5–10 after randomization (from September 2017 to February 2018). In the final stage, the participants were offered job search assistance as they finished the work practice and

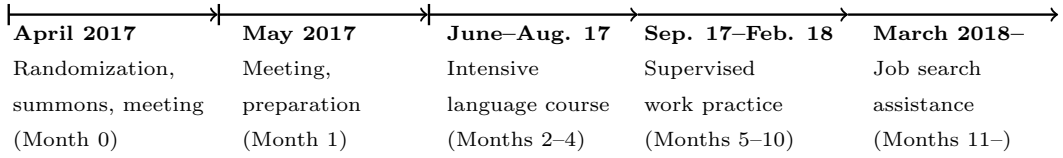
[6] All newly arrived refugees having received a residence permit and registering at the Swedish Public Employment Service are automatically enrolled in a general establishment program. Individuals can stay within the program for a maximum of two years. During the two-year establishment period, individuals can participate in many different activities, including language training, work practice, preparatory education, labor market training programs, etc. The local initiative *Välkommen till Framtiden* we evaluate in this paper took place during this two-year period.

[7] In order to manage language differences between the refugees, the PES set up two separate information meetings. Each meeting had relevant interpreters present.

were ready to take a job. The job search assistance typically started in March 2018, 11 months after randomization.[8]

Next, we describe the program in a little bit more detail, show how the objectives of the program match actual realizations as observed in the data, and present information on what the control group does during the program period.

Figure 1: Illustration of the different stages of the experiment/program.



Note: The randomization was conducted on April 22, 2017. The first information meetings were held on April 26 and April 27. The second information meeting, for those who had accepted the offer to participate in the program, was held on May 22. The activities started on May 29, 2017. Job search assistance could start before March 2018.

2.1 The components of the program

2.1.1 Intensive language training

In Sweden, all newly arrived refugees who register as unemployed at the Swedish PES are offered Swedish language classes, so-called Swedish for Immigrants (SFI).[9] This typically amounts to 15 hours of teaching per week throughout the two-year establishment period (in addition to this, individuals are assumed to study and practice Swedish on their own).

One of the aims of the Gothenburg integration program is to provide more intensive language training in the initial phase of the establishment period. Hence, all participants are obliged to participate in language classes (close to) full time (40 hours a week) during the first three months of the program. Individuals are assigned to different skill levels based on their initial knowledge of Swedish. However, since all language classes take place simultaneously, all participants are able to meet each other at these sessions. Those who follow the program combine the basic SFI language courses with higher level work-specific language training. This is important since they must be ready to start an internship within three months where they practice a real job at a real workplace.

In addition to the initial SFI language training, the program includes other, more general courses, including a course introducing the participants to Swedish society,[10] one course teaching workplace rules,[11] one course teaching how to be service-minded, and introductory courses in

[8] Earlier assistance was given to participants who were ready to start employment before March 2018.

[9] Swedish for Immigrants (SFI) is provided free of charge by the local government sector. Anyone older than 16 can participate. There are three different tracks: track 1, which is offered to illiterates, includes four levels of courses (A–D), track 2 includes three levels (B–D), and track 3, which is typically offered to those with some higher education, includes two levels (C and D).

[10] This is in addition to a 60-hour basic Swedish society course taken by everyone in the establishment program. These courses concern practical matters, including how to apply for a rental apartment, how to pay bills, etc. The extra introduction given to the treatment group adds more detailed information.

[11] This course informs the participants what is expected of them as employees, such as being on time and reporting

mathematics and IT. All extra courses are given in Swedish and they run during the first 8 months (i.e., until December 2017).

After the initial months of the program, the participants enter work practice. This means that the language training continues at a slightly slower pace, as the participants take Swedish classes two days per week (roughly 15 hours). During the work practice, the participants are expected to continue SFI teaching until they reach the required levels for being able to enter the regular Swedish education system (levels C or D). In order to get additional job-specific language training, the participants are encouraged to write down words and expressions they encounter during their work practice, which are then discussed during the next language class.

As described above, the integration program aimed to intensify the language training during the early phase of the establishment period. However, it is not necessarily true that the participants received more SFI training overall. Since the program focuses on other activities at a later stage (i.e., work practice and job search), the control group should in theory have time to catch up. Hence, we may expect that the only thing that is different between the two groups is the distribution of teaching hours, with more intensive training in the beginning for those in the treatment group compared to those in the control group.

Figures 2–4 compare treated and non-treated individuals in our RCT in terms of language training.[12] Starting with Figure 2, we draw two conclusions. First, it is clear that the treatment group had more Swedish training than the control group during months 2–4 of the program (roughly 40 more hours per month from June to August).[13] Second, there is no evidence that the control group catches up later on, at least not during the first year. From month five of the program, the treated and control groups had roughly the same number of hours of language training per month, corresponding to 15 hours per week. This indicates that the participants received more Swedish training overall.[14]

Figure 3 adds additional information by showing the share of individuals with a grade from the SFI language courses. The question we ask here is to what extent attending extra classes raises the participants' language skills. The share with a grade increases sharply in the treatment group relative to the control group just after the end of the intensive initial phase. We take this as evidence of a positive effect of the program on the level of Swedish language skills.

Figure 4 looks at differences with respect to extra courses. The figure confirms that the treatment group participated in other, more general courses as well and that these courses were offered during almost the entire first year of the program. Since the courses are taught in Swedish, they also add language skills.

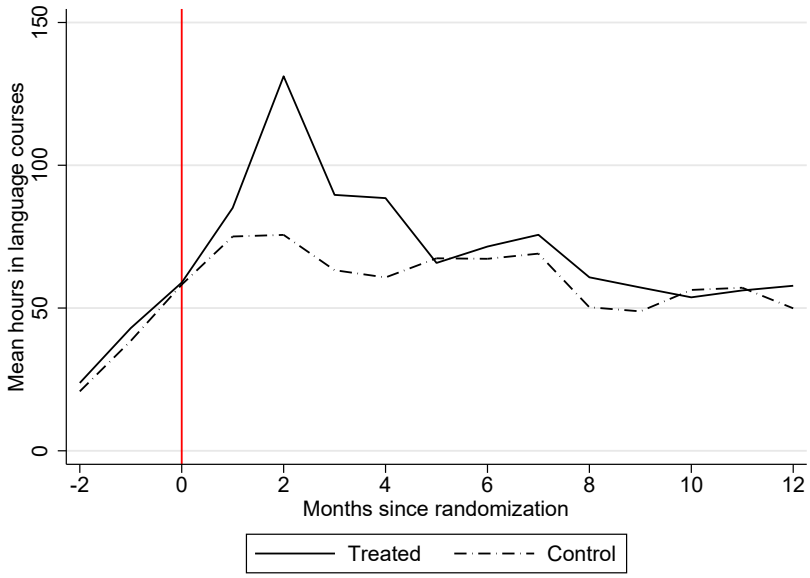
sick leave to the employer. It also considers more general matters, such as how to get unemployment insurance in Sweden.

[12] Here, we consider the second wave of participants, since we have a control group that enables comparisons. It is likely that the sharp contrasts between the program and the baseline PES services also apply to the other three waves.

[13] Individuals in the control group seem to have roughly 50 hours of language training per month, or about 15 hours per week. This is the amount of language training individuals normally get in the basic service at the PES. The number of hours for the treatment group is lower than the 160 hours per month (40 hours per week) we would expect. However, since we use intention-to-treat (ITT) for these comparisons and since not all individuals who were assigned to the treatment group receive treatment, our measures constitute lower bounds.

[14] We have confirmed that this conclusion holds if we consider the second year as well.

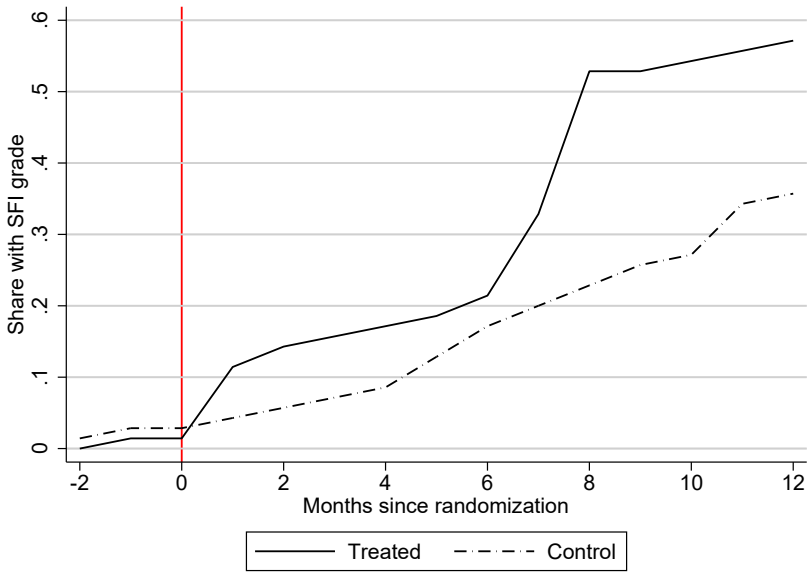
Figure 2: Participating in Swedish language courses.



Notes: The figure shows the extent to which treated and control individuals participate in Swedish language courses during the first year of the program.

Source: Own calculations based on data obtained from the Educational Unit at the City of Gothenburg.

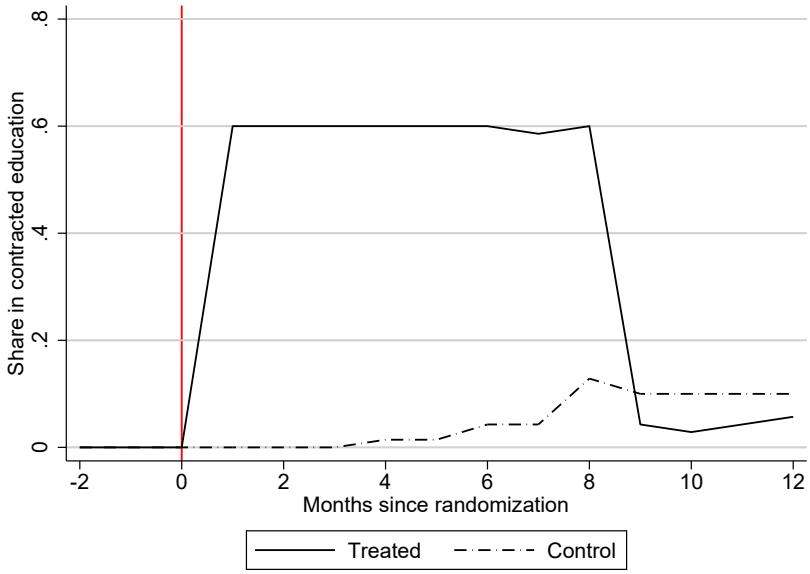
Figure 3: Share with a grade from Swedish language courses.



Notes: The figure shows the extent to which treated and control individuals have a grade from SFI.

Source: Own calculations based on data obtained from the Educational Unit at the City of Gothenburg.

Figure 4: Participating in extra courses.



Notes: The figure shows the extent to which treated and non-treated individuals participate in extra courses offered through the municipality's adult education during the first year of the program. Individuals in the treatment group were registered on May 29, which is why the share is high already in May.

Source: Own calculations based on data obtained from the Educational Unit at the City of Gothenburg.

2.1.2 Intensive work practice

After the initial three months of language training, the participants continue to the next level of the program as they start work practice within the real estate company AB Framtiden in Gothenburg.[15] The work practice proceeds three days a week for at least six months. The participants in the wave we study could choose between three different job tracks when entering the program: indoor cleaning, working with outdoor environments (gardening, snow shoveling, maintenance, etc.), and as real estate caretakers.[16] During their work practice, the participants attend the same work-related courses as all new employees (e.g., to get certificates to handle machines and tools).

Each participant is appointed 1–3 supervisors, who have the main responsibility for the intern during the time of the work practice. All supervisors are hand-picked from the regular staff – typically one of the more senior employees – at the workplace where the work practice occurs. The supervisors are responsible for introducing the participants to the specific work tasks and to make sure that they gain the skills needed to work independently. There are two main reasons for appointing more than one supervisor per individual. First, this makes the participant less dependent on one person and, second, it makes it less burdensome for the supervisor. The supervisors were paid SEK 50 per day (roughly 6 US dollars) for the added responsibility.[17]

[15] The company has four subsidiaries responsible for managing existing properties, and two subsidiaries responsible for building new properties. The work practice typically takes place at the four subsidiaries responsible for existing properties.

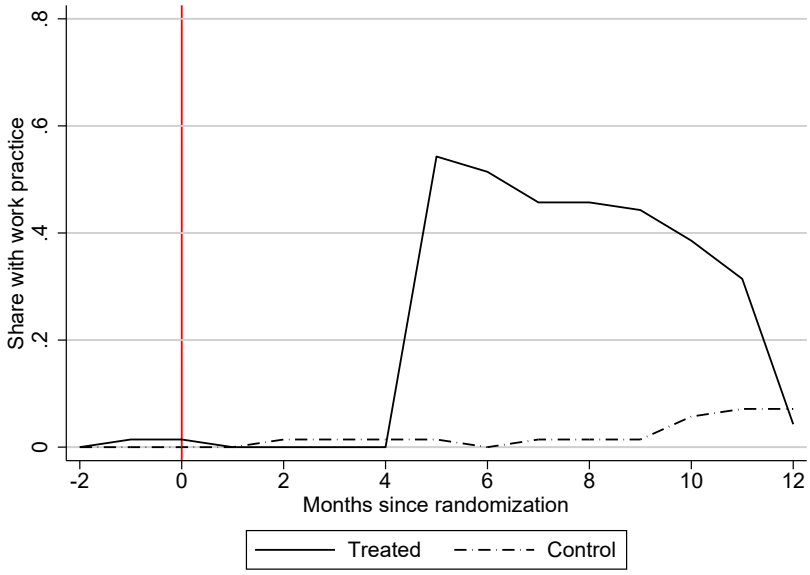
[16] The treatment group was informed about these job tracks at the first information meeting.

[17] According to the real estate company, the supervisors' workload increases at the beginning of the work practice

When allocating the participants to the work practice at the real estate company, two aspects are taken into account. First, in order to challenge the participants to practice Swedish, they are assigned supervisors who do not speak their native language. Second, to get the participants used to commuting, they are assigned a workplace outside their own residential area. In addition, the regular staff are explicitly asked to include the participants in the social life at the workplace so that they can start creating a network in Sweden.

Since our data enables us to observe all the activities that the individuals participate in (at any given point in time), we can capture any differences in terms of work practice between the treatment and the control group. Figure 5 shows a clear difference in participation rates between treated and non-treated individuals: while there is almost no (or at least very little) supervised work practice in the control group, the treatment group exhibits a large increase during months five to ten of the program. The timing of the sharp increase is expected, given that the initial full-time language training ends four months after the randomization (see Figure 1). We also note that the difference between treated and non-treated individuals drops to zero 12 months after randomization, which is also expected given that the work practice is not supposed to last longer.[18]

Figure 5: Participating in work practice.



Notes: The figure shows the extent to which treated and control individuals participate in the activity denoted ”work practice with supervisor” (measured by PES activity code 87).

Source: Own calculations based on data obtained from the Swedish Public Employment Service.

period when they introduce the participants to their new tasks in addition to performing their own usual work. However, after about two months, the general workload reduces as the interns gain skills. This information was given to us in November 2019 when we met with representatives from AB Framtiden, including one of the supervisors, at one of the workplaces involved in the program.

[18]The activity we examine in Figure 3 is called ”work practice with supervisor” (measured by PES activity code 87). The reason why the share of treated individuals involved in this activity is less than 100% is that not everyone assigned to the treatment participated in the program.

2.1.3 Job search assistance

The participants are informed that they will get professional help finding a job if they pass the initial two stages. Hence, there is a clear incentive to stay in the program in the longer run. After the work practice ends, efforts begin in terms of finding potential employers or, alternatively, help the participants get further regular education. The job search assistance in this stage is performed by professional caseworkers at the PES in Gothenburg. The caseworkers also receive help from representatives from the real estate company, who contribute by providing insights on local labor demand. In practice, the job search assistance is about helping the participants find suitable vacant jobs matching their level of skills, helping with applications, and offering support in terms of interview training. The participants in the wave we follow were matched to several various firms, including subcontractors to the real estate company.[19]

2.1.4 Organizational design

A final important component is a well-functioning collaboration between firms and the local government sector. Arguably, language training, work practice, and job search assistance are not new measures in themselves – they have been used in many other countries and are also available to the individuals in our control group. What we consider new in this context, besides the high treatment intensity, is that the program manages to reduce negative lock-in effects by substantially shortening lead times. Since the participants continuously move to the next level through the program, they immediately make use of their recently acquired human capital gains (language skills, work-task skills, social networking, etc.). The time gains are accomplished as a result of the tight collaboration between the three responsible program owners. Representatives from the three program owners have follow-up meetings on a regular basis to discuss the participants' progress, jointly define individual-specific actions, and decide on any necessary changes due to new information.

2.2 What does the control group do?

In the previous section, we get a good grasp of what the individuals in the treatment group do during the first year after randomization. Evidently, they seem to closely follow the intentions of the new program as they take intensive language classes followed by supervised work practice. While we have presented a detailed description of the integration program, we know less concerning what non-treated individuals do during the first year after the randomization.

We have already shown that the control group gets 15 hours of language training per week throughout the first year, corresponding to the minimum level in the baseline services at the PES. We have also shown that virtually no one in the control group participates in the type of work practice (with a supervisor) offered to the treatment group. However, since there are many other various activities in the baseline service that an individual can participate in, we would like to get a more comprehensive picture.

We first rule out that the control group is exposed to the same level of program intensity as the treatment group. To this end, we use the PES data and identify all activities that may be defined as

[19] Roughly 11 percent of the individuals in our treatment group were hired by subcontractors. In a few cases, the participants were hired directly by the real estate company where they had done their work practice.

intensive. We define intensive activities as activities where individuals are at a workplace or where they get full-time training. This includes different types of work practice and all the different labor market training programs administered by the PES.[20]

The first row of Table 1 shows that only a small fraction of the control group participates in intensive activities during the first year, which is in sharp contrast to what we observe in the treatment group. Overall, these findings indicate that the control group to a large extent ends up in low-intensive activities after the randomization. The final rows in Table 1 confirm that this is the case.[21]

Table 1: Participation in different activities for non-treated individuals (percent).

	(1) 3 months	(2) 6 months	(3) 9 months
High-intensive activities			
Training/work practice	1.43	2.86	2.86
Low-intensive activities			
Own job search	54.29	58.57	51.43
Preparatory activities	38.57	21.43	30.00
Other	5.71	17.14	15.71

Notes: High-intensive activities include work practice (measured by PES activity codes 54 and 87) and labor market training programs (code 81). The low-intensive category includes job seekers responsible for their own job search (codes 11, 14, 97, 98, or de-registered) and job seekers participating in preparatory activities (code 80, 83, and 68). The category of "Other" includes our outcome variables defined in Section 5.

2.3 Brief summary of the program under study

The data shows clear-cut differences between the new integration program and the baseline services at the PES. The individuals in the treatment group (i) get more intensive language training (including work-related language training) during the initial months, (ii) participate in work practice with help from supervisors (including social networking) during the first year, and (iii) get intensified job search assistance when they are ready to be matched to an employer. What we want to examine in this paper is whether the high level of ambition in this program can help individuals with very limited ties to the labor market – newly arrived low-educated refugees with limited Swedish language skills – find employment.

[20] Work practice is measured by PES activity codes 54 and 87, labor market training programs are captured by PES activity code 81.

[21] The finding that newly arrived immigrants get low-intensive assistance is in line with results in earlier studies (see, e.g., Andersson Joona, 2020).

3 Experimental design and estimation model

3.1 Experimental design

In total, four waves of individuals have embarked on the program, one per year during 2016–20.[22] In each wave, there have been around 50 slots available. Our study follows the wave where treatment was randomized, the one starting in 2017.

To be eligible for the program, an individual had to meet two main criteria. First, only those with less than high school education were eligible. Second, participation required recently having received a Swedish residence permit.[23] Based on these criteria, the PES in Gothenburg identified 140 potential participants in April 2017.

Out of the 140 potential participants, 70 individuals were randomly drawn to the treatment group and 70 were randomly drawn to the control group. Since previous studies show that the effects of active labor market programs (ALMP) differ between men and women (Card et al., 2018), we block-randomized based on gender to get better precision. This makes the share of men and women in the treatment group identical to the shares in the target population.[24] All 70 individuals in the treatment group were summoned to an information meeting regarding the program, which took place at the PES in Gothenburg at the end of April 2017. The 70 individuals who were summoned were first given a description of the program by representatives from the three program owners. After the introduction, all 70 potential participants were offered to participate in the program. 44 individuals accepted the offer and started the program.

When evaluating the program, we estimate intention-to-treat (ITT) effects or the average treatment effect of being offered the possibility to participate in the program. Since the procedure with summons and information meetings is the procedure typically used by the PES, we argue that the ITT analyses produce the most policy-relevant effect estimates. The 70 potential participants who were randomly allocated to the control group participated in the baseline services provided by the PES in Gothenburg.

In this paper, we evaluate the short-run effects of the integration program. Short run here means within two years, and we measure the effects in each month during the two years.[25] The outcome variables are measured via the data held by the PES (see section 4 below).

3.2 Estimation

Since the assignment mechanism for allocating individuals into treatment and control groups in our stratified randomized experiment is known and controlled by us, we can apply Fisher's approach for calculating exact p-values. Basing the inference on exact p-values instead of relying on the

[22] The first wave started in October 2016, the third wave started in April 2018, the fourth wave started in January 2020.

[23] Since the program includes work practice, individuals who were not ready for employment in the short run were excluded from the pool of potential candidates.

[24] Based on their shares in the target population, we block-randomized 27 women and 43 men to the treatment group. The randomization was carried out by researchers at the Swedish Public Employment Service in Stockholm.

[25] We take the month when we randomized potential participants and the individuals in the treatment group were called to the information meeting about the program, April 2017, as the starting point for the evaluation and the end of May 2019 as the endpoint for the two-year evaluation (since the actual program started on May 29, 2017).

asymptotic properties of the estimators is arguably beneficial in our case with an original sample of 140 individuals. For this reason, we will rely on Fisher's exact test in the analysis.

Fisher's sharp null hypothesis

We adopt the null hypothesis that Fischer himself focused on, which is arguably the most obvious sharp hypothesis (Imbens and Rubin (2015), p. 63). Fisher's sharp null hypothesis is given by:[26]

$$H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \qquad \forall i \tag{1}$$

The sharp null of a zero treatment effect for each individual links observed outcomes to all potential outcomes (i.e., under the null, we know both the realized and the non-realized potential outcomes). Given a test statistic, we can thus calculate its exact distribution under the null by using all possible permutations of the randomization assignments. Inference is then based on the estimated distribution of test statistics.

Assignment mechanism

Given the assignment mechanism for a block-randomized experiment with two strata,[27] we have $\binom{87}{43} \times \binom{53}{27}$ possible combinations. Since calculating a statistic for each and every one of all these possible combinations is not manageable, we will make 10,000 random draws from the set of possible combinations and base the inference on the distribution from those draws.

Test statistic

As test statistic, we will use the absolute value of the difference in the average observed outcomes in the two strata with the relative sample sizes in each strata as weights for combining the two differences.[28] Using the assignment mechanism in equation (2) to get at (random draws from) all possible combinations of the randomization assignments under the null hypothesis stated in equation (1), we calculate the exact distribution of our statistic of interest, F^{τ}. By comparing whether the estimated statistic in equation (3), $\hat{\tau}$, is in line with F^{τ}, we can then perform an inference

[26] Compare with the null hypothesis of no average treatment effects: $H_0 : E[Y_i(1)] - E[Y_i(0)] = 0$.

[27] The assignment mechanism for a block-randomized experiment with two strata, m and f, is given by (Imbens and Rubin, 2015, p. 191):

$$Pr(W) = \left(\binom{N(m)}{N_t(m)} \right)^{-1} \left(\binom{N(f)}{N_t(f)} \right)^{-1} for W \in W^{+} \tag{2}$$

where W is a treatment indicator that takes on the value 1 for observations in the treatment group and 0 for observations in the control group, $N(m)$ is the number of males (87), $N(f)$ is the number of females (53), $N_t(m)$ is the number of males assigned to the treatment group (43), $N_t(f)$ is the number of females assigned to the treatment group (27), $W^{+} = W$ such that $\sum_{i:males} W_i = N_t(m)$ and $\sum_{i:females} W_i = N_t(f)$.

[28] There is a large set of possible test statistics that can be used, but according to Imbens and Rubin (2015), section 5.5, the one we use in this paper is the most popular. Formally, it is given by:

$$\hat{\tau} = \frac{N(m)}{N(m)+N(f)} \left(\overline{Y}_t^{obs}(m) - \overline{Y}_c^{obs}(m) \right) + \frac{N(f)}{N(m)+N(f)} \left(\overline{Y}_t^{obs}(f) - \overline{Y}_c^{obs}(f) \right) \tag{3}$$

based on that comparison (if few values in F^{τ} are higher, in absolute terms, than $\hat{\tau}$ then $\hat{\tau}$ can be considered significant).[29]

4 Data and summary statistics

4.1 Data

Our study on labor market effects of the integration program benefits from comprehensive administrative data collected by the Swedish Public Employment Service. To construct different types of outcome measures, we use daily records on each job seeker's unemployment status. This includes information on open unemployment, active labor market programs, all types of subsidized employment, part-time and temporary work, and, finally, regular employment and education. The PES registers also contain information on individual characteristics, such as age, gender, education, country of origin, and date of entry to Sweden (see Table 2 for the variables we have access to).

4.2 Balancing on covariates

Table 2 provides results from balancing tests based on Fischer exact tests. The first and second columns present the mean values and standard deviations for treatment and control groups, respectively. The third column shows differences in mean values across treated and control groups, and the final column provides the exact p-values from the Fischer test (testing the sharp null hypothesis that the mean values for the two groups equal zero). The p-values range from 0.193 to 1, indicating that the randomization managed to balance treated and non-treated individuals based on pre-determined background characteristics.

[29] Formally, the exact p-values are calculated as:

$$p = \frac{1}{J} \sum_{j=1}^{J} 1(|\tilde{\tau}_j| \geq |\hat{\tau}|) \tag{4}$$

Table 2: Mean values and balancing on covariates, p-values from Fischer exact tests (with 10,000 replications).

	(1) Treated	(2) Control	(3) Difference	(4) P-values
Women	0.386	0.371	0.014	1.000
	(0.490)	(0.487)		
Age	37.96	35.70	2.257	0.193
	(10.65)	(9.324)		
No formal education	0.171	0.143	0.029	0.825
	(0.380)	(0.352)		
Up to ten years of school	0.743	0.800	-0.057	0.546
	(0.440)	(0.403)		
Upper secondary school or more	0.0857	0.0571	0.029	0.740
	(0.282)	(0.234)		
Born in Syria	0.543	0.557	-0.014	1.000
	(0.502)	(0.500)		
Born in Eritrea	0.157	0.171	-0.014	1.000
	(0.367)	(0.380)		
Born in Somalia	0.114	0.0714	0.043	0.567
	(0.320)	(0.259)		
Born in rest of the world	0.186	0.200	-0.014	1.000
	(0.392)	(0.403)		
Time since arriving in Sweden	630.5	692.6	-62.055	0.452
	(371.9)	(550.5)		
Share with permanent residence permit	0.443	0.529	-0.086	0.392
	(0.500)	(0.503)		
Time since receiving residence permit	222.0	233.7	-11.739	0.402
	(77.51)	(82.09)		
Observations	70	70	140	140

Notes: Standard deviations in parentheses.

5 Results

5.1 Mean differences in employment

Figure 6 graphically illustrates the effects of the integration program by comparing treated and non-treated individuals during the first two years after randomization. The top graph of Figure 6 displays differences in employment rates across treated and control groups. Our employment outcome includes all types of employment except job creation schemes. In particular, we define

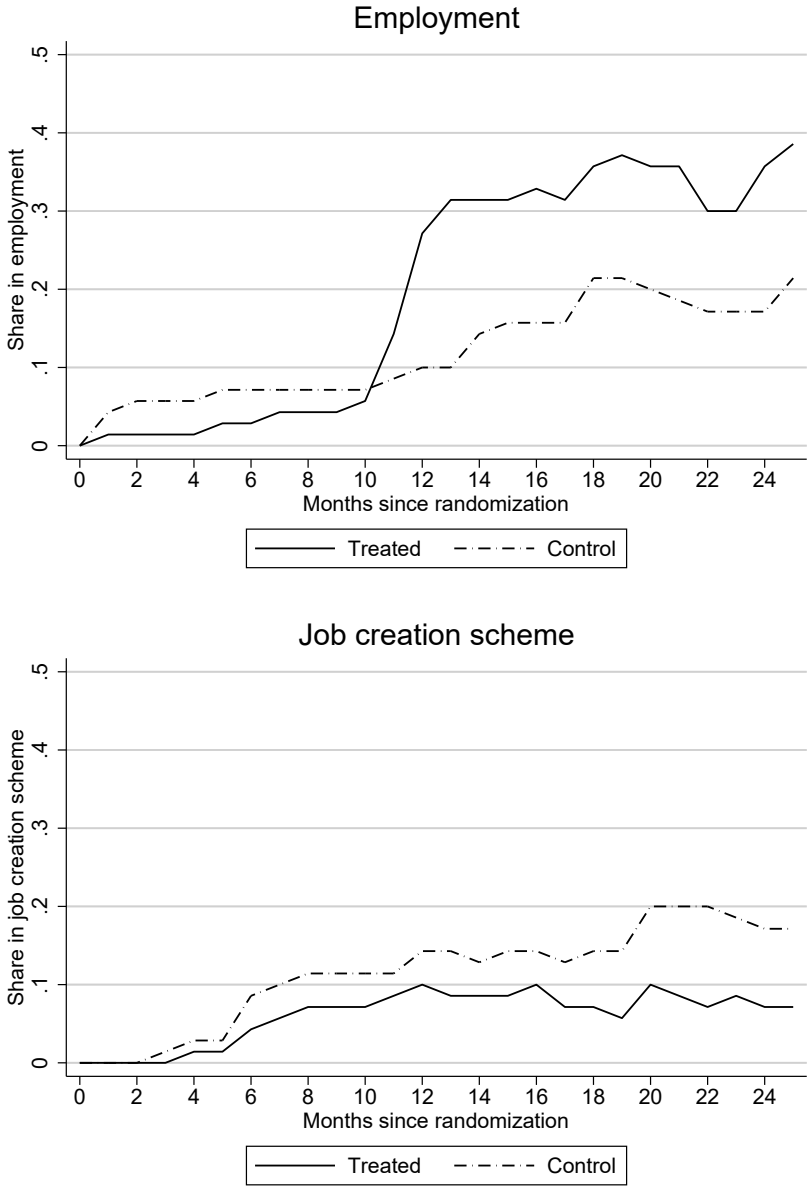
individuals as employed if they have non-subsidized jobs or if they are employed using the so-called *New Start Job* subsidy. All newly arrived unemployed immigrants are qualified for the subsidy, but they need to find an employer (private or public) willing to employ them.[30]

Both treated and non-treated individuals exhibit low employment rates (below 10 percent) during the first ten months. The slightly lower rates for the treatment group, relative to the control group, during this time period are expected since treated individuals participate full-time in language training and work practice. Hence, there seems to be a small lock-in effect associated with the program in the short run. The control group exhibits a slow and gradual increase in employment over the two years that we study, with average employment reaching 10 percent after the first year and 21 percent after the second year. In contrast, the treatment group exhibits a sharp and immediate increase in employment after 10 months, which is when their work practice ends. The magnitude of the immediate effect is remarkable, corresponding to around 25 percentage points. Just above 30 percent of the individuals in the treatment group are employed when leaving the program after one year, and the effect on employment seems to last over the entire second year (if anything, the effect increases during the second year).

In our analysis of labor market effects of the program, we also consider job creation schemes. The question we ask is whether the control group has a higher risk of ending up in what we look upon as created employment. We define job creation schemes as employment with large subsidies, mainly the so-called *Extratjänster*. This subsidy is only available for jobs in the public sector and covers the entire wage cost for the public sector employer, meaning that the employer has zero wage costs. In addition, having this type of employment does not entitle individuals to UI benefits. The bottom graph of Figure 6 indicates that the individuals in the control group to a larger extent end up in job creation schemes, especially if we consider the second year.

[30]The subsidy is given for one year at a time for a maximum of two years. It is typically used for job seekers who have been unemployed for more than 6 months. The subsidy amounts to as much as 2.5 times the payroll tax up to a full-time salary of SEK 20,000 (around USD 2,000) per month.

Figure 6: Mean rates in treatment and control groups.



Notes: The figure compares treatment and control groups per month after randomization. The first graph considers employment and the second graph considers job creation schemes.

5.2 Inference using Fisher's test

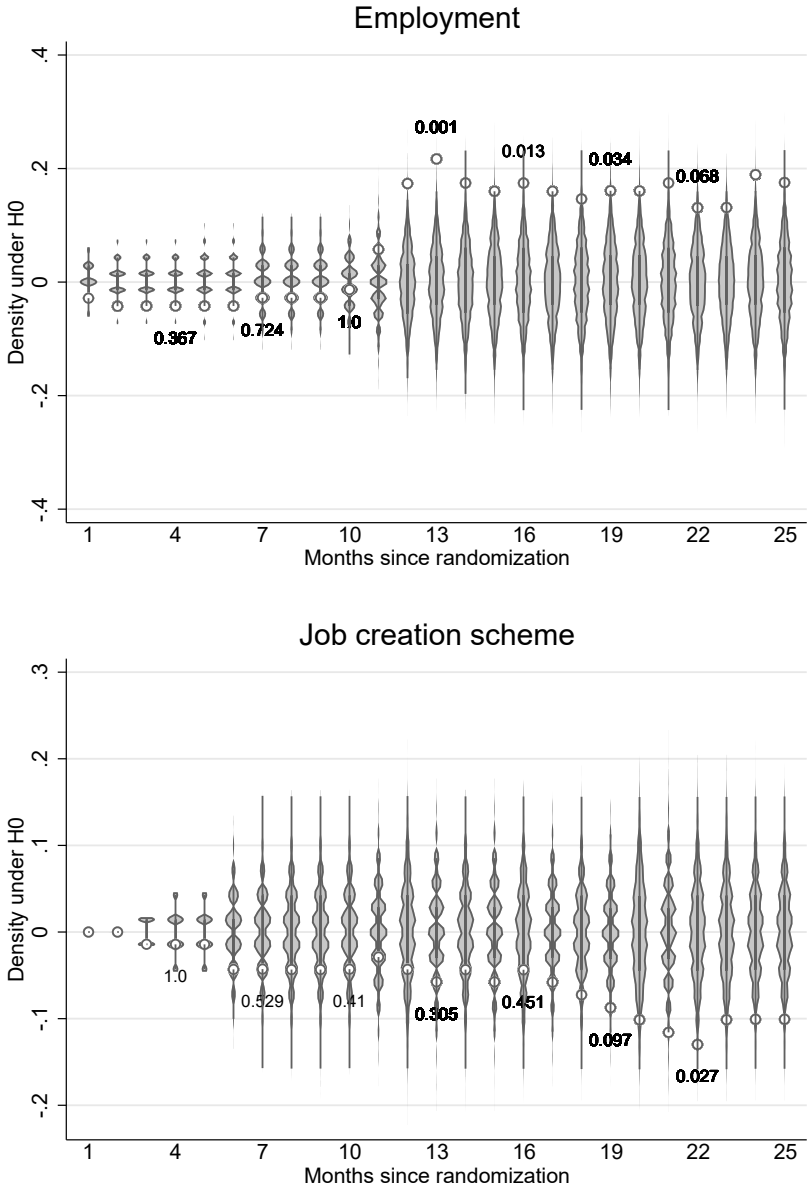
The two violin plots in Figure 7 present our estimates for the effects of the integration program *Välkommen till Framtiden*. We include monthly effects during the first and second year after randomization. For each month, we present both the full distribution of all estimated coefficients under the null hypothesis (the ”violins”)[31] and the actual point estimate from the randomization (the unfilled circular dots). For some months, the figure also shows the estimated exact p-values. The full distribution of point estimates and p-values are presented in Table 3 in the Appendix.

Figure 6, Figure 7, and Table 3 present convincing evidence of positive effects of the program on employment: around 30% of the individuals in the treatment group are employed each month during the first year after the program ends compared to an average of approximately 15% in the control group (cf. Figure 6). Inference based on Fisher’s exact test shows that the 15-percentage-point difference in employment is most likely not given by chance (from month 12 and onward, 11 out of 14 exact p-values are below 0.04 and they are all below 0.07; cf. Table 3 in the Appendix).

The violin plot at the bottom of Figure 7 brings us back to the question of whether the control group ends up in created employment. We find some evidence of significant effects during the second year, indicating that non-treated individuals to a larger extent participate in job creation schemes (cf. estimates in Figure 7 and Table 3 in the Appendix).

[31] We draw 10,000 random samples from all possible combinations originating from the allocation mechanism, meaning that the violin distributions contain 10,000 point estimates.

Figure 7: Effects on employment.



Notes: The figure shows the effects of the integration program per month after randomization. The first graph considers employment and the second graph considers job creation schemes. The estimated effects are given by unfilled circles and the exact p-values are provided for some month-specific estimates.

6 Conclusions

In this paper, we have evaluated a new ambitious program targeted at low-educated, newly arrived refugees in Gothenburg, Sweden. The program contains intensive language training, work practice with close supervision, being matched with an employer or further education, and extended cooperation between the involved program owners.

The program starts approximately once a year, and for the second wave, starting in May 2017, 70 randomly assigned individuals from the pool of 140 eligible individuals (where 140 was the universe of eligible individuals in Gothenburg by the time of the start of the second wave) were given the opportunity to participate. This group is used to evaluate the effects on employment probabilities up to two years after the start of the program.

We reach two main conclusions. First, we find that the program has clear positive effects on employment. During the second year after randomization (the first year after the end of the program), approximately 30% of the individuals in the treatment group are employed each month while the corresponding figure for the individuals in the control group is approximately 15%. Inference based on Fisher's exact test shows that the 15 percentage point difference in employment in the treatment and control group is most likely not given by chance (the exact p-values are typically below 0.04). This is an intention-to-treat effect (44 of the 70 individuals offered to participate in the program actually started the program) that we consider the most policy-relevant treatment effect. Since we estimate the average treatment effect of being offered a place in the new integration program, we estimate the effect of how the PES actually allocates places in the program. Second, we find that the control group to a larger extent participates in job creation schemes.

One potential concern is that the integration program might give rise to displacement effects. We argue that displacement is a small concern in the case under study. There are **three** reasons for this argument. The first reason is that the program focused on jobs where there is a shortage of labor (the real estate company involved in the program, representing the demand side of the labor market, had identified jobs where it was hard to find enough workers). The second reason is that each cohort of eligible individuals is rather small, hardly affecting other unemployed individuals' employment probabilities. Scaling up the program for the group under study (low-educated, newly arrived refugees) implies implementing the program in different municipalities and/or at different points in time in a given municipality, since the target group is quite small for a given municipality at a given point in time. This means that the ”scaling up” is not likely to result in displacements effects or general equilibrium effects. The final reason is that even if the program leads to low-educated refugees displacing some unemployed individuals who are closer to the labor market, it can be argued that this is a displacement that can be accepted.

Extending the program to include more refugees is doable. An important policy conclusion from the analyses in this paper is that policymakers should involve the demand side of the labor market at an early stage, both in order to identify jobs where there is a shortage of workers and to help out in designing the program. Since this is a conclusion with bearing in all cities experiencing an influx of low-educated refugees, the findings in this paper are likely to have important external validity. Put differently, the Gothenburg integration program can be implemented in most, if not all, Western countries experiencing refugee immigration.

The positive short-run effects found in this study are not only important *per se*, they are also promising in terms of finding positive outcomes in the longer run. We are in the process of collecting new data from Statistics Sweden. This will provide us with a richer set of variables for the individuals in the experiment, allowing us to examine the effects of the program in relation to more outcomes (for instance, if they receive welfare) and in the longer run. We will then also be in a good position for making a credible cost-benefit analysis of the new integration program.

References

Andersson Joona, P. (2020). Flykting- och anhöriginvandrares väg till arbete. SNS-Rapport.

Andersson Joona, P. and Nekby, L. (2012). Intensive Coaching of New Immigrants: An Evaluation Based on Random Program Assignment. *The Scandinavian Journal of Economics*, 114(2):575–600.

Arendt, J. N., Bolvig, I., Foged, M., Hasager, L., and Peri, G. (2020). Integrating Refugees: Language Training or Work-First Incentives? Working Paper 26834, National Bureau of Economic Research.

Åslund, O. and Johansson, P. (2011). Virtues of SIN: Can Intensified Public Efforts Help Disadvantaged Immigrants? *Evaluation Review*, 35(4):399–427.

Battisti, M., Giesing, Y., and Laurentsyeva, N. (2019). Can job search assistance improve the labour market integration of refugees? Evidence from a field experiment. *Labour Economics*, 61.

Berman, E., Lang, K., and Siniver, E. (2003). Language-skill complementarity: Returns to immigrant language acquisition. *Labour Economics*, 10(3):265–290.

Butschek, S. and 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(1):48.

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

Chiswick, B. R. (1991). Speaking, Reading, and Earnings among Low-Skilled Immigrants. *Journal of Labor Economics*, 9(2):149–170.

Clausen, J., Heinesen, E., Hummelgaard, H., Husted, L., and 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.

Cortes, K. (2004). Are refugees different from economic immigrants? Some empirical evidence on the heterogeneity of immigrant groups in the United States. *Review of Economics and Statistics*, 86(2):465–480.

Dustmann, C., Fasani, F., Frattini, T., Minale, L., and Schönberg, U. (2017). On the economics and politics of refugee migration. *Economic Policy*, 32(91):497–550.

Dustmann, C. and Soest, A. v. (2001). Language Fluency and Earnings: Estimation with Misclassified Language Indicators. *The Review of Economics and Statistics*, 83(4):663–674.

Fasani, F., Frattini, T., and Minale, L. (2020). Lift the Ban? Initial Employment Restrictions and Refugee Labour Market Outcomes. CEPR Discussion Papers 14765.

Hangartner, D. and Sarvimäki, M. (2017). Dealing with the refugee crisis: Policy lessons from economics and political science. Report, VATT and Economic Policy Council.

Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.

Lochmann, A., Rapoport, H., and Speciale, B. (2019). The effect of language training on immigrants' economic integration: Empirical evidence from France. *European Economic Review*, 113:265–296.

Ruiz, I. and Vargas-Silva, C. (2018). Differences in labour market outcomes between natives, refugees and other migrants in the UK. *Journal of Economic Geography*, 18(4):855–885.

Sarvimäki, M. and Hämäläinen, K. (2016). Integrating Immigrants: The Impact of Restructuring Active Labor Market Programs. *Journal of Labor Economics*, 34(2):479–508.

Appendix A Additional table

Table 3: Estimated effects (τ) and p-values using Fischer's exact test

	Employment		Job creation	
Month since randomization	τ	P-value	τ	P-value
1	-.0284	.618	0	1
2	-.0425	.363	0	1
3	-.042	.367	-.0141	1
4	-.042	.367	-.0142	1
5	-.0422	.446	-.0142	1
6	-.0422	.446	-.0431	.33
7	-.0277	.724	-.0428	.529
8	-.0277	.724	-.0433	.403
9	-.0277	.724	-.0433	.403
10	-.0132	1	-.0429	.41
11	.0576	.3	-.0289	.588
12	.174	.0075	-.0431	.461
13	.217	.0013	-.0576	.305
14	.175	.0109	-.0434	.428
15	.16	.0241	-.0576	.306
16	.174	.0129	-.0435	.451
17	.16	.0224	-.0579	.27
18	.146	.0539	-.0725	.18
19	.161	.0341	-.0873	.0968
20	.161	.0297	-.101	.101
21	.175	.0179	-.116	.0577
22	.131	.0675	-.13	.0268
23	.131	.0667	-.101	.0915
24	.189	.0087	-.101	.0759
25	.175	.0201	-.101	.0759