

Aldashev, Alisher; Thomsen, Stephan L.; Walter, Thomas

Working Paper

Short-term training programs for immigrants: do effects differ from natives and why?

ZEW Discussion Papers, No. 10-021

Provided in Cooperation with:

ZEW - Leibniz Centre for European Economic Research

Suggested Citation: Aldashev, Alisher; Thomsen, Stephan L.; Walter, Thomas (2010) : Short-term training programs for immigrants: do effects differ from natives and why?, ZEW Discussion Papers, No. 10-021, Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim

This Version is available at:

<https://hdl.handle.net/10419/30206>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

Discussion Paper No. 10-021

**Short-Term Training Programs
for Immigrants:
Do Effects Differ from Natives and Why?**

Alisher Aldashev, Stephan L. Thomsen,
and Thomas Walter

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 10-021

**Short-Term Training Programs
for Immigrants:
Do Effects Differ from Natives and Why?**

Alisher Aldashev, Stephan L. Thomsen,
and Thomas Walter

Download this ZEW Discussion Paper from our ftp server:

<ftp://ftp.zew.de/pub/zew-docs/dp/dp10021.pdf>

Die Discussion Papers dienen einer möglichst schnellen Verbreitung von
neueren Forschungsarbeiten des ZEW. Die Beiträge liegen in alleiniger Verantwortung
der Autoren und stellen nicht notwendigerweise die Meinung des ZEW dar.

Discussion Papers are intended to make results of ZEW research promptly available to other
economists in order to encourage discussion and suggestions for revisions. The authors are solely
responsible for the contents which do not necessarily represent the opinion of the ZEW.

Non-technical summary

Labor market integration of immigrants provides a difficulty in many countries. As a consequence, immigrants are substantially overrepresented in welfare systems. Despite forming a substantial share of all welfare recipients, relatively little is known about the impact of welfare-to-work programs on labor market outcomes of this group. This paper explicitly focuses on immigrants and evaluates the effects of a major German welfare-to-work program, namely off-the-job short-term training, on the probability of exiting the welfare system by taking up employment. The core questions are: Are programs similarly effective for immigrants and natives? And, if differences in effects are observable between both ethnic groups, what are the causes of these differences? Are they due to observable differences in socio-demographic characteristics or are they due to unobservable differences which must be attributed to the immigrant status?

To answer these questions, we use a sample of about 80,000 immigrants and 80,000 natives from comprehensive register data of the inflows into welfare in 2006. The effects of natives are estimated to benchmark the effects for immigrants. The data provide detailed information about socio-demographic characteristics, employment history, program participation and the outcome variable of interest; in addition, they enable identification of immigrants beyond the concept of citizenship. Four different types of training are distinguished: aptitude tests, job search training, skill provision, and combined training programs. For the estimation of the treatment effects, we employ propensity score matching estimators in a dynamic setting, where treatment effects vary conditionally on the preceding duration in welfare. To answer the question whether differences in effects are caused by differences in the composition of the native and immigrant population in the welfare system (e.g. due to differences in education or in the age structure) or due to an immigrant fixed effect we suggest and apply a matching based decomposition of differences in treatment effects.

Our estimation results show that the considered training programs exhibit substantial effect heterogeneity. For aptitude tests we observe on average positive employment effects. While in the sample of women treatment effects are larger for natives than for immigrants, the picture is ambiguous for men depending on the timing of the training. Aptitude tests starting in the second or third quarter of welfare receipt generate larger treatment effects for immigrants, whereas native men benefit more from tests in the first quarter. The difference in treatment effects of natives and immigrants in the first quarter is mainly due to differences in observable characteristics between the two ethnic groups. Keeping all covariates constant, immigrants tend to benefit even more from aptitude tests than natives.

Job search training is ineffective for men irrespective of being an immigrant or not. Native women benefit from this form of training, while immigrant females face negative treatment effects. The large difference in treatment effects of native and immigrant women cannot be explained by observable characteristics and must instead be attributed to an immigrant fixed effect. Holding everything else constant, immigrant females participating at job search training have a nearly 15 percentage point lower treatment effect than native participants. Even though the negative immigrant fixed effect fades away nine months after the program starts, this result gives cause for serious concern. Job search training decreases rather than increases employment chances of female immigrants.

In contrast, female immigrants clearly benefit from skill provision, which is a program exhibiting positive effects in general when assigned early during the welfare spell. For this form of training the immigrant fixed effect increases over time and amounts to 14 percentage points one year after program start. Thus, when netting out observable differences between immigrants and natives, the former have on average a 14 percentage points larger treatment effect than the latter.

For the combined training programs, we do not find statistically significant differences in any subgroup. These results reflect the finding that combined programs are rather ineffective for both ethnic groups and for both genders. The general ineffectiveness of combined training programs might be due to the characteristics of the targeted group, since combined programs are in particular assigned to those persons who were out of labor force for a substantial fraction of the final two years before treatment. These persons are likely to face multiple obstacles for employment uptake, which might not be remediable by combined training programs.

Das Wichtigste in Kürze

Der Anteil von Immigranten in Systemen der sozialen Sicherung ist in vielen Ländern der OECD stark überproportional. In Deutschland hatten im Jahr 2006 mehr als 34% aller erwerbsfähigen Hilfebedürftigen einen Migrationshintergrund, wohingegen der Anteil in der gesamten Bevölkerung lediglich 19.5% betrug. Trotz dieser starken Betroffenheit von Hilfebedürftigkeit ist bisher wenig über die Wirkung von arbeitsmarktpolitischen Maßnahmen auf die Arbeitsmarktchancen der Immigranten bekannt. Diese Studie untersucht die Beschäftigungswirkungen von kurzen außerbetrieblichen Trainingsmaßnahmen für Personen mit Migrationshintergrund, die ein zentrales Element der arbeitsmarktpolitischen Aktivierung in Deutschland sind. Gibt es Unterschiede in der Wirksamkeit von Trainingsmaßnahmen zwischen Immigranten und Einheimischen? Worauf sind potentielle Unterschiede zurückzuführen? Haben sie ihre Ursache in beobachtbaren Unterschieden der soziodemographischen Charakteristika beider Gruppen oder sind sie in unbeobachtbaren Unterschieden begründet, die durch den Migrationshintergrund an sich bedingt sind?

Um diese Fragen zu beantworten, verwenden wir Geschäftsdaten der Bundesagentur für Arbeit mit umfangreichen Informationen zu jeweils 80,000 Immigranten und Deutschen ohne Migrationshintergrund, die im Jahr 2006 in den Rechtskreis des SGB II zugegangen sind. Wir unterscheiden vier verschiedene Trainingsmaßnahmen: Eignungsfeststellungen, Bewerbungstrainings, Vermittlung von Kenntnissen und Maßnahmekombinationen. Die Wirkungen der Maßnahmen für Deutsche ohne Migrationshintergrund bilden den Referenzmaßstab. Zur Berechnung der Maßnahmeeffekte verwenden wir einen dynamischen Propensity Score-Matching-Ansatz. Zur Beantwortung der Frage, ob Unterschiede in den Maßnahmeeffekten zwischen Personen mit und ohne Migrationshintergrund auf eine unterschiedliche soziodemographische Zusammensetzung beider Gruppen zurückzuführen sind oder auf dem Migrationshintergrund an sich beruhen, schlagen wir eine Dekomposition der Effekte vor.

Die empirischen Ergebnisse zeigen eine erhebliche Heterogenität in den Wirkungen der Trainingsmaßnahmen. Bei den Eignungsfeststellungsmaßnahmen beobachten wir im Durchschnitt positive Beschäftigungseffekte. Während bei den Frauen ohne Migrationshintergrund die Effekte stärker ausgeprägt sind als für Immigrantinnen, ergibt sich für Männer kein eindeutiges Bild. Hier variiert die relative Stärke der Effekte je nach dem Zeitpunkt des Maßnahmeeinsatzes. Die Unterschiede der Maßnahmeeffekte zwischen Personen mit und ohne Migrationshintergrund lassen sich vor allem auf Unterschiede in beobachtbaren Merkmalen beider Gruppen zurückführen. Kontrolliert man für alle diese Merkmale zeigt sich sogar, dass Immigranten und Immigrantinnen stärker von den Eignungsfeststellungen profitieren als einheimische erwerbsfähige Hilfebedürftige.

Bewerbungstrainings sind bei Männern unabhängig vom Migrationshintergrund ohne signifikante Wirkung. Frauen ohne Migrationshintergrund profitieren von dieser Form des Trainings, wohingegen die Beschäftigungschancen von Immigrantinnen negativ beeinflusst werden. Dieser große Wirkungsunterschied ist nicht auf beobachtbare Unterschiede zwischen den beiden ethnischen Gruppen zurückzuführen, sondern entsteht durch den Migrationshintergrund an sich. Kontrolliert man für alle beobachtbaren Merkmale haben Immigrantinnen einen um 15 Prozentpunkte niedrigeren Eingliederungserfolg als deutsche Teilnehmerinnen. Auch wenn sich dieser substantielle Unterschied über die Zeit abschwächt und etwa 9 Monate nach Maßnahmebeginn verschwindet, macht er deutlich, dass Bewerbungstrainings anders als intendiert die Beschäftigungschancen von Immigrantinnen verschlechtern anstatt verbessern.

Im Gegensatz zu den Bewerbungstrainings zeigt sich für die Vermittlung von Kenntnissen, dass Immigrantinnen von diesen Maßnahmen besonders profitieren. Ein Jahr nach Maßnahmebeginn ist der Maßnahmeeffekt für Frauen mit Migrationshintergrund um 14 Prozentpunkte höher als für Frauen, die in allen Merkmalen identisch sind, aber keinen Migrationshintergrund aufweisen. Bei den Männern zeigen sich keine signifikanten Unterschiede in der Maßnahmewirkung zwischen Immigranten und Deutschen ohne Migrationshintergrund. Bei den Maßnahmekombinationen lässt sich für keine der betrachteten Personengruppen eine signifikante Wirkung erkennen.

SHORT-TERM TRAINING PROGRAMS FOR IMMIGRANTS: DO EFFECTS DIFFER FROM NATIVES AND WHY?

Alisher Aldashev*

Kazakh-British Technical University of Almaty

Stephan L. Thomsen†

University of Magdeburg & ZEW Mannheim

Thomas Walter‡

ZEW Mannheim

This version: April 7, 2010

Abstract

We evaluate the effects of different short-term training programs on the employment chances of immigrant and native welfare recipients in Germany. In particular, we investigate whether program effects differ between both groups and what might cause these potential differences. In a first step, we evaluate program effects separately for immigrants and natives using propensity score matching estimators. To explain potential differences in effects between the groups, we suggest and apply a decomposition method based on the matching procedure that allows identification of differences due to observable characteristics and differences related to an immigrant fixed effect in a second step.

Keywords: Immigrants, short-term training programs, evaluation, decomposition, matching, Germany

JEL Classification: I38, C14, J61

*Alisher Aldashev is Assistant Professor at Kazakh-British Technical University of Almaty. Address: Kazakh British Technical University, International School of Economics, 59 Tole bi Str., Almaty 050000, Kazakhstan.

†Corresponding author: Stephan L. Thomsen is Assistant Professor of Labor Economics at Otto-von-Guericke-University Magdeburg and Research Associate at ZEW, Mannheim. Address: Otto-von-Guericke-University, Department of Economics and Management, PO Box 4120, D-39016 Magdeburg, e-mail: stephan.thomsen@ovgu.de, phone: +49 391 6718431, fax: +49 391 6711700.

‡Thomas Walter is Research Assistant at the Centre for European Economic Research (ZEW), Mannheim. Address: L7, 1, D-68161 Mannheim, e-mail: walter@zew.de, phone: +49 621 1235363, fax: +49 621 1235225.

1 Introduction

We evaluate the effects of four different short-term off-the-job training programs on the probability of exiting the welfare system by taking up employment for native and immigrant welfare recipients in Germany. In particular, we are interested in the effects for immigrant welfare recipients. Are programs similarly effective compared to natives? And, if differences in effects are observable between natives and immigrants, what are the causes of these differences? To answer these questions, we use a sample of about 160,000 observations from comprehensive register data of the inflows into welfare in 2006. These data of native and immigrants provide detailed information about sociodemographic characteristics, employment history, program participation and the outcome variable of interest; in addition, they enable identification of immigrants beyond the concept of citizenship. For the estimation of the treatment effects, we employ propensity score matching estimators in a dynamic setting (see Sianesi, 2004), where treatment effects vary conditionally on the preceding duration in welfare. To answer the question whether differences in effects are caused by differences in the composition of the native and immigrant population in the welfare system (e.g. due to differences in education or in the employment history) or due to an immigrant fixed effect we suggest and apply a matching based decomposition of differences in treatment effects.

Although there is a substantial literature analyzing the impacts of short-term training programs (see Kluge, 2006 for an overview on the international evidence) and there are a number of studies analyzing the programs in Germany¹, relatively little is known about the impacts on labor market outcomes for immigrants. Nevertheless, analyzing the effects for immigrants is important. Within the group of welfare recipients immigrants are clearly over-represented. In 2006, more than 34% of all welfare recipients were immigrants (Bundesministerium für Arbeit und Soziales, 2009) while the corresponding share of the population was only about 19.5% (see Statistisches Bundesamt, 2006). Immigrants have lower participation rates and higher unemployment rates not only in Germany but in most European countries, see OECD (2008) for a comprehensive description. Successful labor market integration of these persons is therefore an issue in most countries, and different

¹Short-term training programs for unemployed individuals in Germany have been evaluated, for example, by Hujer et al. (2006), Biewen et al. (2007) and Lechner and Wunsch (2008). In addition, programs for welfare recipients have been studied, e.g. by Wolff and Jozwiack (2007), Kopf (2009) and Huber et al. (2009).

integration plans have been adopted. Given the over-representation in welfare and unemployment, we would perceive that immigrants have larger difficulties for labor market integration compared to natives due to a lower degree of employability.

Participation in short-term training programs should help to improve employability and the programs are intended to increase the search efficiency and to improve productivity. However, whether these programs reach their purposes for immigrants (as for natives) is not clear ex ante and varying impacts can occur for a number of reasons. Clearly, successful integration depends on labor demand. If immigrants' productivity is too low and programs are not able to increase the productivity sufficiently to meet the required standards, there will be no effects. Similarly, if potential employers apply some kind of statistical discrimination with respect to immigrants then even if productivity is improved by participation the probability of placement could be lower compared to natives. Even in the absence of demand side effects, differences in placement may result from a different value of the programs for immigrants compared to natives. For example, to train immigrants in formally writing job applications may be counterproductive if the traditional search strategy is to rely on networks and contacts within the community. On the other hand, program participation could reveal unexpectedly high levels of productivity of treated immigrants to caseworkers, who might have undervalued these persons before assignment due to a lower average productivity among immigrants. This learn effect might induce caseworkers to increase their placement effort for treated immigrants which in turn might lead to larger treatment effects of short-term training for immigrants than for comparable native welfare recipients.

Determining the source of differences in program effectiveness between the two ethnic groups is important. If, for example, differences in program effectiveness are driven by differences in the composition of native and immigrant welfare recipients it implies a general potential for welfare agencies to improve the targeting of programs to participants. If, on the other hand, differences are due to the immigrant characteristic, then this points to discrimination in the effectiveness of short-term training programs and the question arises whether the use of programs for specific ethnic groups is reasonable at all. Clearly, both possible explanations for differences in program effects must cause concern among policy makers. However, since effect differences due to an immigrant fixed effect per se are especially problematic and affect more than one third of the welfare population, we will mainly focus on the contribution of the immigrant fixed effect to the observed differences

in the effectiveness of training programs. This paper thus contributes, on the one hand, to the small international literature on the effects of training and active labor market policy (ALMP) for immigrants, and, on the other hand, to the comprehensive literature on program evaluation.

The remainder of this paper is organized as follows. Section 2 presents details of the German welfare system and introduces the training programs of interest. In section 3 we discuss the related literature. The data used in the empirical analysis is described in section 4. Our evaluation approach and the propensity-score based decomposition method of differences in treatment effects between natives and immigrants is discussed in section 5. In section 6 we present the estimation results. The final section concludes. In addition, we provide a Data Appendix for selected descriptive statistics of our estimation sample.

2 Institutional Background

The German welfare system was substantially reformed at the beginning of 2005 with the introduction of the new Social Code II (*Sozialgesetzbuch II*).² Until 2005, welfare recipients were eligible for social assistance (SA) if they had not contributed to unemployment insurance before. In addition, persons whose unemployment benefit (UB) claims had expired were eligible for unemployment assistance (UA). If UA was too low to provide a minimum living standard, a combination of UA and SA was granted. In contrast to UB, UA and SA were both means-tested. With the welfare reform of January 2005, both programs were replaced by the so-called unemployment benefits II scheme (UBII). As opposed to UA, which replaced up to 57% of the previous net earnings, UBII (as former SA) does not depend on former earnings. The means-test takes into account the wealth and income of all individuals living in the household. At the beginning of 2005, UBII benefits for a single individual without children amounted to EUR 345 in West Germany and to EUR 331 in East Germany. Meanwhile, the level of UBII in East Germany was adjusted to the Western level and UBII was slightly raised in both parts to compensate for inflation (359 Euro since July 09). Moreover, UBII welfare payments also include compulsory social

²This reform was the last part of a series of four major reforms of the German labor market which were enacted between 2003 and 2005. These reforms have become known as ‘Hartz reforms’ named after the chairman of the commission proposing the reforms. Since the reform of the welfare system is the last of the four reforms it is also referred to as the ‘Hartz IV reform’. See Jacobi and Kluge (2007) for a description of all four ‘Hartz reforms’.

insurance contributions, rents and housing costs. Additional expenses for special needs may also be covered.

In order to be eligible for UBII, persons have to be aged 15 to 64 years and be able to work for at least 15 hours per week. It is important to note that unemployment is not a prerequisite for receipt of UBII. Individuals who are employed but whose household income is too low are also eligible for UBII. Claimants capable of work have to register with the local welfare agency and are obliged to participate in welfare-to-work programs. This obligation marks an important change in German welfare policy. Namely, for the first time welfare recipients became a target group of labor market activation. Before 2005, hardly any effort was made to reintegrate these persons into the labor market and welfare solely relied on passive benefit payments. Since 2005, the welfare recipients' rights and duties in the activation process are set out in a so-called 'integration contract' (*Eingliederungsvereinbarung*), an agreement between the welfare agency and the benefit recipient containing obligations with respect to program participation and job search activities, as well as detailing the services provided by the welfare agency. The integration contract is usually set up after the first meeting of a welfare recipient with the caseworker. The caseworker counsels and advises the welfare recipient and decides about placement in one of the various ALMP programs.

Include Table 1 about here

Table 1 provides some selected figures concerning the number of entitled persons to UBII and the corresponding spending. As becomes obvious, on average about 5 million people were entitled to UBII benefits; however, referring to the years from 2006 onwards a slight decline from about 5.4 million to 5.0 million persons could be established. The spending amounted to more than 30 billion Euro per year for passive UBII benefits. Corresponding to the shift in the number of entitled persons, spending declined slightly between 2006 and 2008. In contrast, the figures for the spending on ALMP emphasize the increased importance of the newly introduced need to activate the former welfare recipients. Whereas in 2005 only 3.1 billion Euro were spent overall, this figure increased by more than 50 percent up to 4.7 billion Euro in 2008. Within the scope of ALMP programs, short-term training programs are a quite frequently used measure. During the last years, between 411 and about 628 thousand UBII recipients per year have participated in these programs.³

³Short-term training programs were introduced in Germany with the enactment of Social Code III

The primary purpose of short-term training programs is to improve the employment prospects of the participating individuals. For this reason, programs consist of three different types of measures (modules) that can be accomplished separately or in combination and allow a flexible implementation in line with the specific needs of the welfare recipients and the options of the local welfare agencies as well. The first type of courses are aptitude tests (*Eignungsfeststellungen*) which last for up to four weeks. These tests are used to assess the suitability of participants in terms of skills, capability and labor market opportunities for specific occupations. During the assessment process occupation specific skills are provided which shall help to improve employment chances in the respective occupations. The measures of the second type of short-term training programs aim at improving the applicant's presentation and job search abilities (*Überprüfung der Verfügbarkeit/Bewerbertraining*). The activities support the individual's efforts to find work or efforts by the welfare agency to place him/her, especially through job-application training, counseling on job search possibilities or measures assessing the person's willingness and ability to work (work-tests). Measures of the second type are promoted for up to two weeks and will be referred to as job search training in this analysis. The third type contains practical training of the participants (for up to eight weeks) providing necessary skills and techniques required for placement in employment or vocational training (*Vermittlung notwendiger Kenntnisse und Fertigkeiten*). The courses cover, for example, specific working techniques like business administration or computer courses. We will refer to this form of training as skill provision. Finally, combinations of modules, e.g., a job aptitude test followed by a computer course, could be granted for a maximum of twelve weeks. This is the fourth type of training and will be referred to as combined training programs. Financial support during the training is provided by the FEA and covers course costs, examination fees, and travel grants as well as child care. In addition, participants receive UBII payments. Decisions about support of courses and placement of welfare recipients are made by the welfare agencies. Support is authorized on recommendation or with the approval of the agency only and activities are often initiated by caseworkers. However, short-term training programs may be initiated by welfare recipients as well. Short-term training programs could be provided on-the-job within firms

(*Sozialgesetzbuch III*) in 1997/1998. They replaced the former short-term qualification measures (*kurzzeitige Qualifizierungsmaßnahmen*), training measures for UB and UA recipients and employment counseling measures (*Maßnahmen der Arbeitsberatung*). In 2005, the rules from Social Code III were adopted in Social Code II.

and off-the-job. If provided off-the-job, activities are conducted by specialized service providers (*Bildungsträger*). Evaluation of the treatment effects of on-the-job courses may be complicated due to potential windfall gains of the supporting employers that have to be considered. For this reason, we concentrate the analysis on off-the-job courses only.

The institutional set-up of short-term training programs implies two channels through which programs affect the job search of the participants and, therefore, the employment chances and the probability of leaving welfare. On the one hand, the modules that improve or support the job placement on part of the welfare agency or the self-contained job search of the participants can be expected to improve the search behavior of the participants by increasing the intensity as well as the efficiency of the search efforts. More efficient job search will lead to an increase in the job offer arrival rate, which increases the probability of leaving welfare. However, it will make job seekers more selective with respect to potential job offer and induces a negative indirect effect on the transition. The overall effect is then the sum of the positive direct and the negative indirect effect. Van den Berg (1994) provides sufficient conditions for the wage offer distribution that ensure a positive net effect. On the other hand, participation in short-term training could improve the job-relevant skills and therefore increase the job opportunities of the participants. Increasing the skills is equivalent to increasing productivity which enables participants to apply for jobs associated with on average higher wages. In terms of job search theory this equals a shift of the wage offer distribution to the right. According to Mortensen (1986), an increase in the mean of the wage offer distribution increases the reservation wage by an amount less than the increase in the mean, and, therefore, this will increase the probability of leaving welfare as well. Clearly, the theoretical perspective implies positive effects of participation in a short-term training program on the probability of leaving welfare and the probability of taking up employment. However, for the theoretically derived positive effects to hold in reality there have to be potential employers willing to engage the participants in the programs. Despite the decrease in the reservation wage due to the increased search effort or the higher productivity obtained in the practical training, participants may still possess productivities too low to be remunerated by the market. In that case, there will be no positive effects of participation.

3 Related Literature

Referring to the literature that analyses ALMP programs with a particular focus on immigrants, Clausen et al. (2009) evaluate the effects of ALMP programs on the hazard rate into regular employment for newly arrived immigrants in Denmark.⁴ The programs are part of the integration policies specifically designed for facilitating the labor-market integration of newly arrived immigrants (introduced in 1999). The emphasis is on programs taking account of language skills. All in all, six different programs are evaluated. Within these programs so-called counseling and upgrading programs come closest to the short-term training programs we analyze here. The counseling and upgrading programs provide counseling regarding employment and education options but may also include voluntary unpaid work, adult education and supplementary training. The results show negative effects of counseling and upgrading which seems to be in contrast to the literature in that most previous studies find positive effects of counseling; however, the authors mitigate their results by noting that the effects of counseling and upgrading are only significant in the larger of the two samples used.

In addition, Cohen-Goldner and Eckstein (2009) evaluate a government provided training programme for highly-skilled female immigrants from the Former Soviet Union (FSU) in Israel. For estimation of effects, they apply dynamic programming and results show that training has no significant impact on the mean offered wage in blue-collar jobs, but does increase the mean offered wage in white-collar jobs. Nevertheless, training increases the probability of receiving a job offer significantly. However, these programs are not directly comparable to the short-term training in Germany as programs last for six months with 26 hours of study per week and participation rates are clearly higher (about 47% of all immigrants in Israel).

A third study is provided by Hämäläinen and Sarvimäki (2008). They evaluate the effects of integration plans for immigrants in Finland, which have the aim to promote integration, equality and freedom of choice by providing measures that help to achieve information and skills needed in Finnish society. The integration plan provides an individualized pathway containing measures of acquiring language skills, preparatory and/or vocational training, career counseling, rehabilitation, work practice and so forth; typically various measures are

⁴In addition, Rosholm and Vejlin (2010) analyze the effects of reducing income transfers to refugee immigrants. However, this is a change of passive labor market policy.

combined in paths in which one measure precedes another; the integration plan is aborted if an immigrant finds permanent, full-time employment (or becomes full-time student). Based on a regression discontinuity design estimator, the results show positive effects of the integration plan. The authors explain the positive effects to have mainly arisen from individually tailored plans combined with better co-ordination of the existing resources. Hence, it implies that individualized plans provide a fairly cost-efficient way to support the integration of immigrants (or at least of those who are likely to participate in the labor force). Unfortunately, the analysis does not go into detail how important training courses are for the labor market success of the immigrants.

In contrast to the foreign programs, which are specifically designed for (newly arrived) immigrants, German welfare-to-work programs including short-term training are identical for immigrants and natives. The effects of short-term training programs on the employment chances of welfare recipients in Germany have been analyzed by Wolff and Jozwiack (2007), Huber et al. (2009), and Kopf (2009) already. The studies vary with respect to the time horizon and the data used for estimation, but all apply propensity score matching estimators. Wolff and Jozwiack (2007) use register data similar to those used here and also consider immigrants as a subgroup in the analysis. However, the definition of immigrant status is not as detailed as in our study and the authors do not distinguish between different training modules. They find that short-term off-the-job training programs are on average ineffective among immigrant men to increase the probability of exiting the welfare system by taking up employment. For women with migration background significantly negative treatment effects are estimated for the first six months after the program start, which then fade away towards the end of the observation period (20 months). The employment chances of women without migration background living in West Germany are also reduced twenty months after the treatment, while there is no significant effect on native women in East Germany and on native men irrespective of the location.

Huber et al. (2009) use combined administrative and survey data to evaluate three types of welfare-to-work programs including short-term training, but they also do not distinguish between different training modules. Pooling all modules together, they estimate positive employment effects of short-term training, which are mainly driven by the subsample of persons without migration background. The estimated effects for immigrants are positive as well, but they are statistically not significant. Insignificance might be due to the relatively low number of treated observations with migration background.

Kopf (2009) uses the same data as Wolff and Jozwiack (2007) but distinguishes five types of off-the-job training programs including application training, work tests, aptitude tests, skill training courses, and combined programs. Her distinction is similar to the one used by us, but we pool application training and work tests into one category since they have overlapping contents. Kopf (2009) runs separate estimations for men and women and for East and West Germany, but does not consider immigrants. She finds that application training has negative locking-in effects lasting up to one year in the subsample of East German men. After the locking-in phase effects are close to zero and statistically insignificant in all subgroups. Work tests exhibit a shorter locking-in period than application training does. Nevertheless, only for West German men are significantly positive effects observed one and a half years after the program start. Aptitude tests show positive effects for men in both parts of Germany and for East German women, whereas West German women do not seem to benefit. For skill training the estimates are significantly positive within all subgroups, but again effects are least pronounced among West German women. Combined training programs are rather ineffective. Here, effects become significantly positive only for West German men 18 months after the program start and for East German women 6 to 13 month after treatment so that the effect in this group is only temporary.

4 Data

For the empirical analysis, we use a sample of all inflows into welfare in Germany from January, 1st 2006 to December, 31st 2006. The data stem from administrative records of the Federal Employment Agency (*Bundesagentur für Arbeit*, FEA) and were provided by the Institute for Employment Research, Nuremberg. To ensure that inflows in the data are not short-term recurrences of welfare episodes, for example due to false reporting or data errors, only persons are regarded who have not been registered in welfare for at least three months before the sampling date. The data were merged from five different sources of administrative records. The main source is the *Integrated Employment Biography* data set (*Integrierte Erwerbsbiographien*, IEB), which provides comprehensive information with regard to the socio-demographic situation, the labor market history, and the participation in ALMP programs. The detailed data allow distinction of the four types of off-the-job short-term training programs (three modules and program combinations) for the empirical analysis. The information included in IEB covers the years 1990 to 2007 and, thus, provides

a sufficient source of background information for inflows into UBII in 2006. These data allow for quite a detailed characterization of the current situation and the labor market chances of the UBII recipients. However, since UBII entitlement is means-tested with consideration of the wealth and the income of further household members, we merge information on further persons living in the households that are recorded in the *Benefit History Master Records (Leistungshistorikgrunddatei, BHMR)*.

In the empirical analysis, we distinguish the following ethnic groups: Immigrants comprise all foreigners and naturalized persons. Foreigners are persons who do not possess German citizenship. The naturalized group contains, on the one hand, German resettlers from Eastern Europe, and, on the other hand, naturalized foreigners. Although citizenship is recorded in IEB as well, identification of naturalized immigrants and German resettlers from Eastern Europe could not be obtained from this dataset. To identify resettlers we consider the information on the immigration date recorded in the *Job Seeker Statistics (Arbeitsuchendenstatistik, ASU)* dating back to 1990, which explicitly contains the information on resettler status. To identify naturalized immigrants, we use the information from the IEB for the years 1990 to 2007 and in addition the *Employment History Records (Beschäftigtenhistorik, EHR)* for the years 1975 to 1989. A person with German citizenship at the sampling date who were recorded being a foreigner in any spell since 1975 is treated as a naturalized. Unfortunately, the administrative records of the FEA contain neither information about the place of birth nor about the parents of the individual. Moreover, since minors (persons under 18 years of age) do not appear in any of these data sources, we are neither able to identify immigrants who were naturalized at an early age nor distinguish first and second generation immigrants.

As the main purpose of ALMP is to eliminate welfare dependency, we could use the drop-off rate from welfare as an outcome variable to evaluate the effects of short-term training programs and to decompose the differences in the effects. However, elimination of welfare dependency does not solely depend on the direct effects for the individual under study but may result from changes in the household as well, e.g. if the income of the partner increases. Therefore, we estimate the effects of training on the drop-off rate from welfare conditional on employment uptake of the individual. This outcome variable measures whether the training is able to improve the situation of a treated individual such that there is a transition to employment and welfare dependency is terminated. The variable can be observed on a monthly basis until July 2008 and has been merged from the *Employment*

*Statistics Register (Beschäftigtenstatistik, ESR).*⁵

In line with the two empirical questions of the paper, i.e. the evaluation of the program effects for immigrants and natives and the decomposition of effect differences, the analysis sample was drawn in a 1:1 ratio of immigrants and native Germans on regional level. In a first step, 80,000 immigrants were randomly drawn from the total inflow population into welfare in 2006. Then in a second step, for each immigrant randomly drawn from a welfare agency district, one native German was drawn from the same district resulting in an overall sample of about 160,000 welfare recipients. Therefore, immigrant-native German ratios are balanced across districts and should mitigate regional imbalances in the distribution of immigrants that could affect the estimates.

For the analysis presented here, the sample is restricted to unemployed welfare recipients aged 18 to 57 years at the sampling date. Although unemployment is not a prerequisite for receiving welfare benefits, it is required for participation in full-time short-term training programs. In addition, welfare recipients younger than 18 years are excluded so that the estimates are not affected by compulsory schooling. Welfare recipients aged 58 years and above are eligible for so-called relaxed welfare receipt. Within this scheme active job search is not required for benefit entitlement and claimants can rely on welfare until (early) retirement age. The final sample for the analysis contains 82,774 observations of which slightly more than half are natives (43,344) and the rest are immigrants (39,430). Using the information in the IEB, we identify for each person the first assigned program during the welfare spell and evaluate participation against nonparticipation in any other program at the time starting the program.

As can be seen from the descriptive statistics provided in the Data Appendix, 4,628 of the 43,344 natives (2,851 men and 1,777 women) are assigned to one of the four considered short-term training programs during the first year of their welfare spell. Among the immigrants the ratio of participants is lower, with only 3,871 individuals (2,599 men and 1,272 women) treated. Despite this difference in the participation ratios, the mix of

⁵It has to be noted that due to delays in reporting by employers, the information available in the ESR has an up to two-year time lag. Therefore, in a first step the FEA forecasts the information and then in a second step the forecast is replaced by the actually reported information. In consequence, assessing contemporary effects of welfare-to-work programs is possible, but the results will be based purely on forecasted employment information. As the evaluation of program effects should be based on actually reported, rather than forecasted information, our observation period ends in July 2008. Data were extracted in February 2009. However, as the time lag between the corresponding date of information and the extraction from the ESR for our analysis amounted to only eight months, the relation between reported and forecasted data was extensively checked. Based on the results of Fröhlich, Kaimer, and Stamm (2004), the share of forecasted data used in the analysis amounts to between four and ten percent at maximum.

assigned programs is similar in both ethnic groups. Aptitude tests are most frequently assigned as first program both for natives and immigrants. Nearly one third of all assigned short-term training programs are aptitude tests. Skill provision and combined training programs have a share of about 25% each. Job search training is used with the lowest frequency resulting in a share of somewhat less than 20%.

Selection into the different training programs is mainly driven by the employment biography of individuals. While sociodemographic characteristics are fairly similar among participants in the four considered programs, we observe substantial differences with respect to the time spent in employment, unemployment and out of the labor force prior to program start. These differences can be noticed both for natives and immigrants.

Combined training programs are assigned in particular to those persons who were out of the labor force for a very long period during the final two years before treatment. These individuals face multiple disadvantages when trying to get back to employment. Thus, the combined training program attempts to comprehensively tackle these disadvantages. Aptitude tests and skill provision are mainly targeted at those persons with a high incidence of unemployment. This reflects the fact that these programs are used to learn about the suitability of participants for different occupations and to refresh general human capital which might have been depreciated during unemployment. In contrast, job search training is focused on individuals with fairly good employment records who recently entered unemployment. These persons still have a valuable human capital stock but need support for writing job applications and attending job interviews.

Even though native and immigrant participants in the different training programs have similar employment biographies, they are distinct with respect to some sociodemographic characteristics. Considerable differences exist in terms of household composition. Immigrants are less frequently single and, thus, household size for immigrants is on average larger than for natives. Moreover, the variation in educational achievement is larger for immigrants than for natives. We observe a relatively large share of immigrants without any school leaving certificate, but also a noticeable share of persons with a university entrance diploma, especially among women. In addition, differences between the two ethnic groups are apparent in the lower end of the age distribution. While the share of treated immigrants aged between 18 and 24 is lower than for natives, the opposite is true for 25 to 34 aged individuals. The differences are more pronounced for men than for women.

However, despite these differences, there is sufficient overlap in the distribution of covariates of natives and immigrants so that both groups are comparable with respect to the effectiveness of training programs.

5 Evaluation Approach

5.1 Estimation of Treatment Effects

The evaluation of the treatment effects of a participation in a short-term training program on the drop-off rates from welfare has to consider the set-up of the comprehensive system of ALMP in Germany. This system is characterized by a wide array of programs which take place continuously over time and are open to welfare recipients who meet certain eligibility criteria, where participation can take place at different points of time during the welfare spell. Recent empirical literature highlights the need to consider the timing of treatment in the unemployment spell when evaluating treatment effects, see e.g. Abbring and van den Berg (2003), Sianesi (2004), Thomsen (2007), Fredriksson and Johansson (2008), or Hujer and Thomsen (2010). Whereas standard evaluation literature usually deals only with binary information, i.e. whether an individual has been subject to treatment or not, this literature points out the importance of information on the timing of treatment events as it conveys useful information for the identification of the treatment effect and has implications for the definition of the comparison groups. Specifically, the starting point of the program within the individual welfare spell may be an important determinant for the selection of participating individuals, as well as for the type of program the individual is assigned to.

The basis of the empirical analysis is given by the potential outcome approach of causality, comprehensively described in Heckman et al. (1999) and variously attributed to e.g. Neyman (1923), Roy (1951) and Rubin (1974). Following the conventional notation, let Y^1 and Y^0 denote the two potential outcomes, where Y^1 is the outcome when the individual participates in the program, and Y^0 is the outcome, when the individual does not participate. Since the individual cannot be in both states at the same time, one of the potential outcomes is unobservable and direct estimation of the treatment effect is impossible.⁶

⁶Imbens (2000) and Lechner (2002) generalize the framework for situations where a whole range of programs is available. Although we analyze a number of different types of short-term training programs, the focus of the analysis are the effects of participation compared to non-participation in that program

Therefore, to identify the treatment effect we have to provide an estimate of the unobserved state. We focus on the average effect of treatment on the treated (ATT) at some given elapsed welfare duration. Conditioning on the elapsed welfare duration is sensible in the German context for a reason first raised by Sianesi (2004). She argues that in a comprehensive ALMP system a person will join a program at some point, provided the individual remains in welfare long enough. Consequently, the reason why an individual is not observed as participating in a program is that the person has already left the welfare system, or the time horizon of the analysis is too short. Obviously, although participation in a program is not mandatory in Germany, like it is for instance in Sweden, it tends to be true that benefit recipients become more likely to participate in any program the longer they remain on welfare. The argument is therefore reasonable for the evaluation of German training programs as well.

In line with that, participation and non-participation have to be defined dynamically, i.e. with respect to the point in time in which the comparison is made. According to Sianesi (2004), persons who have neither entered a program nor left welfare up to a specific point in time are defined as non-participants of interest or ‘waiters’ (in the sense that they are waiting to be allocated to a program). Thus, non-participation can be interpreted as the default state for each individual, and everybody is a non-participant until entering a program or leaving to take up a job. In this context, it should be noted that individuals who are defined as non-participants at the moment we start our comparison may enter a program at a later point in time. The evaluation approach in the dynamic setting could be formalized as follows. Let $U = \{0, \dots, U_{\max}\}$ define the discrete elapsed welfare duration of the individual since registration at the local welfare agency. Furthermore, let u denote the point of time during the welfare spell in which the program of interest starts and D_u the treatment indicator with the discrete time index. $D_u = 1$ if the individual starts a program at time u of the welfare spell, $D_u = 0$ if the individual remains on welfare at u . Program effects are estimated for time t , i.e. the time since the program started. The hypothetical outcomes for time t given a treatment at time u are then defined as $Y_{t,u}^1$ for individuals who received the treatment at u and $Y_{t,u}^0$ for individuals who did not receive the treatment at least up to time u . The parameter of interest for each u is the average effect in t for individuals starting a program in period u of their welfare spell compared and not relative effects of comparing one type of short-term training with another. Therefore, we forgo the distinction of J different available programs in the description.

to not joining at u :

$$\begin{aligned}
\Delta_{t,u}^{ATT} &= E(Y_{t,u}^1 - Y_{t,u}^0 | D_u = 1, D_1 = \dots = D_{u-1} = 0) \\
&= E(Y_{t,u}^1 | D_u = 1, D_1 = \dots = D_{u-1} = 0) \\
&\quad - E(Y_{t,u}^0 | D_u = 1, D_1 = \dots = D_{u-1} = 0).
\end{aligned} \tag{1}$$

Whereas the first term is identified in the data by the observed outcome of the participants, the second term has to be estimated. Simply using the observable non-participants' outcomes to approximate the unobservable participants' outcomes without treatment may lead to biased estimates due to self-selection.

To solve the selection problem we apply a propensity score matching estimator. The basic idea of the matching approach is to find, in a large group of non-participants, those individuals who are similar to the participants in all relevant pre-treatment characteristics X ('statistical twins'). However, it is well known that matching can become hazardous when X is of high dimension. To deal with this dimensionality problem, Rosenbaum and Rubin (1983) suggest the use of the propensity score $p(X) = E(D = 1|X)$, i.e. the probability of participation in a program, summarizing the information of the relevant covariates X into a single index function. However, for the ATT to be identified with matching, the so-called *conditional independence assumption* (CIA, $Y^0 \perp\!\!\!\perp D|X$ in the static binary case, Lechner, 1998) has to be imposed. It states that, conditional on the set of relevant (observable) covariates X , the non-participation outcome Y^0 is independent of the participation decision.

For the dynamic case, we have to invoke an adjusted version, the *dynamic conditional independence assumption* (DCIA):

$$Y_{t,u}^0 \perp\!\!\!\perp D_u | p(X_u), D_1 = \dots = D_{u-1} = 0, \tag{2}$$

i.e. the hypothetical outcome at time t after not participating up to time u is independent of program participation at time u , conditional on the propensity score $p(X_u)$ measured at time u . The DCIA ensures that treated and non-treated individuals are comparable in their non-treatment outcomes at time t conditional on $p(X_u)$, conditional on claiming welfare benefits up to time $u - 1$, and conditional on not receiving treatment before u . In addition, the availability of non-participating analogues for the participants must be guaranteed (*common support*), i.e. $Pr(D = 1|X_u) < 1$ (Smith and Todd, 2005a).

5.2 Implementation

For the DCIA to hold, it is necessary to observe all covariates that, conditional on having spent a given welfare duration u , jointly influence the participation decision at that time (D_u) and the outcome variable where such a decision is postponed further ($Y_{t,u}^0$). In line with that, we condition on previous welfare experience by stratifying the welfare duration in quarters. Using this kind of aggregation is useful for consideration of differences due to the timing of treatments since we expect the probabilities of entering a program or employment to remain relatively constant within quarters of the welfare spell. For the propensity scores, we have estimated separate probit models for each group, each treatment, gender, and the first four quarters of welfare receipt. Each probit estimates the probability of starting a program in quarter u , conditional on X , conditional on having reached the welfare duration of $u \in \{1, \dots, 4\}$ quarters, and conditional on not having received a treatment before u in the welfare spell. Hence, we analyze the effects of a training program for groups of individuals that join within the first year of the welfare spell. The outcomes are measured monthly from the first month of the sequent quarter after (potential) participation onwards until July 2008 due to the time horizon of the analysis.⁷ The treatment effects are estimated using kernel density matching on the estimated propensity score. Standard errors were calculated by bootstrapping with 250 replications.

With regard to the variables selected as relevant to solve the potential self-selection bias, the comprehensive data at hand provides a sufficient basis. In the empirical specification of the propensity score models, we use 21 categories of variables comprising socio-demographic information like age, marital status, or the number of children, the qualification of the individual and information characterizing the employment, unemployment, and welfare history of the participants dating in some cases dating back until 1990. The specifications for the final models used in the estimations were obtained by estimating probit regressions starting with the full set of variables and a stepwise dropping of jointly insignificant variable-blocks (indicated by F -tests) in order to provide a parsimonious specification. For this reason, the model specifications vary across the probit models estimated

⁷For programs assigned in the first quarter of the welfare spell we have an observation period of at least 16 months for each observation. The last entry into the welfare system in our sample is December 31st, 2006. Thus, a program in the first quarter could be assigned until March 31st, 2007. In this case, the observation period for the outcomes is April 2007 until July 2008. Consequently, for programs assigned in the second quarter we have an observation period of 13 months. In the third quarter the observation period lasts for 10 months and in the fourth quarter for 7 months.

for ethnic groups, quarters of program start, gender, and the programs considered.

The estimated propensity score should guarantee that the included variables are balanced between treatment and comparison group. To check the balancing property of the estimated propensity score, we applied a procedure suggested by Smith and Todd (2005b):

$$\begin{aligned}
X_{k_u} = & \beta_0 + \beta_1\hat{p}(X_u) + \beta_2\hat{p}(X_u)^2 + \beta_3\hat{p}(X_u)^3 + \beta_4\hat{p}(X_u)^4 \\
& + \beta_5D + \beta_6D\hat{p}(X_u) + \beta_7D\hat{p}(X_u)^2 + \beta_8D\hat{p}(X_u)^3 + \beta_9D\hat{p}(X_u)^4. \quad (3)
\end{aligned}$$

Eq. (3) was estimated for each variable X_{k_u} included in the propensity score of program participation in quarter u . Afterwards, the null hypothesis of β_5 to β_9 being jointly zero was tested. The test indicates, whether there are differences due to the treatment indicator conditional on a quartic polynomial of the propensity score. If ideal balancing is achieved all those coefficients should be zero.

Obviously, caseworkers play a crucial role in the process of assignment to programs. Turning down a placement could be sanctioned by benefit revocation and, hence, caseworkers can be assumed to have the final word in the participation decision. If the caseworkers act on unobservable information that is correlated with the individual's potential labor market outcomes, the DCIA would be violated. However, it is not very likely that caseworkers have referred to further unobservable information than the large set of variables recorded. The data used in this analysis were collected by the caseworkers and supplemented by their own subjective assessment of the qualification and placement restrictions of the individuals. Moreover, it should be noted that - to bias the estimates - any further unobserved information has to jointly influence the participation decision and the outcomes. Given the large set of variables we considered relevant and we controlled for in the estimations, we assume that caseworkers act idiosyncratically given the observable characteristics of the individuals and the subjective assessments.

For interpretation of the results, one has to bear in mind that the chosen comparison group does not reflect a no-program state, but rather possibly postponed participation. If we choose as the comparison group those individuals who have been observed to never participate in the data, this may invalidate the DCIA, as we have to condition on future outcomes. For unbiased estimation we have to rule out anticipatory effects, else people would behave differently conditional on future outcomes or treatments. If for example, non-participants would know in advance to be treated later and when, then matching

could not solve the selection problem and we would overestimate the treatment effect since the non-participants have no reason to leave welfare instantly for work. In contrast, if people dread the prospect of being treated and, again, they know when to be treated in the future they will leave for work and the program effect is underestimated since non-participants would differ significantly even after matching from the participants. However, it is important to note that this is only the case if people know exactly that they will be treated and when. In line with that, Abbring and van den Berg (2003) point out that the exclusion of anticipatory effects does not rule out that the individuals know and act on the determinants of assignment to treatment or labor market outcomes, i.e. individuals are allowed to adjust their optimal behavior to the determinants of the treatment process, but not to realization of the treatment. This is not a problem for the analysis as long as treated and non-treated individuals anticipate the chances of these events conditional on propensity score and the elapsed welfare duration in a certain quarter in the same way. Hence, with respect to the assignment process during the individual welfare spell people may know the determinants, but it is unlikely that they know the realizations of the future events. For that reason, we assume our estimates not to be affected by anticipatory effects.

5.3 Decomposition of Differences in Treatment Effects

Considering effect heterogeneity in the treatment effects between ethnic groups for a particular program can be used to reveal important insights. Assuming that identical programs are provided, differences could be, on the one hand, due to differences in the composition of the groups, i.e. the distribution of characteristics that are relevant for program and labor market success may be different. Hence, when conditioning on all these variables no further differences should occur. However, on the other hand, if residual differences would remain between the compared ethnic groups these differences are solely due to the ethnic group attachment of the individual and might be interpreted as potential discrimination. An important question in the context of providing ALMP for immigrants is whether potential discrimination is identified as the unexplained part of the gap in the difference of the treatment effects. To analyze the extent of the potential discrimination, we suggest and apply the following decomposition procedure.

To abbreviate notation, we suppress the indicators of the dynamic setting. Starting point for the decomposition is the raw differential Δ_{Dif}^{ATT} of the differences in the ATTs between

immigrants and native Germans:

$$\Delta_{Dif}^{ATT} = \Delta_{Mig}^{ATT} - \Delta_{nG}^{ATT}, \quad (4)$$

with

$$\Delta_{Mig}^{ATT} = E\left(Y^1 - Y^0 | X_{Mig}, D = 1\right) \Big|_{Mig=1}, \quad (5)$$

and

$$\Delta_{nG}^{ATT} = E\left(Y^1 - Y^0 | X_{nG}, D = 1\right) \Big|_{Mig=0}. \quad (6)$$

Here, Δ_{Mig}^{ATT} denotes the ATT for the immigrants and Δ_{nG}^{ATT} is the ATT for the native Germans who participated in the program under consideration. Both ATTs were estimated according to the procedure described in the previous section.

To highlight the differences in the raw differential, we have added the relevant conditions in eq. (5) and (6). Mig is a dummy variable taking value 1 if the group of interest are immigrants, and 0 if native Germans are considered. Moreover, the ATT of the immigrants (eq. 5) is conditional on the observable characteristics X_{Mig} of the participating immigrants and the ATT for the native Germans (eq. 6) is conditional on the characteristics X_{nG} of the participants in that group.

Accordingly, we could decompose the raw differential in eq. (4) into a part which is explained by differences in observable characteristics and a residual part which cannot be explained by observables:

$$\Delta_{Dif}^{ATT} = \Delta_{explained}^{ATT} + \Delta_{residual}^{ATT}. \quad (7)$$

The first term on the right-hand side denotes the part of the difference in the ATTs for immigrants and native Germans that is explained by differences in observable characteristics (e.g. due to a different age or qualification structure). This part is defined as

$$\Delta_{explained}^{ATT} = E\left(Y^1 - Y^0 | X_{Mig}, D = 1\right) \Big|_{Mig=0} - E\left(Y^1 - Y^0 | X_{nG}, D = 1\right) \Big|_{Mig=0}. \quad (8)$$

It is the difference in ATTs for the native participants when conditioning first on the observable characteristics X_{Mig} of the participating immigrants and second on the observables X_{nG} of the participating native Germans. If X_{Mig} and X_{nG} are identical, $\Delta_{explained}^{ATT}$ will be 0 and the difference in ATTs for immigrants and natives is not attributable to differences in observables between the two ethnic groups. However, if $X_{Mig} \neq X_{nG}$, then $\Delta_{explained}^{ATT}$ will in general be nonzero and measure differences in ATTs between immigrants and natives due to observable characteristics.

The second term on the right-hand side of eq. (7) denotes the difference in the ATTs for immigrants and native Germans that is solely due to unobservable differences between the two groups. Holding the observable characteristics constant, i.e. assuming all individuals to possess the characteristics X_{Mig} of the immigrants, the difference is defined as:

$$\Delta_{residual}^{ATT} = E\left(Y^1 - Y^0 | X_{Mig}, D = 1\right) \Big|_{Mig=1} - E\left(Y^1 - Y^0 | X_{Mig}, D = 1\right) \Big|_{Mig=0}. \quad (9)$$

It is the difference in ATTs between immigrants and natives when conditioning in both cases on the covariates X_{Mig} of the participating immigrants. If covariates X_{Mig} are valued equally in both ethnic groups, then $\Delta_{residual}^{ATT} = 0$ and the difference in ATTs does not depend on unobservable characteristics. However, if covariates X_{Mig} are valued differently, then $\Delta_{residual}^{ATT}$ is non-zero and measures the unexplained part of the raw differential Δ_{Dif}^{ATT} .

Thus, the proposed decomposition of the differences in the treatment effects is similar to a difference-in-differences estimator. It allows the ceteris paribus identification of the difference in program effects that is due to variation in observable characteristics, i.e. differences in the composition of the immigrant and native participants in the particular program, and of the part that is due to belonging to the immigrant group. The latter relates to unobservable differences between immigrants and native Germans. We will refer to this part as an immigrant fixed effect.

To estimate the difference that is due to unobservable differences (eq. 9), we have to match participating immigrants with comparable participating native Germans, i.e. $X_{Mig} = X_{nG}$. To do so, we apply a matching procedure similar to that described above. In the first step, we estimate the ATTs separately for both ethnic groups and all considered training programs. In the second step, we keep only the participants in each sample and match treated immigrants and treated native Germans conditional on the distribution of the observable characteristics of the treated immigrants. Outcome variable in this matching step is the individual treatment effect from training for each participant. Therefore, the resulting effect of the second matching step gives us the average difference in program effects between immigrants and natives which is due to the immigrant fixed effect keeping all observable characteristics constant. Analogously to the estimation of the program effects, we also use a kernel density matching estimator for the second matching step.

6 Empirical Results

6.1 Quality of the Estimates

For the estimation of program effects we stratify our data by ethnic group, gender and quarter of program start. In total we are able to estimate treatment effects for 51 different strata.⁸ To obtain valid treatment effects it is crucial that the covariates included in the propensity score estimation are balanced between treatment and comparison group after matching. As a balancing test we apply the procedure suggested by Smith and Todd (2005b). Results of this test are summarized in Table 2. The test is passed in 95% or 2,355 of 2,481 cases at the 1% significance level. Thus, balancing is not ideal in every case but sufficient to obtain valid treatment effects. The matching quality is similar for men (95%, 1,387 of 1,460 tests passed) and women (95%, 968 of 1,021) as well as for natives (95%, 1,151 of 1,217) and immigrants (95%, 1,204 of 1,264). Even at the 5% level 2,270 of the total 2,481 tests are passed and 2,184 at the 10% level.

Include Table 2 about here

The exact specifications of the estimated 51 propensity scores cannot be presented here, but are available upon request from the authors. Results reveal, that especially age, educational attainment, professional qualification, household composition, region, and employment history within the last six years, in particular during the last 24 months before entering the welfare system, are relevant factors that must be accounted for when estimating the effects of short-term training programs. It turns out that these covariates are also important in the second matching step when decomposing differences in training effects between immigrants and natives. In this matching step, we detail the employment history even further resulting in a large number of variables for the final specification of the propensity scores. As can be seen from Table 3, covariates are balanced very well and matching quality is of the same high degree as in the first matching step.

Include Table 3 about here

⁸In 13 strata the number of treated individuals is too small to estimate valid treatment effects. See Table 2 for the affected strata.

6.2 Program Effects

The estimated program effects and corresponding t-values are displayed in Tables 4 to 7. The effects are estimated separately for natives and immigrants, for men and women and for each quarter. As can be seen from Table 4, aptitude tests have a positive impact on the probability of native and immigrant males to take up employment providing a sufficient income above the subsistence level. This positive impact is independent of the exact timing of the training. However, the absolute size and the significance of the effects differ across quarters and also between the two considered ethnic groups. For aptitude tests starting in the first quarter of welfare receipt, we observe larger treatment effects for natives than for immigrants. Three months after starting the program, the average treatment effect on the treated for male natives amounts to 6.68 percentage points, whereas the corresponding value for men with migration background is 4.64 percentage points. Thus, native participants in aptitude tests have a nearly 7 percentage points larger probability to take up employment providing a sufficient income than in a situation without training. Even though the estimated effect for immigrants is about 2 percentage points lower, it is still of considerable size. During the middle of our observation period the estimated effect for immigrants increases up to 7 percentage points, but then slightly decreases afterwards. One year after program start, it amounts to 6.53 percentage points. For native males, we observe treatment effects ranging between 8 and 9 percentage points six to nine months after starting the program and an effect of 9.39 percentage points one year after the training. Consequently, at the end of the observation period the difference in treatment effects between natives and immigrants is somewhat larger than at the beginning.

Include Table 4 about here

In contrast to the first quarter, we observe that for aptitude tests starting in the second quarter after the inflow into welfare, the treatment effects are larger for immigrants than for natives. During the whole observation period, the estimated effect for immigrants amounts to slightly more than 10 percentage points, whereas the corresponding estimate for natives ranges between 7.6 and 9.7 percentage points. Thus, for men with migration background aptitude tests starting in the second quarter of welfare dependency are more effective than tests starting immediately after the inflow into welfare, while for natives almost no difference between the first two quarters is detected. In the third quarter, the

effectiveness of aptitude tests further increases for immigrants. Nine months after starting the program, treated immigrants have a 15.7 percentage points larger probability to find employment than without the training. In contrast, treatment effects for natives are lower than in the first two quarters. At the end of the observation period, the estimated treatment effect amounts to 6.4 percentage points and is only slightly significant. In the fourth quarter, the picture is again reversed. We now find large treatment effects with a magnitude of more than 15 percentage points for men without migration background, while the effect for immigrants is lower and amounts to about 10 percentage points.

For women we also find positive effects of aptitude tests, but observe a more uniform pattern of the estimated effects. Irrespective of the quarter of program start, native females profit more from aptitude tests than women with a migration background. In the first quarter, the probability to take up a job providing a sufficient income within one year after the training increases for a female native participant by about 9.5 percentage points. In contrast, female immigrants participating in aptitude tests face only a slightly significant increase of 5.2 percentage points. In the second quarter all estimated employment effects are insignificant for immigrant females. For native women employment effects are highly significant and amount to more than 10 percentage points. Similar employment effects for native females are found in the third quarter. In this quarter, employment effects are also positive for female immigrants but somewhat lower in magnitude compared to natives and only slightly significant.

Include Table 5 about here

In contrast to aptitude tests, job search training is rather ineffective (see Table 5). For native males we estimate insignificant employment effects in all considered quarters. For male immigrants the estimated effects are also insignificant in the second and third quarter. Only in the first quarter we observe significantly positive effects nine months after starting the program of about 6.4 percentage points. However, this value represents a maximum only and does not describe a long lasting effect.

For female immigrants participating in job search training during the first quarter of their welfare spell, we observe negative employment effects at the beginning of the observation period. The probability to take up a job and thereby to leave the welfare system is reduced by about 5 percentage points in the first six months after the program starts. After this

locking-in period, the sign of effects turns positive, but the estimates are not statistically significant. In contrast, the corresponding figures for native women are positive and statistically significant throughout the whole observation period. In the first nine months after training the treatment effect amounts to about 6 percentage points and then rises further to 9 percentage points one year after program start. In the second quarter, employment effects for native women are significantly positive in the middle of the observation period reaching a maximum of 12 percentage points nine months after program start, but then slightly decrease. For female immigrants no treatment effects could be obtained for the second quarter since the number of treated individuals was too low.

Contrary to job search training, skill provision seems to be more effective (see Table 6). For native men who start training in the first quarter of the welfare spell, we observe positive employment effects. The training increases the probability of participants to take up a job and to leave the welfare system by more than 10 percentage points in the second half of the observation period. For male immigrants we also observe positive employment effects. However, compared to native men effects are smaller in magnitude and only slightly significant. One year after program start the probability to find a job is increased by 6 percentage points. In the remaining quarters, employment effects are mostly insignificant for natives and immigrants. Only for immigrants participating in the third quarter, we observe increasingly positive effects during the observation period.

Include Table 6 about here

In contrast to the picture observed for men, we find for women and the first quarter that skill provision is more effective among participants with migration background than for natives. While for native women employment effects amount to nearly 6 percentage points one year after program start, we estimate considerably larger effects for immigrants of about 12.6 percentage points. As opposed to this picture, we find for the second quarter that employment effects for female immigrants are insignificant, whereas effects are significantly positive for female natives. For this group, the probability to take up a job increases by 12.5 percentage points six months after program start and then remains on a level of about 10 percentage points until the end of the observation period. In the third quarter all estimated effects are positive but insignificant, while we observe increasingly negative treatment effects throughout the whole observation period in the fourth quarter.

Even though some training modules show a positive impact on the probability to take up employment, the combination of two or three modules in one program is rather ineffective (see Table 7). For native men we find no significant effect, irrespective of the quarter considered. A similar picture arises for men with migration background for the first and second quarter. However, in the third quarter we find increasingly positive employment effects for this group during the observation period albeit the degree of statistical significance is low. Nine months after program start the estimated effect amounts to 7.4 percentage points, but the effect is significant only at the 10% level. A similar development of the estimated treatment effects is found for immigrant females participating in combined training programs in the first quarter of the welfare spell. Here, the effect amounts to 6.6 percentage points at the end of the observation period. In the second quarter the corresponding estimate is slightly larger amounting to 7.9 percentage points, but again the effect is significant only at the 10% level. In the third quarter, we do not detect any significant effect. For women without migration background, we estimate insignificant employment effects in all considered quarters. The general ineffectiveness of combined training programs might be due to the characteristics of the targeted group. As has been noted above, combined training programs are in particular assigned to those persons who were out of labor force for a substantial fraction of the final two years before treatment. These persons are likely to face multiple obstacles for employment uptake, which might not be remediable by combined training programs.

Include Table 7 about here

To summarize our results, we find pronounced differences in the effectiveness of the considered training programs. For aptitude tests we observe positive employment effects. While in the case of women treatment effects are larger for natives than for immigrants, the picture is ambiguous for men depending on the timing of the training. In some quarters men with migration background profit more from the training, while in others native men do better. Skill provision also shows positive employment effects especially when it takes place right at the beginning of the welfare spell. As in the case of aptitude tests, treatment effects differ between natives and immigrants and between men and women. For women and the first quarter we observe that this form of training is more effective among participants with migration background. In contrast, for men and the first quarter we find that natives do better. Irrespective of the migration background the effect of job search training

is rather limited. Here, we only observe significantly positive employment effects for male immigrants nine months after program start and significantly positive employment effects for native females throughout the whole observation period in the first quarter of the welfare spell. Similarly, the combination of all three training modules is quite ineffective. While neither native men nor women benefit from the combined training program, we only find slightly significantly positive employment effects for men with migration background in the third quarter and for immigrant females in the first two quarters.

Our results are in line with previous empirical evidence. Huber et al. (2009) do not distinguish between different training modules, but find on average positive employment effects of short-term training. These positive effects might be driven by aptitude tests and skill provision. Kopf (2009) uses in her study a distinction of training programs which is similar to the one used by us. She also estimates positive employment effects of aptitude tests and skill provision, while job search training and combined training programs seem to be rather ineffective. However, our estimated treatment effects for aptitude tests and skill provision are somewhat larger than those reported by her. This might be due to the different time horizon of the analysis (2006 in our case vs. 2005 in Kopf, 2009) or the different sampling of the data. While we use an inflow sample, the study by Kopf (2009) is based on a stock sample. Wolff and Jozwiack (2007) use the same data as Kopf (2009) but do not distinguish between different training modules. They find that short-term off-the-job training programs are on average ineffective among immigrant men, while they significantly decrease employment chances of women with migration background in the short-run. Our results indicate that the adverse effects for female immigrants might be caused by job search training. However, we also find that aptitude tests and skill provision exhibit positive employment effects in this subgroup. So, training programs do not in general reduce employment chances of women with migration background. The same is true for male immigrants. In this subgroup we also find positive employment effects of aptitude tests and skill provision.

6.3 Decomposition Results

The previous subsection showed that the treatment effects of the considered training programs differ between native and immigrant participants. Therefore, the question arises what might cause these differences. Are they due to differences in the observable charac-

teristics of the two groups or are they due to unobservable differences subsumed in the immigrant fixed effect? To disentangle the influence of both possible explanations we decompose the differences in the treatment effects between natives and immigrants in two parts: the part which is caused by differences in observables and the residual part due to the immigrant fixed effect. Differences due to the immigrant fixed effect are of major policy concern, since in this case discrimination in the effectiveness of training programs is present. Therefore, in the following we concentrate on differences in treatment effects due to unobservables. Table 8 displays the relevant results.⁹

The first row of each block in the table depicts the raw differential of differences in the ATTs between immigrants and native Germans for the respective program. This raw differential is calculated from the results presented in the previous subsection. The p-value denotes statistical significance of the difference in ATTs of natives and immigrants. The third row of each block in the table is denoted by $\Delta_{residual}^{ATT}$ and shows the estimated part of the raw differential which is due to the immigrant fixed effect. In other words, $\Delta_{residual}^{ATT}$ indicates by how much the treatment effect of a program is changed due to the migration background holding all other factors fixed.

The entry 0.0135 in the top left block of Table 8 states that male immigrant participants in an aptitude test have on average a 1.35 percentage points larger treatment effect concerning employment uptake three months after program start than native participants with identical sociodemographic characteristics. Therefore, immigrants benefit more from aptitude tests than natives holding all other characteristics constant. However, this immigrant fixed effect is not statistically significant as can be seen from the t-value. Six months after program start, we observe a similar picture. The immigrant fixed effect is positive but insignificant. During the following months, the immigrant fixed effect increases and reaches a maximum of 6.39 percentage points nine months after program start. In this month, the effect is statistically significant at the 10% level. At the end of the observation period, the immigrant fixed effect decreases and amounts to about 3 percentage points one year after assignment. Even though the effect again lacks statistical significance it is of considerable size. Thus, the difference in program effectiveness between immigrants and natives is not insubstantial.

For women participating in aptitude tests we arrive at a similar conclusion. Female natives

⁹We only decompose the differences in the treatment effects for the first quarter, since in the other quarters the number of program participants is too small.

have a larger treatment effect than immigrants resulting in a negative raw differential. The negative sign of the raw differential is driven by differences in observable characteristics between the two ethnic groups. Keeping all covariates constant, immigrants tend even to benefit more from aptitude tests than natives as indicated by the positive immigrant fixed effect. However, again this positive immigrant fixed effect is not statistically significant although it is of considerable size.

When looking at job search training targeted at male welfare recipients, we do not estimate statistically significant immigrant fixed effects. In contrast, we observe for women that in the short-run immigrant participants clearly benefit less than natives from this form of training. The negative raw differential indicates that the treatment effect for immigrants is smaller than for natives. As can be seen from the p-value, differences in ATTs between both groups are statistically significant during the first six months of the observation period. The large gap in treatment effects is caused by the immigrant fixed effect. Keeping everything else constant, immigrant females participating in job search training have a nearly 15 percentage points lower treatment effect than native participants.¹⁰ Even though the negative immigrant fixed effect fades away nine months after program start, this is an alarming result. Job search training decreases rather than increases employment chances of female immigrants.

In contrast, female immigrants clearly benefit from skill provision. For this form of training the immigrant fixed effect increases over time and amounts to 14 percentage points one year after program start. Thus, netting out observable differences between immigrants and natives, the former have on average a 14 percentage points larger treatment effect than the latter.¹¹ For men, we do not find statistically significant differences between the two ethnic groups. Here, as opposed to women, the immigrant fixed effect is of negative sign at the end of the observation period indicating that immigrants tend to benefit less from the training than natives.

For the combined training programs, we do not find statistically significant differences either. While the immigrant fixed effect tends to be negative for men, it tends to be positive for women. These results reflect the finding that combined programs are rather ineffective for both ethnic groups and for both genders.

¹⁰Due to large standard errors, the immigrant fixed effect is only slightly significant in the first three months after program start.

¹¹Again, due to large standard errors, the immigrant fixed effect is only slightly significant.

Include Table 8 about here

To sum up, we find that aptitude tests yield larger treatment effects for natives than for immigrants. The lower treatment effects for immigrants are due to differences in observable characteristics. Keeping all observables constant, immigrants even tend to benefit more from this training module than identical natives. Job search training impacts similarly on native and immigrant men, once it is controlled for all observable differences. In contrast, female immigrant participants are clearly disadvantaged. Job search training does not seem to be designed for this group and might not meet the needs of female immigrants. Better suited courses are needed for this group. In contrast, skill provision seems to meet the needs of female immigrants, who benefit more from this program than identical natives. However, this result might indicate that caseworkers undervalue the productivity of immigrants but learn from the training about unexpected opportunities to place the treated. This might, of course, be beneficial for the treated but gives concern to a poor performance of the untreated. The effectiveness of combined training programs does not differ significantly between immigrants and natives with identical characteristics.

7 Conclusion

Based on comprehensive administrative data on immigrant and native welfare recipients in Germany, we have evaluated the employment effects of four different short-term off-the-job training programs. In particular, we have investigated whether program effects differ between the two ethnic groups and what causes these differences.

Our estimation results show that the considered training programs exhibit substantial effect heterogeneity. For aptitude tests we observe on average positive employment effects. While in the sample of women treatment effects are larger for natives than for immigrants, the picture is ambiguous for men depending on the timing of the training. Aptitude tests starting in the second or third quarter of welfare receipt generate larger treatment effects for immigrants, whereas native men benefit more from tests in the first and fourth quarter. The difference in ATTs of natives and immigrants in the first quarter is mainly due to differences in observable characteristics between the two ethnic groups. Keeping all covariates constant, immigrants tend to benefit even more from aptitude tests than natives.

Job search training is ineffective for men irrespective of the migration background. Native women benefit from this form of training, while immigrant females face negative treatment effects. The large difference in treatment effects of native and immigrant women cannot be explained by observable characteristics and must instead be attributed to the immigrant fixed effect. Holding everything else constant, immigrant females participating in job search training have a nearly 15 percentage point lower treatment effect than native participants. Even though the negative immigrant fixed effect fades away nine months after program start, this result gives cause for serious concern. Job search training decreases rather than increases employment chances of female immigrants. Therefore, immediate action has to be taken by welfare agencies to develop better suited programs.

In contrast, female immigrants clearly benefit from skill provision, which is a program exhibiting positive effects in general when assigned early during the welfare spell. For this form of training the immigrant fixed effect increases over time and amounts to 14 percentage points one year after the program begins. Thus, netting out observable differences between immigrants and natives, the former have on average a 14 percentage points larger treatment effect than the latter. Caseworkers might undervalue the productivity of female immigrants, but could learn from the training and intensify their placement effort once they are aware of the true potential of the treated. This could also explain the positive immigrant fixed effect observed for aptitude tests. However, for men participating in skill provision we do not find statistically significant differences between the two ethnic groups.

For the combined training programs, we do not find statistically significant differences in any subgroup. These results reflect the finding that combined programs are rather ineffective for both ethnic groups and for both genders. The general ineffectiveness of combined training programs might be due to the characteristics of the targeted group, since combined programs are in particular assigned to those persons who were out of the labor force for a substantial fraction of the final two years before treatment. These persons are likely to face multiple obstacles for employment uptake, which might not be remediable by combined training programs.

Our results reveal that not all training programs impact equally on native and immigrant welfare recipients. Differences are especially pronounced for women participating in job search training and skill provision. While immigrants benefit more than natives from skill provision, they are clearly disadvantaged by job search training. The underlying reasons

for this opposing trend cannot be pinned down on the basis of our analysis. Further research is also needed with respect to the large size of standard errors in the decomposition which prevents us in some cases to detect significant immigrant fixed effects albeit effects are of considerable size. One explanation for this could be a high degree of heterogeneity in the immigrant group (despite having controlled for many observed characteristics). For example, the immigrant fixed effect could differ across various country-of-origin groups. Possibly, considering more homogenous ethnic groups could shed more light onto this problem. On the other hand, our results also indicate a general potential for welfare agencies to improve the targeting of programs at participants based on observable characteristics. Again, further research is needed to advise welfare agencies on how to use programs in the most effective and efficient way. This should contribute to reducing the high number of welfare recipients in Germany.

Acknowledgements

The paper has benefited from discussions at the ZEW Workshop on Evaluation of Policies Fighting Social Exclusion in Mannheim, 2010, and the Conference on Migration and Migration Policy at University of Maastricht, 2010. We wish to thank Moritz Hennig and Ann-Kathrin Köbler for helpful research assistance. Financial support from the project *Evaluation of the Effects of Basic Social Care for Job Seekers with Migration Background in Germany* commissioned by the Federal Ministry of Labor and Social Affairs (BMAS) is gratefully acknowledged. Stephan L. Thomsen thanks the Stifterverband für die Deutsche Wissenschaft (Claussen-Simon-Stiftung) for financial support.

References

- ABBRING, J. H., AND G. J. VAN DEN BERG (2003): “The Nonparametric Identification of Treatment Effects in Duration Models,” *Econometrica*, 71(5), 1491–1517.
- BIEWEN, M., B. FITZENBERGER, A. OSIKOMINU, AND M. WALLER (2007): “Which Program for Whom? Evidence on the Comparative Effectiveness of Public Sponsored Training Programs in Germany,” Discussion Paper No. 07-42, ZEW.
- Bundesagentur für Arbeit (2006): “Arbeitsmarkt 2005,” Amtliche Nachrichten der Bundesagentur für Arbeit 54, Nuremberg.
- Bundesagentur für Arbeit (2007): “Arbeitsmarkt 2006,” Amtliche Nachrichten der Bundesagentur für Arbeit 55, Nuremberg.
- Bundesagentur für Arbeit (2008): “Arbeitsmarkt 2007,” Amtliche Nachrichten der Bundesagentur für Arbeit 56, Nuremberg.
- Bundesagentur für Arbeit (2009a): “Arbeitsmarkt 2008,” Amtliche Nachrichten der Bundesagentur für Arbeit 57, Nuremberg.
- Bundesagentur für Arbeit (2009b): “Jahresbericht 2008,” SGB II - Sozialgesetzbuch Zweites Buch - Grundsicherung für Arbeitsuchende 3, Nuremberg.
- Bundesministerium für Arbeit und Soziales (2009): “Wirkungen des SGB II auf Personen mit Migrationshintergrund. Abschlussbericht. Hauptband.,” Forschungsbericht 395, Berlin.
- CLAUSEN, J., E. HEINESEN, H. HUMMELGARD, L. HUSTED, AND M. ROSHOLM (2009): “The Effect of Integration Policies on the Time Until Regular Employment of Newly Arrived Immigrants: Evidence from Denmark,” *Labour Economics*, 16, 409–417.
- COHEN-GOLDNER, S., AND Z. ECKSTEIN (2009): “Estimating the Return to Training and Occupational Experience: The Case of Female Immigrants,” *Journal of Econometrics*, doi:10.1016/j.jeconom.2009.09.009.
- FREDRIKSSON, P., AND P. JOHANSSON (2008): “Dynamic Treatment Assignment - The Consequences for Evaluations Using Observational Data,” *Journal of Business & Economics Statistics*, 26(4), 435–445.

- FRÖHLICH, S., S. KAIMER, AND M. STAMM (2004): “Beschreibung und Qualitätsanalyse – Verbleibsnachweis in Beschäftigung,” Internal Working Paper, IAB, Nuremberg.
- HÄMÄLÄINEN, K., AND M. SARVIMÄKI (2008): “Moving Immigrants from Welfare to Work: The Impact of an Integration Program,” unpublished paper, Government Institute for Economic Research (VATT), Helsinki.
- HECKMAN, J. J., R. J. LALONDE, AND J. A. SMITH (1999): “The Economics and Econometrics of Active Labor Market Programs,” in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3A, chap. 31, pp. 1865–2097. Elsevier, Amsterdam.
- HUBER, M., M. LECHNER, C. WUNSCH, AND T. WALTER (2009): “Do German Welfare-to-Work Programmes Reduce Welfare and Increase Work?,” Discussion paper no. 2009-03, University of St. Gallen.
- HUJER, R., AND S. L. THOMSEN (2010): “How Do Employment Effects of Job Creation Schemes Differ with Respect to the Foregoing Unemployment Duration,” *Labour Economics*, 17(1), 38–51.
- HUJER, R., S. L. THOMSEN, AND C. ZEISS (2006): “The Effects of Short-Term Training Measures on Individual Unemployment Duration in Western Germany,” Discussion Paper No. 06-065, ZEW, Mannheim.
- IMBENS, G. W. (2000): “The Role of the Propensity Score in Estimating Dose-Response Functions,” *Biometrika*, 87(3), 706–710.
- JACOBI, L., AND J. KLUVE (2007): “Before and After the Hartz Reforms. The Performance of Active Labour Market Policy in Germany,” Discussion paper 41, RWI Essen.
- KLUVE, J. (2006): “The Effectiveness of European Active Labor Market Policy,” Discussion Paper No. 2018, IZA.
- KOPF, E. (2009): “Short-term Training Variety for Welfare Recipients,” Iab discussion paper 17/2009, IAB Nuremberg.
- LECHNER, M. (1998): *Training the East German Labour Force. Microeconomic Evaluations of Continuous Vocational Training after Unification*. Physica-Verlag, Heidelberg.

- LECHNER, M. (2002): “Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods,” *Journal Of The Royal Statistical Society Series A*, 165(1), 59–82.
- LECHNER, M., AND C. WUNSCH (2008): “What Did All the Money Do? On the General Ineffectiveness of Recent West German Labour Market Programmes,” *Kyklos*, 61(1), 134–174.
- MORTENSEN, D. (1986): “Job Search and Labor Market Analysis,” in *Handbook of Labor Economics Vol.II*, ed. by O. Ashenfelter, and D. Card, pp. 849–919. Elsevier, Amsterdam.
- NEYMAN, J. S. (1923): “On the Application of Probability Theory to Agricultural Experiments. Essays on Principles. Chapter 9,” *Statistical Science*, 5(4), 463–480, Reprint in English from 1990.
- OECD (2008): “International Migration Outlook: SOPEMI - 2008 Edition,” Social Issues/Migration/Health 14, Paris.
- ROSENBAUM, P. R., AND D. B. RUBIN (1983): “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, 70(1), 41–50.
- ROSHOLM, M., AND R. VEJLIN (2010): “Reducing Income Transfers to Refugee Immigrants: Does Start-help Help You Start?,” *Labour Economics*, 17, 258–275.
- ROY, A. D. (1951): “Some Thoughts on the Distribution of Earnings,” *Oxford Economic Papers*, 3(2), 135–145.
- RUBIN, D. B. (1974): “Estimating Causal Effects to Treatments in Randomised and Nonrandomised Studies,” *Journal of Educational Psychology*, 66, 688–701.
- SIANESI, B. (2004): “An Evaluation of the Active Labour Market Programmes in Sweden,” *The Review of Economics and Statistics*, 86(1), 133–155.
- SMITH, J. A., AND P. TODD (2005a): “Does Matching Overcome LaLonde’s Critique of Nonexperimental Estimators?,” *Journal of Econometrics*, 125(1/2), 305–354.
- (2005b): “Does Matching Overcome LaLonde’s Critique of Nonexperimental Estimators? Rejoinder,” *Journal of Econometrics*, 125(1/2), 365–375.

- Statistisches Bundesamt (2006): *Leben in Deutschland. Haushalte, Familien und Gesundheit - Ergebnisse des Mikrozensus 2005*. Statistisches Bundesamt, Wiesbaden.
- THOMSEN, S. L. (2007): *Evaluating the Employment Effects of Job Creation Schemes in Germany*, vol. 36 of *ZEW Economic Studies*. Physica-Verlag, Heidelberg.
- VAN DEN BERG, G. J. (1994): "The Effects of Changes of the Job Offer Arrival Rate on the Duration of Unemployment," *Journal of Labor Economics*, 12(3), 478–498.
- WOLFF, J., AND E. JOZWIACK (2007): "Does short-term training activate means-tested unemployment benefit recipients in Germany?," Discussion paper no. 29/2007, IAB Nuremberg.

Table 1: Unemployment Benefits II and Short-term Training Programs

	2005	2006	2007	2008
Persons entitled to UBII (avg. annual stock) ^a	4,981,748	5,392,166	5,276,835	5,009,656
Spending for UBII (in billion Euro) ^b	32.8	34.7	31.5	30.2
Spending for ALMP (overall, in billion Euro) ^b	3.1	3.8	4.2	4.7
New Participants in Short-term Training Programs ^a	410,900	533,634	519,783	627,739

^a Figures obtained from Bundesagentur für Arbeit (2006, 2007, 2008, 2009).

^b Figures obtained from Bundesagentur für Arbeit (2009b).

Table 2: Results for Smith and Todd (2005b) balancing test

		Men				Women			
		$p > .1$	$p > .05$	$p > .01$	Regressors	$p > .1$	$p > .05$	$p > .01$	Regressors
Aptitude tests									
Quarter 1	Natives	28	31	35	37	42	44	44	46
	Immigrants	66	70	72	74	40	42	44	44
Quarter 2	Natives	50	54	56	58	38	38	39	39
	Immigrants	36	36	38	40	51	53	54	59
Quarter 3	Natives	36	37	38	38	38	40	42	45
	Immigrants	66	67	68	74	27	29	30	32
Quarter 4	Natives	30	31	33	35	/	/	/	/
	Immigrants	45	48	53	55	/	/	/	/
Job search training									
Quarter 1	Natives	58	61	63	65	41	43	44	48
	Immigrants	51	54	59	66	47	48	49	50
Quarter 2	Natives	41	42	45	46	31	32	32	32
	Immigrants	73	74	77	80	/	/	/	/
Quarter 3	Natives	23	24	24	26	/	/	/	/
	Immigrants	/	/	/	/	/	/	/	/
Quarter 4	Natives	/	/	/	/	/	/	/	/
	Immigrants	/	/	/	/	/	/	/	/
Skill provision									
Quarter 1	Natives	48	51	52	54	44	46	49	52
	Immigrants	55	55	60	62	51	53	53	55
Quarter 2	Natives	33	36	37	42	35	35	36	38
	Immigrants	34	35	37	39	63	65	67	69
Quarter 3	Natives	57	58	61	68	40	43	47	50
	Immigrants	32	32	32	33	40	40	41	44
Quarter 4	Natives	45	49	52	55	34	34	36	39
	Immigrants	53	56	58	59	25	26	27	29
Combined training programs									
Quarter 1	Natives	51	55	56	62	36	39	40	44
	Immigrants	76	77	80	83	52	56	56	59
Quarter 2	Natives	50	53	54	57	32	32	32	34
	Immigrants	43	43	44	44	40	42	43	48
Quarter 3	Natives	34	36	37	38	27	28	29	30
	Immigrants	27	28	28	31	33	33	34	35
Quarter 4	Natives	36	36	38	39	/	/	/	/
	Immigrants	/	/	/	/	/	/	/	/

Remarks: Fields marked by a / indicate that the number of treated individuals in the respective cell was too low to estimate valid treatment effects.

Table 3: Smith and Todd (2005b) balancing test for decomposition

	Men				Women			
	$p > .1$	$p > .05$	$p > .01$	Regressors	$p > .1$	$p > .05$	$p > .01$	Regressors
Aptitude tests	50	53	57	57	63	63	65	65
Job search training	63	64	68	69	68	70	70	70
Skill provision	43	44	46	47	50	50	51	52
Combined programs	44	46	48	50	78	80	81	81

Table 4: Effects of aptitude tests

Month after program start:	Men				Women			
	3	6	9	12	3	6	9	12
Quarter 1	Treated: 507 natives and 448 immigr.				Treated: 297 natives and 180 immigr.			
Natives	0.0668 3.76	0.0867 4.42	0.0834 4.04	0.0939 4.61	0.0490 2.42	0.0690 3.05	0.0673 2.74	0.0948 3.71
Immigrants	0.0464 2.66	0.0441 2.44	0.0703 3.55	0.0653 3.06	0.0425 1.82	0.0727 2.69	0.0608 2.19	0.0524 1.80
Quarter 2	Treated: 260 natives and 213 immigr.				Treated: 99 natives and 85 immigr.			
Natives	0.0965 3.88	0.0848 3.26	0.0760 2.84	0.0947 3.46	0.0675 1.69	0.1205 2.65	0.1170 2.61	0.1159 2.62
Immigrants	0.1007 3.77	0.1032 3.73	0.1129 3.69	0.1020 3.30	0.0227 0.69	0.0194 0.57	-0.0111 -0.35	0.0531 1.21
Quarter 3	Treated: 138 natives and 143 immigr.				Treated: 62 natives and 76 immigr.			
Natives	0.0815 2.43	0.0889 2.51	0.0640 1.86	- -	0.1126 2.25	0.0890 1.79	0.0992 1.79	- -
Immigrants	0.1033 3.17	0.1279 4.02	0.1566 4.32	- -	0.0601 1.63	0.0665 1.60	0.0767 1.79	- -
Quarter 4	Treated: 91 natives and 98 immigr.							
Natives	0.1532 3.24	0.1857 3.71	- -	- -	/ /	/ /	- -	- -
Immigrants	0.1085 2.95	0.0943 2.49	- -	- -	/ /	/ /	- -	- -

Remarks: Displayed are average treatment effects on the treated and corresponding t-values. Standard errors have been obtained through bootstrapping based on 250 replications. Fields marked by a - indicate that no outcome variable could be observed for the respective month. Fields marked by a / indicate that the number of treated individuals in the respective cell was too low to estimate valid treatment effects.

Table 5: Effects of job search training

Month after program start:	Men				Women			
	3	6	9	12	3	6	9	12
Quarter 1	Treated: 303 natives and 270 immigr.				Treated: 189 natives and 102 immigr.			
Natives	-0.0123	0.0307	0.0224	0.0329	0.0621	0.0585	0.0556	0.0910
	-0.65	1.24	0.93	1.24	2.31	1.98	1.82	2.87
Immigrants	-0.0014	0.0193	0.0637	0.0414	-0.0382	-0.0564	0.0131	0.0362
	-0.08	0.85	2.69	1.77	-2.10	-2.87	0.41	0.96
Quarter 2	Treated: 116 natives and 110 immigr.				Treated: 69 natives			
Natives	-0.0073	0.0191	-0.0045	0.0026	0.0703	0.1005	0.1205	0.0848
	-0.24	0.57	-0.13	0.07	1.54	2.06	2.31	1.62
Immigrants	0.0034	-0.0323	0.0126	0.0193	/	/	/	/
	0.12	-1.14	0.33	0.47	/	/	/	/
Quarter 3	Treated: 48 natives							
Natives	-0.0399	-0.0120	0.0191	-	/	/	/	-
	-1.03	-0.25	0.35	-	/	/	/	-
Immigrants	/	/	/	-	/	/	/	-
	/	/	/	-	/	/	/	-
Quarter 4								
Natives	/	/	-	-	/	/	-	-
	/	/	-	-	/	/	-	-
Immigrants	/	/	-	-	/	/	-	-
	/	/	-	-	/	/	-	-

Remarks: Displayed are average treatment effects on the treated and corresponding t-values. Standard errors have been obtained through bootstrapping based on 250 replications. Fields marked by a - indicate that no outcome variable could be observed for the respective month. Fields marked by a / indicate that the number of treated individuals in the respective cell was too low to estimate valid treatment effects.

Table 6: Effects of skill provision

Month after program start:	Men				Women			
	3	6	9	12	3	6	9	12
Quarter 1	Treated: 288 natives and 244 immigr.				Treated: 218 natives and 133 immigr.			
Natives	0.0430	0.0764	0.1192	0.1026	0.0318	0.0556	0.0639	0.0591
	1.83	2.72	4.23	3.60	1.47	2.15	2.30	2.02
Immigrants	0.0218	0.0497	0.0522	0.0617	0.0741	0.0956	0.1153	0.1263
	0.96	1.89	1.87	2.09	2.64	3.14	3.38	3.56
Quarter 2	Treated: 160 natives and 146 immigr.				Treated: 152 natives and 80 immigr.			
Natives	0.0157	0.0380	0.0308	-0.0029	0.0609	0.1253	0.1148	0.0934
	0.57	1.15	0.93	-0.08	1.92	3.26	3.05	2.61
Immigrants	-0.0096	0.0332	-0.0084	0.0032	-0.0052	0.0039	-0.0026	-0.0068
	-0.42	1.15	-0.29	0.10	-0.21	0.14	-0.08	-0.21
Quarter 3	Treated: 94 natives and 82 immigr.				Treated: 72 natives and 52 immigr.			
Natives	0.0124	-0.0115	-0.0016	-	0.0163	0.0201	0.0253	-
	0.36	-0.32	-0.04	-	0.43	0.49	0.55	-
Immigrants	0.0213	0.0527	0.0998	-	0.0489	0.0086	0.0503	-
	0.59	1.21	2.19	-	1.17	0.22	1.10	-
Quarter 4	Treated: 69 natives and 67 immigr.				Treated: 48 natives and 41 immigr.			
Natives	0.0722	0.0688	-	-	-0.0214	-0.0486	-	-
	1.59	1.44	-	-	-0.81	-1.83	-	-
Immigrants	0.1028	0.0563	-	-	-0.0231	-0.0400	-	-
	2.29	1.29	-	-	-1.03	-1.78	-	-

Remarks: Displayed are average treatment effects on the treated and corresponding t-values. Standard errors have been obtained through bootstrapping based on 250 replications. Fields marked by a - indicate that no outcome variable could be observed for the respective month.

Table 7: Effects of combined training programs

Month after program start:	Men				Women			
	3	6	9	12	3	6	9	12
Quarter 1	Treated: 410 natives and 360 immigr.				Treated: 188 natives and 135 immigr.			
Natives	0.0045	0.0306	0.0293	0.0289	0.0116	0.0157	0.0167	0.0066
	0.28	1.52	1.39	1.38	0.50	0.58	0.62	0.23
Immigrants	0.0069	0.0316	0.0208	0.0272	-0.0027	0.0142	0.0533	0.0656
	0.38	1.47	1.01	1.21	-0.13	0.51	1.61	1.87
Quarter 2	Treated: 180 natives and 157 immigr.				Treated: 113 natives and 95 immigr.			
Natives	0.0207	0.0089	0.0374	0.0379	0.0470	0.0459	0.0399	0.0452
	0.75	0.31	1.15	1.07	1.32	1.26	1.06	1.19
Immigrants	0.0082	0.0271	0.0054	0.0519	0.0586	0.0512	0.0742	0.0786
	0.31	0.94	0.18	1.57	1.73	1.48	1.92	2.02
Quarter 3	Treated: 90 natives and 96 immigr.				Treated: 66 natives and 56 immigr.			
Natives	-0.0250	-0.0262	-0.0355	-	0.0107	0.0176	0.0256	-
	-0.89	-0.76	-0.98	-	0.30	0.42	0.55	-
Immigrants	0.0154	0.0588	0.0744	-	0.0247	0.0228	0.0600	-
	0.50	1.48	1.70	-	0.67	0.55	0.15	-
Quarter 4	Treated: 63 natives							
Natives	-0.0407	-0.0108	-	-	/	/	-	-
	-1.31	-0.24	-	-	/	/	-	-
Immigrants	/	/	-	-	/	/	-	-
	/	/	-	-	/	/	-	-

Remarks: Displayed are average treatment effects on the treated and corresponding t-values. Standard errors have been obtained through bootstrapping based on 250 replications. Fields marked by a - indicate that no outcome variable could be observed for the respective month. Fields marked by a / indicate that the number of treated individuals in the respective cell was too low to estimate valid treatment effects.

Table 8: Differences in treatment effects between natives and immigrants

Month after program start:	Men				Women			
	3	6	9	12	3	6	9	12
Aptitude tests								
Δ_{Dif}^{ATT}	-0.0204	-0.0426	-0.0131	-0.0286	-0.0065	-0.0037	-0.0065	-0.0424
p-value	0.4130	0.1192	0.6492	0.3388	0.8389	0.9206	0.8644	0.2953
$\Delta_{residual}^{ATT}$	0.0135	0.0172	0.0639	0.0303	0.0320	0.0605	0.0421	0.0111
t-value	0.41	0.50	1.85	0.75	0.69	1.17	0.75	0.19
Job search training								
Δ_{Dif}^{ATT}	0.0109	-0.0114	0.0413	0.0085	-0.1003	-0.1149	-0.0425	-0.0548
p-value	0.6848	0.7294	0.2463	0.8175	0.0104	0.0071	0.3789	0.2996
$\Delta_{residual}^{ATT}$	0.0044	-0.0039	0.0239	-0.0319	-0.1488	-0.1434	-0.0002	0.0072
t-value	0.09	-0.06	0.36	-0.46	-1.74	-1.52	0.00	0.06
Skill provision								
Δ_{Dif}^{ATT}	-0.0212	-0.0267	-0.0670	-0.0409	0.0423	0.0400	0.0514	0.0672
p-value	0.5028	0.4680	0.0873	0.3101	0.2563	0.3538	0.2633	0.1585
$\Delta_{residual}^{ATT}$	0.0107	0.0231	-0.0778	-0.0591	0.0962	0.1149	0.1298	0.1401
t-value	0.23	0.48	-1.46	-0.99	1.62	1.69	1.65	1.69
Combined training programs								
Δ_{Dif}^{ATT}	0.0024	0.0010	-0.0085	-0.0017	-0.0143	-0.0015	0.0366	0.0590
p-value	0.9205	0.9714	0.7753	0.9580	0.6713	0.9715	0.4068	0.2012
$\Delta_{residual}^{ATT}$	0.0014	-0.0278	-0.0191	-0.0386	0.0387	0.0182	0.0814	0.0426
t-value	0.04	-0.59	-0.39	-0.75	0.50	0.21	0.85	0.39

Remarks: Δ_{Dif}^{ATT} denotes the mean difference in the ATTs between immigrants and native Germans for the respective program and month after program start. The p-values derive from t-tests on the equality of the ATTs in the group of natives and immigrants. $\Delta_{residual}^{ATT}$ is based on the matching approach described in section 4.3 and denotes the estimated difference in the ATTs for immigrants and native Germans that is solely due to unobservable differences between the two ethnic groups, or in other words, which is due to the immigrant fixed effect. t-values denote significance of these immigrant fixed effects. Standard errors have been obtained through bootstrapping based on 250 replications.

A Data Appendix

In this appendix we provide selected descriptive statistics. The means of the variables depicted in Tables A.1 to A.8 refer to participants in the considered training programs before matching. Thus, the number of observations might differ from the number of observations displayed in Tables 4 to 7. The tables are stratified according to the ethnic group and the quarter of program start. The p-values derive from t-tests on the equality of means of the displayed variables between natives and immigrants for the respective quarter.

Table A.1: Means of selected variables for male participants in aptitude tests

	Quarter 1		Quarter 2		Quarter 3		Quarter 4					
	Natives	Immigr.	p-value	Natives	Immigr.	p-value	Natives	Immigr.	p-value			
Age												
18-24	0.302	0.190	0.000	0.281	0.155	0.001	0.250	0.140	0.019	0.261	0.163	0.100
25-34	0.337	0.413	0.016	0.365	0.469	0.022	0.400	0.448	0.420	0.272	0.480	0.003
35-49	0.312	0.346	0.260	0.308	0.329	0.627	0.321	0.357	0.533	0.435	0.316	0.093
50-57	0.049	0.051	0.886	0.046	0.047	0.968	0.029	0.056	0.255	0.033	0.041	0.766
Schooling												
No school leaving certificate	0.087	0.205	0.000	0.065	0.239	0.000	0.064	0.175	0.004	0.120	0.194	0.162
Secondary general school	0.473	0.444	0.367	0.542	0.460	0.075	0.464	0.476	0.850	0.446	0.418	0.706
Intermediate secondary school	0.250	0.161	0.001	0.242	0.150	0.013	0.286	0.175	0.027	0.293	0.153	0.020
Vocational diploma (Fachabitur)	0.057	0.049	0.579	0.062	0.038	0.238	0.079	0.021	0.025	0.033	0.031	0.938
University entrance diploma	0.085	0.094	0.629	0.054	0.061	0.738	0.071	0.077	0.861	0.098	0.102	0.923
Missing	0.047	0.047	0.973	0.035	0.052	0.361	0.036	0.077	0.134	0.011	0.102	0.007
Family status												
Single	0.673	0.422	0.000	0.685	0.324	0.000	0.593	0.308	0.000	0.663	0.286	0.000
Household size												
Number of persons	1.509	2.116	0.000	1.465	2.169	0.000	1.607	2.238	0.000	1.565	2.296	0.000
Region												
East Germany	0.055	0.022	0.009	0.058	0.033	0.203	0.036	0.084	0.089	0.054	0.092	0.325
Labor market history 2 years before program start (measured in number of half-months)												
Employment	14.225	15.855	0.080	11.915	12.300	0.735	9.536	10.203	0.606	8.283	9.755	0.302
Unemployment	15.570	15.839	0.737	19.781	20.000	0.836	24.707	23.776	0.496	28.543	27.847	0.626
Job seeking while employed	1.134	0.569	0.002	1.185	0.610	0.033	0.764	0.517	0.362	0.478	0.490	0.967
Program participation	2.406	3.009	0.166	2.608	3.056	0.478	2.686	1.895	0.306	3.696	1.888	0.035
Out of labor force	14.665	12.728	0.084	12.512	12.033	0.737	10.307	11.608	0.431	7.000	8.020	0.523
Labor market history 4 years before program start (measured in number of half-months)												
Employment	34.970	35.221	0.892	32.392	32.944	0.822	28.871	29.378	0.863	30.022	30.204	0.959
Unemployment	23.793	26.377	0.045	29.423	28.803	0.714	34.550	33.622	0.687	39.500	39.663	0.949
Job seeking while employed	1.448	0.944	0.024	1.854	0.962	0.037	1.157	0.832	0.353	0.935	0.939	0.992
Program participation	4.469	5.509	0.134	5.300	5.667	0.726	5.229	4.350	0.501	6.293	3.531	0.057
Out of labor force	31.320	27.949	0.114	27.031	27.624	0.833	26.193	27.818	0.653	19.250	21.663	0.537
Observations	507	448		260	213		140	143		92	98	

Table A.2: Means of selected variables for female participants in aptitude tests

	Quarter 1		Quarter 2		Quarter 3		Quarter 4					
	Natives	Immigr.	p-value	Natives	Immigr.	p-value	Natives	Immigr.	p-value			
Age												
18-24	0.293	0.239	0.200	0.253	0.151	0.090	0.258	0.263	0.946	0.232	0.042	0.006
25-34	0.279	0.361	0.062	0.364	0.453	0.217	0.226	0.289	0.401	0.304	0.542	0.014
35-49	0.347	0.361	0.752	0.343	0.302	0.554	0.452	0.355	0.253	0.429	0.396	0.738
50-57	0.081	0.039	0.072	0.040	0.093	0.149	0.065	0.092	0.555	0.036	0.021	0.655
Schooling												
No school leaving certificate	0.071	0.139	0.014	0.061	0.209	0.003	0.065	0.132	0.197	0.071	0.208	0.042
Secondary general school	0.354	0.356	0.964	0.364	0.302	0.381	0.403	0.342	0.463	0.339	0.312	0.774
Intermediate secondary school	0.313	0.200	0.007	0.384	0.128	0.000	0.371	0.237	0.088	0.446	0.188	0.005
Vocational diploma (Fachabitur)	0.057	0.056	0.939	0.081	0.105	0.578	0.065	0.053	0.768	0.071	0.083	0.822
University entrance diploma	0.104	0.094	0.727	0.051	0.116	0.103	0.048	0.079	0.473	0.054	0.083	0.550
Missing	0.101	0.156	0.078	0.061	0.140	0.071	0.048	0.158	0.040	0.018	0.125	0.030
Family status												
Single	0.552	0.372	0.000	0.495	0.256	0.001	0.419	0.289	0.113	0.464	0.125	0.000
Household size												
Number of persons	1.502	1.933	0.000	1.455	2.140	0.000	1.645	2.171	0.006	1.464	2.417	0.000
Region												
East Germany	0.067	0.022	0.029	0.071	0.012	0.049	0.113	0.053	0.196	0.107	0.104	0.961
Labor market history 2 years before program start (measured in number of half-months)												
Employment	16.751	15.578	0.460	12.889	9.988	0.161	9.500	9.329	0.934	6.554	8.271	0.368
Unemployment	11.727	12.750	0.405	19.434	17.977	0.473	20.968	22.724	0.363	28.500	23.771	0.024
Job seeking while employed	0.862	0.822	0.889	1.061	0.942	0.836	1.855	0.763	0.139	1.018	0.333	0.147
Program participation	2.340	1.828	0.409	1.222	2.081	0.213	1.726	1.408	0.693	1.732	2.750	0.399
Out of labor force	16.320	17.022	0.697	13.394	17.012	0.144	13.952	13.776	0.945	10.196	12.875	0.288
Labor market history 4 years before program start (measured in number of half-months)												
Employment	36.158	32.106	0.172	32.626	22.314	0.014	25.710	27.895	0.649	23.018	20.938	0.655
Unemployment	17.569	17.694	0.944	25.152	26.500	0.671	27.661	28.355	0.822	38.536	30.542	0.043
Job seeking while employed	1.148	1.067	0.819	1.465	1.523	0.932	3.065	1.171	0.053	1.839	0.521	0.034
Program participation	4.158	3.039	0.216	3.141	3.163	0.984	3.516	2.382	0.450	2.518	7.854	0.032
Out of labor force	36.966	42.094	0.133	33.616	42.500	0.079	36.048	36.197	0.979	30.089	36.146	0.313
Observations	297	180		99	86		62	76		56	48	

Table A.3: Means of selected variables for male participants in job search training

	Quarter 1		Quarter 2		Quarter 3		Quarter 4					
	Natives	Immigr.	p-value	Natives	Immigr.	p-value	Natives	Immigr.	p-value			
Age												
18-24	0.320	0.252	0.072	0.259	0.264	0.932	0.208	0.122	0.259	0.226	0.081	0.096
25-34	0.343	0.363	0.622	0.328	0.373	0.479	0.354	0.429	0.458	0.323	0.486	0.176
35-49	0.264	0.300	0.340	0.353	0.327	0.680	0.333	0.388	0.581	0.323	0.351	0.806
50-57	0.073	0.085	0.577	0.060	0.036	0.405	0.104	0.061	0.447	0.129	0.081	0.524
Schooling												
No school leaving certificate	0.053	0.233	0.000	0.095	0.255	0.001	0.104	0.102	0.973	0.097	0.324	0.024
Secondary general school	0.545	0.452	0.027	0.405	0.436	0.637	0.542	0.490	0.614	0.484	0.459	0.844
Intermediate secondary school	0.231	0.167	0.055	0.259	0.145	0.035	0.250	0.224	0.771	0.194	0.054	0.077
Vocational diploma (Fachabitur)	0.036	0.022	0.322	0.052	0.045	0.828	0.000	0.041	0.161	0.097	0.027	0.230
University entrance diploma	0.076	0.093	0.473	0.086	0.064	0.522	0.104	0.102	0.973	0.097	0.054	0.509
Missing	0.059	0.033	0.142	0.103	0.055	0.176	0.000	0.041	0.161	0.032	0.081	0.402
Family status												
Single	0.686	0.400	0.000	0.672	0.391	0.000	0.646	0.408	0.019	0.710	0.270	0.000
Household size												
Number of persons	1.465	2.126	0.000	1.362	2.100	0.000	1.438	2.184	0.008	1.452	2.000	0.047
Region												
East Germany	0.069	0.037	0.089	0.026	0.027	0.948	0.062	0.000	0.077	0.000	0.000	-
Labor market history 2 years before program start (measured in number of half-months)												
Employment	19.208	21.585	0.117	13.112	14.055	0.634	8.875	13.510	0.063	7.226	6.946	0.900
Unemployment	9.759	11.393	0.094	15.310	15.373	0.966	25.500	23.796	0.431	24.774	28.919	0.106
Job seeking while employed	0.957	0.870	0.748	1.000	0.627	0.254	0.812	0.082	0.011	1.742	0.135	0.075
Program participation	2.617	2.085	0.383	2.319	2.482	0.867	2.438	3.347	0.505	0.774	4.270	0.014
Out of labor force	15.459	12.067	0.020	16.259	15.464	0.718	10.375	7.265	0.212	13.484	7.730	0.063
Labor market history 4 years before program start (measured in number of half-months)												
Employment	40.092	44.663	0.100	33.284	32.009	0.749	30.812	43.102	0.019	23.065	26.432	0.573
Unemployment	17.297	18.481	0.449	22.121	23.309	0.587	34.917	31.857	0.398	31.258	41.757	0.013
Job seeking while employed	1.327	1.141	0.564	1.345	1.391	0.932	2.167	0.408	0.014	2.710	0.216	0.090
Program participation	4.142	3.656	0.572	4.267	4.255	0.994	5.562	4.653	0.709	3.645	7.919	0.137
Out of labor force	33.142	28.059	0.068	34.983	35.036	0.990	22.542	15.980	0.172	35.323	19.676	0.033
Observations	303	270		116	110		48	49		31	37	

Table A.4: Means of selected variables for female participants in job search training

	Quarter 1		Quarter 2		Quarter 3		Quarter 4					
	Natives	Immigr.	p-value	Natives	Immigr.	p-value	Natives	Immigr.	p-value			
Age												
18-24	0.337	0.331	0.909	0.333	0.304	0.725	0.250	0.308	0.632	0.227	0.071	0.233
25-34	0.326	0.331	0.940	0.319	0.357	0.655	0.219	0.308	0.451	0.318	0.429	0.515
35-49	0.253	0.254	0.975	0.275	0.268	0.926	0.344	0.385	0.753	0.364	0.429	0.707
50-57	0.084	0.085	0.987	0.072	0.071	0.982	0.188	0.000	0.019	0.091	0.071	0.842
Schooling												
No school leaving certificate	0.068	0.127	0.082	0.072	0.179	0.070	0.094	0.154	0.493	0.091	0.214	0.31
Secondary general school	0.353	0.441	0.124	0.435	0.411	0.789	0.344	0.154	0.104	0.318	0.571	0.141
Intermediate secondary school	0.311	0.161	0.003	0.261	0.196	0.400	0.312	0.346	0.790	0.364	0.071	0.05
Vocational diploma (Fachabitur)	0.100	0.051	0.125	0.058	0.036	0.566	0.000	0.038	0.271	0.000	0.071	0.215
University entrance diploma	0.084	0.136	0.152	0.130	0.107	0.693	0.094	0.192	0.287	0.136	0.000	0.158
Missing	0.084	0.085	0.987	0.043	0.071	0.503	0.156	0.115	0.660	0.091	0.071	0.842
Family status												
Single	0.563	0.322	0.000	0.522	0.286	0.007	0.469	0.269	0.124	0.500	0.143	0.03
Household size												
Number of persons	1.568	1.907	0.002	1.565	1.929	0.030	1.438	2.269	0.001	1.636	1.929	0.412
Region												
East Germany	0.068	0.042	0.345	0.072	0.125	0.325	0.000	0.000	0.045	0.071	0.749	
Labor market history 2 years before program start (measured in number of half-months)												
Employment	19.921	19.059	0.691	13.913	14.500	0.828	15.500	7.538	0.027	9.545	11.214	0.675
Unemployment	9.105	9.780	0.634	16.087	17.571	0.531	21.344	18.346	0.276	25.727	29.214	0.341
Job seeking while employed	0.900	0.975	0.842	0.812	0.732	0.867	0.562	0.885	0.575	2.227	0.571	0.375
Program participation	2.168	1.280	0.286	0.841	2.232	0.179	1.406	0.577	0.265	1.455	1.286	0.911
Out of labor force	15.905	16.907	0.658	16.348	12.964	0.251	9.188	20.654	0.001	9.045	5.714	0.402
Labor market history 4 years before program start (measured in number of half-months)												
Employment	40.221	36.398	0.327	31.029	30.857	0.973	37.500	16.115	0.003	29.318	31.000	0.86
Unemployment	14.763	15.161	0.856	21.145	26.464	0.165	29.719	21.385	0.072	32.182	43.429	0.16
Job seeking while employed	1.132	1.381	0.593	1.203	0.875	0.539	1.250	1.308	0.945	4.318	1.071	0.447
Program participation	3.368	2.983	0.742	1.333	3.929	0.112	2.094	0.846	0.174	2.727	1.286	0.622
Out of labor force	36.516	40.076	0.396	41.290	33.875	0.211	25.438	56.346	0.000	27.455	19.214	0.398
Observations	190	118		69	56		32	26		22	14	

Table A-5: Means of selected variables for male participants in skill provision

	Quarter 1		Quarter 2		Quarter 3		Quarter 4					
	Natives	Immigr.	p-value	Natives	Immigr.	p-value	Natives	Immigr.	p-value			
Age												
18-24	0.299	0.152	0.000	0.213	0.171	0.362	0.245	0.110	0.021	0.087	0.075	0.794
25-34	0.330	0.413	0.044	0.281	0.329	0.368	0.415	0.524	0.148	0.275	0.463	0.023
35-49	0.312	0.390	0.056	0.412	0.404	0.882	0.298	0.354	0.433	0.565	0.328	0.005
50-57	0.059	0.045	0.476	0.094	0.096	0.949	0.043	0.012	0.229	0.072	0.134	0.238
Schooling												
No school leaving certificate	0.059	0.201	0.000	0.119	0.205	0.039	0.085	0.183	0.055	0.087	0.164	0.176
Secondary general school	0.500	0.447	0.213	0.519	0.397	0.033	0.404	0.585	0.016	0.478	0.463	0.857
Intermediate secondary school	0.233	0.182	0.142	0.188	0.171	0.712	0.298	0.098	0.001	0.261	0.134	0.065
Vocational diploma (Fachabitur)	0.073	0.034	0.045	0.038	0.027	0.621	0.053	0.061	0.825	0.043	0.045	0.971
University entrance diploma	0.094	0.091	0.909	0.112	0.103	0.784	0.128	0.024	0.011	0.087	0.060	0.546
Missing	0.042	0.045	0.828	0.025	0.096	0.008	0.032	0.049	0.570	0.043	0.134	0.063
Family status												
Single	0.635	0.360	0.000	0.625	0.377	0.000	0.766	0.341	0.000	0.681	0.299	0.000
Household size												
Number of persons	1.514	2.201	0.000	1.438	2.062	0.000	1.319	2.329	0.000	1.536	2.239	0.002
Region												
East Germany	0.035	0.023	0.402	0.044	0.062	0.484	0.106	0.049	0.161	0.029	0.015	0.580
Labor market history 2 years before program start (measured in number of half-months)												
Employment	16.264	18.189	0.151	12.137	13.144	0.517	9.064	12.512	0.054	6.812	8.701	0.260
Unemployment	15.052	14.750	0.769	20.031	18.575	0.277	23.319	23.927	0.693	27.812	28.731	0.618
Job seeking while employed	1.385	0.629	0.006	0.637	0.712	0.785	1.011	0.463	0.154	0.449	0.299	0.573
Program participation	3.354	2.523	0.158	2.356	2.466	0.872	2.223	1.939	0.711	4.203	2.597	0.203
Out of labor force	11.944	11.909	0.980	12.837	13.103	0.883	12.383	9.159	0.116	8.725	7.672	0.605
Labor market history 4 years before program start (measured in number of half-months)												
Employment	38.326	39.360	0.684	31.169	29.603	0.629	27.149	35.610	0.041	26.913	29.358	0.570
Unemployment	22.635	22.799	0.918	28.956	28.699	0.910	32.330	32.366	0.989	39.130	39.806	0.833
Job seeking while employed	1.830	1.197	0.091	1.050	1.288	0.608	1.691	1.000	0.319	1.043	0.433	0.294
Program participation	5.955	4.784	0.239	4.506	4.610	0.929	4.649	4.659	0.995	6.478	5.493	0.658
Out of labor force	27.253	27.860	0.828	30.319	31.801	0.690	30.181	22.366	0.080	22.435	20.910	0.752
Observations	288	264		160	146		94	82		69	67	

Table A.6: Means of selected variables for female participants in skill provision

	Quarter 1		Quarter 2		Quarter 3		Quarter 4					
	Natives	Immigr.	p-value	Natives	Immigr.	p-value	Natives	Immigr.	p-value			
Age												
18-24	0.242	0.233	0.849	0.204	0.200	0.944	0.097	0.173	0.217	0.208	0.293	0.364
25-34	0.283	0.391	0.036	0.355	0.338	0.788	0.292	0.250	0.611	0.375	0.244	0.188
35-49	0.365	0.323	0.425	0.388	0.400	0.861	0.528	0.462	0.471	0.354	0.415	0.564
50-57	0.110	0.053	0.068	0.053	0.062	0.757	0.083	0.115	0.555	0.062	0.049	0.782
Schooling												
No school leaving certificate	0.023	0.150	0.000	0.026	0.188	0.000	0.083	0.038	0.319	0.083	0.171	0.216
Secondary general school	0.374	0.323	0.333	0.375	0.387	0.853	0.347	0.269	0.360	0.333	0.268	0.511
Intermediate secondary school	0.384	0.263	0.021	0.342	0.125	0.000	0.347	0.288	0.494	0.375	0.195	0.064
Vocational diploma (Fachabitur)	0.068	0.030	0.123	0.046	0.075	0.364	0.028	0.077	0.211	0.062	0.024	0.393
University entrance diploma	0.114	0.158	0.239	0.112	0.138	0.570	0.111	0.192	0.208	0.021	0.122	0.059
Missing	0.037	0.075	0.111	0.099	0.087	0.783	0.083	0.135	0.362	0.125	0.220	0.240
Family status												
Single	0.584	0.286	0.000	0.553	0.287	0.000	0.375	0.269	0.220	0.458	0.220	0.018
Household size												
Number of persons	1.580	1.797	0.033	1.586	2.138	0.000	1.764	2.096	0.102	1.833	2.195	0.115
Region												
East Germany	0.037	0.030	0.747	0.046	0.062	0.593	0.097	0.058	0.429	0.021	0.024	0.911
Labor market history 2 years before program start (measured in number of half-months)												
Employment	16.849	16.925	0.966	13.401	13.250	0.937	9.889	11.558	0.476	6.229	6.732	0.806
Unemployment	14.644	13.331	0.363	17.211	18.487	0.448	22.042	19.519	0.208	24.312	27.780	0.127
Job seeking while employed	0.858	0.902	0.888	1.007	0.863	0.746	0.847	0.500	0.409	0.271	0.683	0.400
Program participation	2.411	1.774	0.378	2.342	1.700	0.480	2.236	1.519	0.545	2.062	2.146	0.938
Out of labor force	13.237	15.068	0.337	14.039	13.700	0.881	12.986	14.904	0.490	15.125	10.659	0.118
Labor market history 4 years before program start (measured in number of half-months)												
Employment	36.986	34.805	0.519	33.368	31.013	0.559	24.708	28.500	0.434	18.896	23.732	0.353
Unemployment	22.064	21.098	0.661	22.776	24.275	0.529	31.167	26.654	0.179	32.562	31.634	0.791
Job seeking while employed	1.324	1.361	0.930	1.480	1.138	0.507	1.236	0.885	0.560	0.792	1.561	0.302
Program participation	4.205	3.158	0.324	4.447	2.800	0.248	5.069	2.212	0.183	3.271	2.195	0.401
Out of labor force	31.420	35.579	0.257	33.928	36.775	0.537	33.819	37.750	0.520	40.479	36.878	0.581
Observations	219	133		152	80		72	52		48	41	

Table A.7: Means of selected variables for male participants in combined training programs

	Quarter 1		Quarter 2		Quarter 3		Quarter 4					
	Natives	Immigr.	p-value	Natives	Immigr.	p-value	Natives	Immigr.	p-value			
Age												
18-24	0.351	0.233	0.000	0.328	0.261	0.183	0.256	0.240	0.802	0.254	0.136	0.102
25-34	0.276	0.400	0.000	0.378	0.382	0.934	0.389	0.406	0.810	0.349	0.356	0.939
35-49	0.276	0.300	0.456	0.228	0.306	0.106	0.311	0.312	0.984	0.365	0.407	0.640
50-57	0.098	0.067	0.122	0.067	0.051	0.544	0.044	0.042	0.926	0.032	0.102	0.121
Schooling												
No school leaving certificate	0.117	0.275	0.000	0.100	0.204	0.007	0.100	0.292	0.001	0.032	0.271	0.000
Secondary general school	0.522	0.431	0.011	0.444	0.510	0.234	0.500	0.406	0.201	0.556	0.492	0.483
Intermediate secondary school	0.183	0.119	0.015	0.294	0.121	0.000	0.189	0.167	0.694	0.222	0.102	0.073
Vocational diploma (Fachabitur)	0.044	0.033	0.450	0.039	0.038	0.975	0.044	0.031	0.639	0.032	0.034	0.947
University entrance diploma	0.071	0.056	0.390	0.078	0.076	0.963	0.089	0.031	0.097	0.111	0.068	0.408
Missing	0.063	0.086	0.231	0.044	0.051	0.780	0.078	0.073	0.901	0.048	0.034	0.705
Family status												
Single	0.698	0.394	0.000	0.722	0.420	0.000	0.667	0.375	0.000	0.698	0.322	0.000
Household size												
Number of persons	1.412	2.050	0.000	1.333	2.032	0.000	1.422	2.260	0.000	1.270	2.288	0.000
Region												
East Germany	0.020	0.008	0.193	0.028	0.025	0.896	0.022	0.010	0.526	0.032	0.034	0.947
Labor market history 2 years before program start (measured in number of half-months)												
Employment	12.576	15.992	0.001	10.511	11.261	0.582	9.089	9.896	0.629	5.317	8.746	0.035
Unemployment	14.385	14.192	0.836	18.978	19.854	0.488	23.978	26.010	0.218	28.413	29.322	0.631
Job seeking while employed	0.890	0.558	0.088	1.000	0.516	0.102	1.022	0.677	0.365	0.825	0.254	0.127
Program participation	3.080	2.625	0.407	2.533	2.331	0.777	1.744	1.458	0.669	2.476	2.102	0.703
Out of labor force	17.068	14.633	0.066	14.978	14.038	0.588	12.167	9.958	0.273	10.968	7.576	0.122
Labor market history 4 years before program start (measured in number of half-months)												
Employment	30.420	36.836	0.002	27.878	30.006	0.471	29.422	29.927	0.895	20.635	30.254	0.022
Unemployment	22.593	22.619	0.985	27.539	30.076	0.198	31.022	36.146	0.052	41.698	41.881	0.961
Job seeking while employed	1.266	1.100	0.552	1.644	1.038	0.178	1.289	1.135	0.727	1.762	0.559	0.074
Program participation	4.837	4.217	0.411	4.939	4.573	0.743	4.167	3.510	0.678	6.397	3.508	0.097
Out of labor force	36.885	31.228	0.026	34.000	30.306	0.290	30.100	25.281	0.268	25.508	19.797	0.261
Observations	410	360		180	157		90	96		63	59	

Table A.8: Means of selected variables for female participants in combined training programs

	Quarter 1		Quarter 2		Quarter 3		Quarter 4					
	Natives	Immigr.	p-value	Natives	Immigr.	p-value	Natives	Immigr.	p-value			
Age												
18-24	0.345	0.281	0.169	0.327	0.326	0.986	0.197	0.268	0.358	0.188	0.091	0.234
25-34	0.246	0.421	0.000	0.327	0.253	0.240	0.288	0.500	0.016	0.354	0.545	0.090
35-49	0.345	0.236	0.017	0.265	0.337	0.265	0.394	0.179	0.009	0.375	0.212	0.122
50-57	0.065	0.062	0.907	0.080	0.084	0.905	0.121	0.054	0.197	0.083	0.152	0.344
Schooling												
No school leaving certificate	0.082	0.191	0.001	0.142	0.242	0.065	0.076	0.125	0.367	0.042	0.273	0.003
Secondary general school	0.332	0.303	0.540	0.398	0.389	0.898	0.470	0.536	0.472	0.375	0.485	0.331
Intermediate secondary school	0.254	0.169	0.037	0.292	0.147	0.013	0.227	0.107	0.081	0.375	0.121	0.011
Vocational diploma (Fachabitur)	0.086	0.039	0.058	0.035	0.063	0.354	0.061	0.054	0.869	0.062	0.061	0.973
University entrance diploma	0.073	0.084	0.682	0.035	0.032	0.880	0.045	0.054	0.838	0.062	0.000	0.147
Missing	0.172	0.213	0.295	0.097	0.126	0.509	0.121	0.125	0.950	0.083	0.061	0.705
Family status												
Single	0.565	0.427	0.006	0.584	0.326	0.000	0.500	0.339	0.075	0.542	0.212	0.003
Household size												
Number of persons	1.612	1.899	0.007	1.425	2.084	0.000	1.455	2.196	0.000	1.729	2.061	0.208
Region												
East Germany	0.017	0.006	0.289	0.044	0.011	0.149	0.015	0.000	0.359	0.042	0.061	0.703
Labor market history 2 years before program start (measured in number of half-months)												
Employment	12.608	12.792	0.902	10.761	12.779	0.275	8.424	6.857	0.452	8.083	9.848	0.460
Unemployment	11.440	11.292	0.905	18.920	17.242	0.346	21.879	23.625	0.433	26.833	29.152	0.309
Job seeking while employed	0.716	0.820	0.747	0.743	1.032	0.435	0.773	0.786	0.978	0.958	0.485	0.412
Program participation	2.478	1.494	0.128	2.690	1.074	0.039	1.015	1.750	0.370	2.667	2.000	0.629
Out of labor force	20.759	21.601	0.670	14.885	15.874	0.673	15.909	14.982	0.742	9.458	6.515	0.249
Labor market history 4 years before program start (measured in number of half-months)												
Employment	26.310	27.646	0.647	26.142	29.726	0.341	25.773	21.286	0.358	22.896	29.364	0.282
Unemployment	16.599	15.876	0.690	25.947	23.221	0.324	30.167	32.929	0.472	35.958	40.697	0.288
Job seeking while employed	1.043	0.961	0.836	1.257	1.758	0.309	0.985	1.125	0.799	1.062	0.515	0.350
Program participation	3.901	1.966	0.032	4.354	1.632	0.016	1.894	2.679	0.542	6.458	4.121	0.450
Out of labor force	48.147	49.551	0.706	38.301	39.663	0.770	37.182	37.982	0.897	29.625	21.303	0.202
Observations	232	178		113	95		66	56		48	33	