Discussion Paper No. 13-056

What Active Labour Market Programmes Work for Immigrants in Europe?

A Meta-Analysis of the Evaluation Literature

Sebastian Butschek and Thomas Walter

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 13-056

What Active Labour Market Programmes Work for Immigrants in Europe?

A Meta-Analysis of the Evaluation Literature

Sebastian Butschek and Thomas Walter

Non-technical summary

Immigrants are a disadvantaged group on European labour markets. In most countries, they are over-represented in unemployment and under-represented in employment. To integrate immigrants into the labour market, European governments use a broad range of Active Labour Market Programmes (ALMPs) including language and introduction courses, job search assistance, training programmes, and subsidised public and private sector employment.

A growing number of empirical evaluation studies investigate the employment effects of ALMPs on immigrants. Yet, there is no clear picture which programmes work for immigrants and which do not. In order to help policymakers allocate resources efficiently we condense the findings of the existing studies by means of a meta-analysis. A meta-analysis identifies the extent to which certain features of the underlying studies (e.g., programmes analysed, methods employed, data used) predict the results of these studies. We focus on the relationship between the type of programme evaluated and the programme's estimated employment effect on immigrants.

We find that only wage subsidies in the private sector can be confidently recommended to European policy-makers. In most evaluation studies, wage subsidies are estimated to increase employment chances of immigrants. On the other hand, most studies present insignificant estimates for the effect of training. The same is true for job-search assistance and public-sector employment.

Das Wichtigste in Kürze

In vielen europäischen Ländern sind Immigranten eine benachteiligte Gruppe auf dem Arbeitsmarkt. Sie sind häufig in Arbeitslosigkeit überrepräsentiert und in Beschäftigung unterrepräsentiert. Um die Beschäftigungschancen von Immigranten zu erhöhen, setzen europäische Regierungen auf vielfältige Maßnahmen der aktiven Arbeitsmarktpolitik. Diese beinhalten Sprach- und Integrationskurse, Unterstützungsmaßnahmen bei der Jobsuche, Trainingsmaßnahmen sowie subventionierte Beschäftigung im öffentlichen und privaten Sektor.

Eine wachsende Anzahl an empirischen Evaluationsstudien untersucht die Beschäftigungseffekte dieser arbeitsmarktpolitischen Maßnahmen für Immigranten. Bisher ergibt sich daraus allerdings noch kein klares Bild, welche Maßnahmen für die Arbeitsmarktintegration von Immigranten förderlich sind und welche nicht. Um den politischen Entscheidungsträgern zu helfen, die Maßnahmen effizient einzusetzen, verdichten wir die Ergebnisse der vorhandenen Studien durch eine Meta-Analyse. Eine Meta-Analyse ermöglicht es zu identifizieren, in welchem Maße bestimmte Elemente der zugrunde liegenden Studien (z. B. die untersuchten Maßnahmen, die verwendeten Methoden und Daten) Einfluss auf die Ergebnisse haben, zu denen diese Studien gelangen. In unserer Analyse sind wir am Einfluss des untersuchten Maßnahmentyps auf die Beschäftigungschancen der teilnehmenden Immigranten interessiert.

Unsere Ergebnisse zeigen, dass nur der Einsatz von Lohnsubventionen im privaten Sektor als beschäftigungswirksam für Immigranten eingestuft werden kann. Die meisten Evaluationsstudien ermitteln für Lohnsubventionen positive Beschäftigungseffekte. Evaluationen von Trainingsmaßnahmen finden hingegen überwiegend insignifikante Effekte. Dasselbe gilt für Unterstützungsmaßnahmen bei der Jobsuche und subventionierte Beschäftigung im öffentlichen Sektor.

What Active Labour Market Programmes Work for Immigrants in Europe?

A Meta-Analysis of the Evaluation Literature

Sebastian Butschek
(ZEW Mannheim)

Thomas Walter
(ZEW Mannheim)

This version: August 21, 2013

Abstract

A growing body of programme evaluation literature recognises immigrants as a disadvantaged group on European labour markets and investigates the employment effects of Active Labour Market Programmes (ALMPs) on this subgroup. Using a meta-analysis, we condense 93 estimates from 33 empirical studies of the effectiveness of four types of ALMPs employed across Europe to combat immigrant unemployment: training, job search assistance, and subsidised public and private sector employment. We find that only wage subsidies in the private sector can be confidently recommended to European policy-makers.

Sebastian Butschek, Centre for European Economic Research (ZEW), L 7, 1, D-68161 Mannheim, Germany. Tel: +49 621 1235238; E-mail: butschek@zew.de

Thomas Walter, Centre for European Economic Research (ZEW), L 7, 1, D-68161 Mannheim, Germany. Tel: +49 621 1235363; E-mail: walter@zew.de

1 Introduction

Immigrants are under-represented in employment and over-represented in unemployment in most European countries. In 2009/2010 their employment rate was on average 2.9 percentage points lower than that of natives across all European OECD countries. At the same time their unemployment rate was 4.3 percentage points higher (see OECD, 2012).

To facilitate immigrants' labour market integration, European governments use a wide range of Active Labour Market Programmes (ALMPs). These include language and introduction courses, job search assistance, training programmes, and subsidised public and private sector employment and involve substantial government spending. While evidence on the effectiveness of single programmes exists, it has not been reviewed systematically to establish empirically which types of programmes actually facilitate immigrants' employment uptake. To shed light on this question we review the small but growing literature evaluating the employment effects of ALMPs on immigrants in Europe. By means of a meta-analysis, we try to identify which ALMPs work for immigrants and which ones do not. Our results should help policymakers employ activation measures more efficiently.

With respect to ALMPs' effects on all unemployed workers (natives and immigrants), recent analyses have strengthened a growing consensus: job search assistance and, to some extent, wage subsidies are effective in the short run while training works in the longer run; subsidised public sector employment (also known as public works), however, is generally ineffective (Heckman et al., 1999, Greenberg et al., 2003, Kluve, 2010, Card et al., 2010). Also, the findings of the ALMP evaluation literature on heterogeneous treatment effects on women or young workers have been reviewed (e.g., Bergemann and van den Berg, 2008, Card et al., 2010).

For immigrants, only two surveys of the literature on the effect of ALMPs exist. Nekby (2008) provides a qualitative review of four studies evaluating labour market programmes for immigrants in the Nordic countries; she concludes that the same types of ALMPs work for immigrants as for the general population of unemployed workers. Rinne (2012) discusses the findings of three studies evaluating language/introduction courses designed for immigrants and eight recent evaluations of general labour market programmes' effects on immigrants. In a similar vein as Nekby, he suggests that "programs that are relatively closely linked to the labor market (e.g., work experience and wage subsidies) appear as the comparatively most

effective programs" (Rinne, 2012, p 19). While both surveys present relevant evidence, neither of them formally aggregates the findings of the studies reviewed.

We provide an accessible quantitative summary of the existing empirical evidence. To this end, we collect the relevant studies following a search protocol and then condense the findings extracted from them in two steps: first, using descriptive analysis and second, performing a meta-analysis with sign and significance of the effect estimate as our outcome variable of interest.

As Stanley (2001) argues, "The most important strength of meta-analysis is that it moves literature reviews away from casual judgments about "good" studies that deserve attention and "poor" studies that should be set aside, and instead provides a replicable statistical framework for summarizing and interpreting the full range of evidence." The key ingredient of such an analysis is then an exhaustive data set of relevant studies. We find 33 micro-econometric papers that estimate 93 short-run treatment effects up to two years after programme start. The interventions evaluated were implemented in the Nordic countries (Norway, Sweden, Denmark, and Finland), Germany, the Netherlands and Switzerland between 1984 and 2007.

Our descriptive analysis looks at the distribution of impact estimates conditional on study characteristics in order to provide an absolute indication of the effectiveness of different types of ALMPs. Performing a meta-analysis of the same sample of effect estimates allows us to go beyond the descriptive analysis in two ways. First, we can control for different study characteristics when investigating which ALMP types are associated with significant or insignificant impact findings, simultaneously addressing such issues as methodological differences or changes in programme effectiveness over time. Second, we can provide a summary measure for whether the evidence suggests that one type of ALMP works better than some other.

We find that subsidised employment in the private sector is significantly more likely estimated to have a positive effect on immigrants' labour market outcomes than training. For the other ALMP types, our meta-analysis yields mostly insignificant results. The descriptive analysis shows that evaluations of training and job search assistance programmes produce predominantly insignificant effect estimates for immigrant participants. Public works seem to perform even worse, receiving many insignificant and negative evaluation results. Therefore, only subsidised employment in the private sector seems to be effective for immigrants.

The remainder of this paper consists of four parts: section 2 provides background information on immigrants on European labour markets; section 3 describes the data and presents some

descriptive analysis; section 4 discusses the findings of our meta-analysis and performs a sensitivity check; section 5 concludes.

2   Immigrants in Europe and their labour market integration

On average, the share of immigrants (defined as foreign-born persons) among the total population amounted to 11.2% in European OECD countries in 2009/2010 (see Table A.1 in the appendix). 1 In almost every country, the share of immigrants in the working age population (age 15 to 64) is even larger than in the total population. It amounts to 13.3% on average across European OECD countries.

Despite substantial heterogeneity in immigrant origins, European countries share the problem of integrating immigrants into the labour market. Immigrants are usually under-represented in employment and over-represented in unemployment. Table 1 shows the employment and unemployment rates of immigrants in our sample of seven European OECD countries in 2009/2010 and how they compare to the respective rates of the native populations. On average, the employment rate for immigrants in these countries is 65.8%. It is 9.3 percentage points lower than the rate for natives. Differences in the employment rate are especially pronounced in Denmark, the Netherlands and Sweden, with a difference of more than 10 percentage points. Correspondingly, unemployment is more prevalent among immigrants than among natives. Immigrants' unemployment rate is 11.6% on average. It is more than twice as high as the natives' unemployment rate (5.4%). Across all European OECD countries, differences in the employment and unemployment rates between immigrants and natives are smaller than in the seven countries investigated here but still sizeable: 4.3 percentage points for unemployment and 2.9 percentage points for employment.

Include Table 1 about here

To combat the high level of unemployment among immigrants and to foster their employment uptake governments use Active Labour Market Programmes (ALMPs). For immigrants, two different categories of ALMPs can be distinguished: first, programmes that are specifically

---

[1] Note that this figure includes only the first generation of immigrants. Unfortunately, comparable data on immigrants across European countries including the second or third generation are not available.

designed for and exclusively targeted at immigrants, and second, general programmes that are also used for the native population. In what follows we will refer to these categories as migrant-specific and general ALMPs, respectively.

General ALMPs comprise four types of interventions (see, e.g., Card et al., 2010):

1) *Training*: This includes all programmes that aim to enhance participants' skills needed for employment uptake (e.g. computer courses or courses providing specific occupational knowledge). Training programmes can be provided either on-the-job within a firm or off-the-job in a classroom.

2) *Subsidised private sector employment*: This category comprises programmes that generate incentives to increase job opportunities in the private sector. One example for such a programme is wage subsidies for employers who hire disadvantaged workers. Wage subsidies can also be paid to workers when they accept a job with a wage below their unemployment benefits or when they start their own business.

3) *Subsidised public sector employment (public works)*: This type of intervention aims at offering temporary job opportunities outside the private sector, mainly for community services. Public-works programmes should be designed in such a way that they do not crowd out regular employment.

4) *Job search assistance and sanctions*: This intervention type has the objective of making the job search process of participants more effective and efficient. Job search assistance is predominantly provided by public employment services and includes counselling and monitoring of job search efforts. In case of a lack of job search effort, sanctions are intended to restore an appropriate level of compliance.

Migrant-specific programmes can be grouped into three categories:

1) *Language training* often not only improves participants' ability to communicate in the host country's main language but also provides information about history, culture and institutions of the host country. One example for such a course is the so-called orientation course in Germany (see, e.g., Liebig, 2007)

2) *Introduction programmes* provide a customised integration plan towards employment uptake. Targeted at newly arriving immigrants, they usually start with language training and continue with either training or subsidised em-

ployment. Throughout the programme job search assistance is provided. See, e.g., Andersson Joona and Nekby (2012) and Sarvimäki and Hämäläinen (2012) for introduction programmes in Sweden and Finland, respectively.

3) *General programmes exclusively for immigrants* comprise general ALMPs other than language courses (training, subsidised private or public sector employment, job search assistance and sanctions) targeted at immigrants (and not at natives). One example for such an intervention is intensified job search assistance programmes, where immigrants are assigned to caseworkers whose caseload is reduced. That is, caseworkers have more time for the counselling and support of each individual. See, e.g., Aslund and Johansson (2011) for a programme of this kind in Sweden.

Whether general programmes or migrant-specific ones are more effective for the integration of immigrants in the labour market is a question of major policy interest. One might expect migrant-specific programmes to be more successful since they are designed for the needs of immigrants whereas general programmes address the needs of average participants, including mostly natives. However, the fact that in practice both programmes coexist in all European countries might be taken to suggest that neither of them is superior or that policymakers are not aware of which programmes work and which ones do not. [2] Empirical studies have not established an answer to this question either. We attempt to address it by means of our meta-analysis (see section 3).

3   Description of the Data

3.1 Estimation Sample

To obtain an exhaustive sample of studies evaluating the effects of ALMPs on immigrants' labour market outcomes, we implemented the following search protocol:

1) Collect studies on ALMPs surveyed by Nekby (2008), Rinne (2012), Kluve (2010), and Card et al. (2010).

---

[2] If the migrant-specific programmes were successful in integrating all newly-arriving immigrants into the labour market and into stable jobs, there would not be any need for participation in general ALMPs later on.

2) Perform internet keyword searches[3] on 27 November 2012 and on 15 April 2013 to find additional studies.

We then identified those studies that met the following selection criteria:

1) Studies that estimate ALMP treatment effects for immigrants.[4]

2) Studies that perform a micro-econometric evaluation of the intervention's effect on individual labour market outcomes, outlining the identification strategy.

3) Studies that evaluate an intervention that roughly fits into one of four ALMP categories (described more fully below): training, wage subsidy, public works, or services/sanctions. We also admitted studies that evaluate the aggregate effect of a country's ALMPs.

Applying these criteria yielded a sample of 34 studies estimating ALMP effects on immigrants' probability of or hazard to employment[5] in seven countries (Denmark, Finland, Germany, Netherlands, Norway, Sweden, Switzerland). Some studies evaluate several programmes or perform their analyses separately by gender or region as well as estimating effects for different points in time during follow-up. For comparability, we focus on short-run estimates, defined as effect estimates based on outcomes observed up to two years after programme start[6]. Where there is more than one such short-term estimate per gender-region-group combination, we choose the latest (most long-term) one. This gives 33 studies providing 93 short-run estimates.[7]

The four ALMP categories we use are as follows:

---

[3] Keywords, in different combinations: ALMP, labour market programmes, labor market programs, migrants, foreign, native, born, citizen, subgroup, sub-group, hetero; search engines used: Google Scholar, EconPapers and Econis.

[4] The definition of immigrants varies across studies. It usually means those with foreign citizenship, the foreign-born or individuals whose parents or grandparents were foreign-born. Most studies estimate heterogeneous ALMP treatment effects for several subgroups, including immigrants. A few studies have a sample of only immigrants.

[5] Instead of looking at employment, one study considers earnings. Another, evaluating the promotion of self-employment, uses yet a different outcome variable: neither unemployed nor in receipt of unemployment.

[6] While we do have information on longer-run outcomes (38 estimates), there is not enough variation in them to permit a separate econometric analysis. One study reports only long-term estimates for 36 and 50 months after the programme (Groß et al., 2006). This study is dropped from the analysis.

[7] See the references for the list of the 33 studies analysed.

1) Classroom and on-the-job training (henceforth "training")

2) Subsidised private sector employment ("wage subsidy")

3) Subsidised public sector employment ("public works")

4) Job-search assistance and sanctions ("services/sanctions")

These are taken from Card et al. (2010) but are fairly standard in the evaluation literature, as exemplified by analogous definitions in Calmfors (1994) and Kluve (2010). We allow for a fifth residual group of "other programmes" (including aggregate ALMP effects and programmes that combine several ALMP types in a single treatment). See also section 2 for a definition of the ALMP categories. [8]

From our sample of 33 studies we extracted information about the programme evaluated and its geographic and chronological setting, the sample studied and the methods applied. We recorded programme type, duration and whether it was designed specifically for immigrants in order to characterise the nature of the treatment. To capture sample characteristics, we included information on whether an effect was estimated for males, females or a mixed group of participants as well as in what country and decade they received the treatment. As methodological proxies we documented the econometric technique used and whether the estimates came from a published paper or a working paper.[9]

3.2 Summary Statistics

The first column of Table 2 summarises the distribution of the short-run estimates we focus on. First, consider the outcome variable: those evaluations finding no effect are most frequent (48 estimates), followed by ones finding significantly positive effects (32), with significantly negative effect estimates less frequent still (13 estimates).

Include Table 2 about here

---

[8] Since only 6 out of the 33 studies analyse migrant-specific programmes, our econometric analysis cannot differentiate between these types of ALMP as outlined in section 2. Instead we define a dummy variable to indicate whether a programme is migrant-specific or general. We then classify migrant-specific programmes as training, wage subsidy, public works, services/sanctions or other programme, depending on their content.

[9] We categorised PhD dissertations as published studies because of the similarities between PhD supervision and the referee process.

Next, lines 2a) to 2e) show that among ALMP types, training programmes dominate (30 estimates). Public works also feature prominently, contributing 23 data points; the third column reveals that this pattern is driven by evaluations for Germany. Both wage-subsidy (16) and services/sanctions (17) each provide about half as many observations as the largest category. There are seven estimates in the residual category (other programmes). Only six of our 93 estimates are for migrant-specific programmes, whereas 87 are for general ones (see lines 2f) and 2g), respectively).

Lines 3a) through 3c) show that 25 estimates are for programmes with a duration of up to four months, while 20 effects are estimated for programmes of at least five months. However, most short-run estimates are for programmes of unknown or mixed duration (48), reflecting some heterogeneity in the level of detail on interventions given in the papers. Lines 4a) through 4c) reveal that about two thirds of estimates are from the 2000s. Next, lines 5a) through 5c) illustrate that matching approaches were the most popular method (55 estimates). From column 3 it is clear that German estimates, based on matching procedures with only one exception, account for this distribution. Finally, lines 6a) and 6b) demonstrate that less than a third of the short-run estimates came from published papers (28), with Nordic evaluations accounting for disproportionately many publications (20)[10].

Comparing estimates by origin reveals that the largest contributor, Germany, differs markedly from Denmark, Finland, Norway and Sweden (Nordic countries) and Switzerland and the Netherlands (other countries). Nordic estimates are relatively optimistic about programme effects while most German estimates are insignificant and effects tend to be more often negative in the other countries. There is more variety in the methods used to evaluate Nordic and Swiss/Dutch programmes than for German programmes. While training is the ALMP type evaluated most frequently in the Nordic countries and Switzerland/Netherlands, evaluations of public works dominate in Germany.

---

[10] We do not intend to suggest that published papers meet different quality standards than working papers, given that we are agnostic about the relative quality of the various refereed journals and opt for estimates from working papers in some cases where the published version no longer presents all heterogeneous effect estimates, e.g., Gerfin and Lechner (2000).

3.3 Descriptive Analysis

In this subsection we present the distribution of the outcome variable (short-run effect: significantly negative, insignificant or significantly positive) conditional on the covariates we extracted from the studies (see Table 3). This serves a dual purpose: one is to provide a flavour of the potential results of the meta-analysis; another is to give some absolute indications of the effectiveness of the programme types evaluated. This is important because our meta-analysis, by virtue of its method, only allows conclusions about the relative effectiveness of different types of programme.

Include Table 3 about here

Lines 1a) to 1e) of Table 3 show that insignificant estimates are the largest category in all types of ALMP except for wage subsidies, where 10 out of 16 estimates are positive. For training and services/sanctions, about half of the estimates are insignificant while for public works, about two thirds are insignificant. For both training and services/sanctions, positive estimates are more frequent than negative ones, while the converse is true for public works. These raw descriptive statistics indicate that wage subsidies seem to have positive employment effects; for the remaining ALMP categories, the evidence mostly points to an employment effect too close to zero to be significant. Because only six of the 93 estimates are for migrant-specific programmes (see lines 1f) and 1g), respectively), it is hard to draw reliable conclusions about their relative effectiveness. We effectively focus on the effect of general ALMPs on immigrants.

The next three lines, 2a) to 2c), seem to suggest that duration analysis is more optimistic about programme effectiveness than matching approaches are. Similarly, lines 3a) to 3c) may point to a deterioration of ALMP quality over time from the 1980s to the 2000s. While programme duration is unknown in most cases, short programmes may have been more effective than longer ones. Finally, published papers seem to find positive effects in a higher fraction of cases. Yet, regional and chronological differences may be confounding all of these potential relationships; multivariate analysis will help disentangle these effects.

4 Empirical Analysis

4.1 Method

We perform an ordered probit analysis with sign and significance of the estimate as the outcome variable. This variable can take three values: significantly negative, insignificantly different from zero, and significantly positive. The explanatory variables of interest are dummies describing the type of ALMP. In addition, we include a number of variables to account for differences in evaluation technique and setting. We focus on the relationship between ALMP type and sign/significance of the short-run effect estimated for each study-gender-region cell.

The index model underlying our estimation is as follows:

$$y_i^* = \alpha_1 WS_i + \alpha_2 PW_i + \alpha_3 SE_i + \alpha_4 OT_i + X_i'\beta + u_i,$$

where WS, PW, SE and OT are dummy variables describing the programme type analysed in study i (wage subsidy, public works, services/sanctions, or other programmes, with training being the omitted category), X is a vector of control variables (study characteristics, programme characteristics, sample characteristics, contextual controls) and u is an error term.

Taking into account the degrees of freedom underlying each effect estimate as well as a measure of effect size (such as a t-statistic) would be an attractive alternative to the sign/significance approach we take (see, e.g., Greenberg et al., 2003 and Stanley, 2005). It would allow us to identify systematic differences in estimated effect size across types of programme, unconfounded by differences in precision. However, the plurality of the underlying econometric estimation techniques makes this unfeasible.[11]

Card et al. (2010) show that the approach we follow would be invalid if the pattern of estimate sign and significance were generated by differences in precision rather than differences in effect size. They also demonstrate that the sign/significance approach is approximately valid when the effective sample size is constant, i.e., when larger samples are offset by more demanding designs. They present evidence that this is the case in their sample of studies and indeed find that the sign/significance approach and an effect size-based analysis on a subsample of studies that use the probability of employment as outcome variable yield similar results.

---

[11]It is not straightforward how test statistics from different types of models can be transformed into a common distribution so that the test statistics can be compared directly. For instance, test statistics from a duration model and a matching model will have different distributions (and degrees of freedom).

While we cannot perform such a check for our smaller set of studies, we can partly rely on their finding in that there is some overlap between our samples of evaluation studies.

Recent theoretical work on meta-analyses has stressed the importance of checking for publication bias (see, e.g., Stanley, 2005). We cannot rule out publication bias as the test statistics of the studies we analyse are not directly comparable. In this paper, however, our focus is the relative effectiveness of different types of ALMPs. As long as the presence of publication bias does not interact with the type of programme evaluated, it will not distort our findings on the relative effectiveness of different types of ALMPs.

4.2   Estimation results

We estimate six specifications of the ordered probit model outlined above, gradually introducing groups of control variables. Specification 1 includes only the type of programme, omitting training. Specification 2 adds study characteristics: whether the study employed duration analysis or some other econometric technique (omitted: matching), and whether the paper is published (baseline: working paper). Specification 3 introduces programme characteristics, namely whether the intervention was designed for immigrants and whether the treatment was short, that is, no longer than four months. In specification 4, sample characteristics enter the equation: participant gender (baseline: pooled estimation for men and women) and treatment in the 2000s (omitted: 1980s or 1990s). Specification 5 adds the unemployment rate in the year that the evaluated programme started as a proxy for the macroeconomic context. Alternatively, specification 6 uses GDP growth as a contextual control. [12]

Table 4 presents the results. In specification 1, which includes only programme type, we obtain positive coefficients for wage subsidy and services/sanctions. The interpretation for these positive coefficients is that studies evaluating wage subsidies and services/sanctions are more likely to find positive employment effects than studies evaluating training. However, the estimated coefficients are not statistically significant. The coefficients on public works and other programmes are negative, though only the coefficient for other programmes is marginally significant. There is no meaningful interpretation for the coefficient on "other programmes" as this is a residual category.

---

[12] Unemployment rates and GDP growth rates were obtained from the Online OECD Employment database; see http://www.oecd.org/employment/employmentpoliciesanddata/onlineoecdemploymentdatabase.htm (accessed on 7 January, 2013).

When including study characteristics (method and publication status) in specification 2, the results remain very similar. None of the ALMP types are significant except for the residual category. Specification 3's programme features cause the wage subsidy coefficient to grow and become significant; all other ALMP type coefficients remain insignificant. Including information about the sample on the right-hand side in specification 4 does little to wage subsidy (still significant) and public works or services/sanctions (still insignificant). When we add contextual control variables in specification 5 (unemployment rate) and 6 (GDP growth rate), the coefficient on wage subsidy remains positive and significant, though only marginally in specification 6. The parameter estimates on public works and services/sanctions are still insignificant.

Based on the Akaike information criterion we choose specification 5 as our preferred one. Specification 5 contains the national unemployment rate at programme start as a contextual control variable. The coefficient on the unemployment rate is positive and significant, suggesting that inferior macroeconomic conditions at the time of treatment are associated with a higher probability of a positive evaluation result.[13]

Almost all other control variables in specification 5 (and across the other specifications) have insignificant coefficient estimates. Exceptions are the dummy for short programme (which is always marginally significant and positive) and the dummy for published paper (which is positive and at least marginally significant in most specifications). The dummy for migrant-specific programmes is insignificant. While this implies that these programmes are equally (in)effective as general ones, this result has to be interpreted with caution as the number of studies analysing migrant-specific programmes is small.

Include Table 4 about here

It is worth re-iterating that our ordered probit analysis only permits relative, not absolute, conclusions on the effectiveness of ALMP types. Thus, our meta-analysis suggests that wage subsidies work better than training. Because the corresponding coefficients are insignificant, we cannot claim with confidence that public works are less effective than training or that ser-

---

[13] This result is in line with the findings of Lechner and Wunsch (2009), who show a positive correlation of the unemployment rate at the start of the programme with the effectiveness of training programmes in Germany.

vices/sanctions are more effective even though coefficients signs consistently point in that direction.

Combining our conclusions from the meta-analysis with the descriptive analysis in section 3 suggests that wage subsidies not only work better than other programmes but also do have positive employment effects on immigrants. The descriptive results appear most pessimistic about public works, as do the meta-analytic results (albeit insignificant), suggesting at the least that this type of programme should be used very selectively. Insignificant programme estimates dominate the descriptive analysis of training and job search assistance. No firm conclusions can be drawn on the suitability of these activation measures for immigrants in the short-run. Additional research on the longer-run employment effects may help clarify the picture.

Our findings are based on a smaller sample of studies and on a more specific group of programme participants than the meta-analyses of Kluve (2010) and Card et al. (2010) but point in a similar direction. Moreover, they are in line with the conclusions that Nekby (2008) and Rinne (2012) arrived at in their qualitative reviews.

4.3 Robustness analysis

To address the potential criticism that our selection criteria were to some extent arbitrary we vary these criteria and re-estimate our preferred specification on the sample of estimates this gives. This provides a simple check on the robustness of our results. Our variation is to tighten the definitions of ALMPs and of the short run.

First, we drop estimates for programmes that, strictly speaking, are not ALMPs. That is, we exclude evaluation studies of temporary agency work, which like wage subsidy programmes make hiring cheaper and may facilitate employer-worker matching, and exclude transfer reduction programmes, which work much like sanctions. Moreover, we exclude evaluations of aggregate ALMP effects. Second, we define the short run as up to twelve months after programme start rather than 24 months.

This gives an alternative sample of 86 estimates from 27 studies. As Table 5 illustrates, our variation results in a very similar pattern of coefficient estimates, providing evidence for the robustness of our findings (see Table A.2 in the appendix for the full results). In the appendix we also include the results of both main and robustness analysis when standard errors are

clustered at the study level[14]. As Tables A.3 and A.4 illustrate, these are practically the same as those presented here.

Include Table 5 about here

5   Conclusion

Immigrants constitute an important group on European labour markets in terms of both the risks they face and the potential they harbour: they are numerous and over-represented in unemployment on the one hand but younger than the native population on the other. This highlights the importance of immigrants' labour market integration. While the full range of ALMPs is used in practice, there is little empirical guidance for policy-makers seeking to facilitate immigrants' employment take-up. In other words, there is not yet a clear indication of what programmes work for immigrants.

To answer this question, we provide a quantitative synthesis of the evidence on ALMPs' effect on immigrants. Using 93 effect estimates extracted from 33 relevant evaluation studies, we perform a meta-analysis of the evaluation results. An ordered probit analysis based on sign and significance of short-run effect estimates suggests that wage subsidies work better for immigrants than training programmes. Public works may be less effective while job search assistance programmes (services/sanctions) may be more effective but estimated coefficients are insignificant. To help interpret these relative statements, we present a detailed descriptive analysis: there, effect estimates for wage subsidy programmes are mostly positive, suggesting that wage subsidies are indeed a promising measure to increase employment rates of immigrants. They should be used more often than public works which reduce employment chances or result in insignificant employment effects at best. The short-run effects of services/sanctions on employment prospects are mostly insignificant. The same is true for training.

At this point, only wage subsidies can be confidently recommended to policymakers trying using general ALMPs to improve immigrants' labour market integration. The reasons for the

---

[14] Several studies present estimates for the effects of multiple programmes. These estimates are based on different samples of people though. This is why we regard the estimates as independent. It may be argued that author's individual research strategies introduce correlation between multiple estimates of one study – in which case clustering at the study level would be appropriate.

disappointing effects of the other programmes need to be clarified in further research. Furthermore, migrant-specific interventions such as language courses and introduction programmes, on which the evidence is still scarce, may be promising; further research in this area is warranted, too.

References

Bergemann, Annette and Gerard J. van den Berg (2008): "Active Labor Market Policy Effects for Women in Europe – A Survey." Annals of Economics and Statistics, 91/92, 385-408.

Calmfors, Lars (1994): "Active labour market policy and unemployment – a framework for the analysis of crucial design features." OECD Economic Studies, 22, 7-47.

Card, David, Jochen Kluve and Andrea Weber (2010): "Active Labor Market Policy Evaluations: A Meta-Analysis." The Economic Journal, 120, 452-477.

Greenberg, David H., Charles Michalopoulos and Philip K. Robins (2003): "A Meta-Analysis of Government-Sponsored Training Programs." International and Labour Relations Review, 57, 31-53.

Groß, Verena, Michael Rothgang and Michael Schumacher (2006): "A comprehensive evaluation of ESF financed labour market policy in Germany." Unpublished Working Paper.

Heckman, James J., Robert J. Lalonde and Jeffrey A. Smith (1999): "The Economics and Econometrics of Active Labor Market Programs." In Ashenfelter, Orley and David Card (eds.): Handbook of Labor Economics, Volume 3A. Amsterdam and New York: Elsevier, 1865-2095.

Kluve, Jochen (2010): "The effectiveness of European active labor market programs." Labour Economics, 17, 904-918.

Liebig, Thomas (2007): "The Labour Market Integration of Immigrants in Germany." OECD Social, Employment and Migration Working Papers, No. 47, OECD Publishing.

Nekby, Lena (2008): "Active labor market programs for the integration of youths and immigrants into the labor market. The Nordic experience." CEPAL Serie Macroeconomía del desarollo No. 73.

OECD (2012): "Settling In: OECD Indicators of Immigrant Integration 2012." OECD Publishing.

Rinne, Ulf (2012): "The Evaluation of Immigration Policies." IZA Discussion Paper No. 6369.

Stanley, Tom (2005): "Beyond Publication Bias." Journal of Economic Surveys 19(3), 309-345.

Stanley, Tom (2001): "Wheat From Chaff: Meta-Analysis As Quantitative Literature Review." Journal of Economic Perspectives 15(3), 131-150.

Studies used in meta-analysis:

Ahmad, Nisar and Michael Svarer (2010): "The Effect of Sanctions and Active Labour Market Programmes on the Exit Rate from Unemployment." Working paper.

Aldashev, Alisher, Stephan L. Thomsen and Thomas Walter (2010): "Short-term training programs for immigrants in welfare: Do effects differ from natives and why?" ZEW Discussion Paper 10-027.

Andersson Joona, Pernilla and Lena Nekby (2012): "Intensive coaching of new immigrants: an evaluation based on random program assignment." Scandinavian Journal of Ecnomics 114(2), 575-600.

Andrén, Thomas and Daniela Andrén (2006): "Assessing the employment effects of vocational training using a one-factor model." Applied Economics, 38, 2469-2486.

Andrén, Thomas and Björn Gustafsson(2004): "Income effects from labor market training programs in Sweden during the 1980s and 1990s." International Journal of Manpower, 25, 8, 688-713.

Aslund, Olof and Per Johansson (2011): "Virtues of SIN: Can intensified public efforts help disadvantaged immigrants?" Evaluation Review 35(4), 399-427.

Bernhard, Sarah, Hermann Gartner and Gesine Stephan (2008): "Wage subsidies for needy job-seekers and their effect on individual labour market outcomes after the German reforms." IAB Discussion Paper 21/2008.

Bernhard, Sarah and Thomas Kruppe (2012): "Effectiveness of further vocational training in Germany." IAB Discussion Paper 10/2012.

Bernhard, Sarah and Joachim Wolff (2008): "Contracting out placement services in Germany." IAB Discussion Paper 5/2008.

Caliendo, Marco and Steffen Künn (2010): Start-up subsidies for the unemployed: long-term evidence and effect heterogeneity." IZA Discussion Paper No. 4790.

Clausen, Jens, Eskil Heinesen, Jans Hummelgaard, Leif Husted and Michael 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

Delander, Lennart, Mats Hammarstedt, Jonas Mansson and Erik Nyberg (2005): "Integration of immigrants: The role of language proficiency and experience." Evaluation Review 29/1, 24-41.

Gerfin, Michael and Michael Lechner (2000): "Microeconometric evaluation of the Active Labour Market Policy in Switzerland." IZA Discussion Paper No. 154.

Hartig, Martina, Eva Jozwiak und Joachim Wolff (2008): " Trainingsmaßnahmen: Für welche unter 25-jährigen Arbeitslosengeld II-Empfänger erhöhen sie die Beschäftigungschancen?" IAB-Forschungsbericht 6/2008.

Heinesen, Eskil, Leif Husted and Michael Rosholm (2011): "The effects of Active Labour Market Policies for immigrants receiving social assistance in Denmark." IZA Discussion Paper No. 5632.

Hohmeyer, Katrin and Joachim Wolff (2007): A fistful of Euros. Does one-euro-job participation lead means-tested benefit recipients into regular jobs and out of unemployment benefit II receipt?" IAB Discussion Paper 32/2007.

Huber, Martin, Michael Lechner, Conny Wunsch and Thomas Walter (2009): "Do German Welfare-to-work programmes reduce welfare dependency and increase employment?" German Economic Review 12(2), 1-23.

Jahn, Elke and Michael Rosholm (2012): "Is Temporary Agency Employment a Stepping Stone for Immigrants?" IZA discussion paper no. 6405

Kjaersgaard, Lene and Eskil Heinesen (2012): "Effects of consecutive Active Labor Market Programs – evidence from immigrants in Denmark." Ph.D. Dissertation.

Lalive, Rafael, Jan C. van Ours and Josef Zweimüller (2002): "The impact of Active Labor Market Programs on the Duration of Unemployment." Institute for Empirical Research in Economics Working Paper No. 41.

Larsson, Laura, Lindqvist, Linus and Oskar Nordström Skans (2005): "Stepping-stones or dead-ends? An analysis of Swedish replacement contracts." IFAU working paper 2005:18.

Prey, Hedwig (2000): "Evaluation of Training Programs in St. Gallen, Switzerland." Schweiz. Zeitschrift für Volkswirtschaft und Statistik.

Richardson, Katarina and Gerard J. van den Berg (2008): „Duration dependence versus unobserved heterogeneity in treatment effects: Swedish labor market training and the transition rate to employment." IFAU Working Paper 2008:7.

Ronsen, Marit and Torbjorn Skardhamar (2009): "Do welfare-to-work initiatives work? Evidence from an activation programme targeted at social assistance recipients in Norway." Journal of European Social Policy 19, 61-77

Rosholm, Michael and Rune Vejlin (2010): "Reducing income transfers to refugee immigrants: Does start-help help you start?" Labour Economics 17, 258-275

Sarvimäki, Matti and Kari Hämäläinen (2012): "Assimilating immigrants. The impact of an integration program." Updated version of HECER Discussion Paper No. 306.

Sianesi, Barbara (2001): "An evaluation of the active labour market programmes in Sweden." IFAU working paper 2001:5

Svarer, Michael (2010): "The Effect of Sanctions on Exit from Unemployment: Evidence from Denmark." Economica 78, 751-778

Thomsen, Stephan L. and Thomas Walter (2010): "Temporary extra jobs for immigrants: Merging lane to employment or dead-end road in welfare?" Labour 24(s1), 114-140.

Van den Berg, Gerard J., Bas van der Klaauw and Jan C. van Ours (2004): „Punitive sanctions and the transition rate from welfare to work." Journal of Labor Economics 22(1).

Walter, Thomas (2012): "Germany's 2005 Welfare Reform. Evaluating Key Characteristics With a Focus on Immigrants." ZEW Economic Studies, Vol. 46, Physica-Verlag, Heidelberg.

Wolff, Joachim and Anton Nivorozhkin (2008): "Start me up. The effectiveness of a self-employment programme for needy unemployed people in Germany." IAB Discussion Paper 20/2008.

Wolff, Joachim and Eva Jozwiak (2008): "Do short-term training programmes active means-tested unemployment benefit recipients in Germany?" LASER Discussion Paper No. 12.

Tables

Table 1: Employment and unemployment rates of immigrants in selected European OECD countries, 2009/2010

Country	Employment rate of foreign-born (in %)	Difference to natives (in percentage points)	Unemployment rate of foreign-born (in %)	Difference to natives (in percentage points)
Denmark	65.6	-10.0	11.8	5.5
Finland	62.1	-6.6	16.3	8.2
Germany	63.8	-8.7	12.2	5.6
Netherlands	65.5	-11.9	7.7	4.2
Norway	66.6	-9.8	9.9	7.0
Sweden	61.7	-12.9	15.8	8.7
Switzerland	75.1	-5.1	7.4	4.2
7-country average (unweighted)	65.8	-9.3	11.6	6.2
European OECD average (unweighted)	63.2	-2.9	12.6	4.3

Source: OECD (2012) and own calculations

Table 2: Characteristics of the estimation sample

	Short-run estimates	Nordic countries	Estimates for Germany	Other countries
1) Estimated programme effect				
a) Negative	13	4	1	8
b) Insignificant	48	13	30	5
c) Positive	32	17	12	3
2) ALMP type				
a) Training	30	9	13	8
b) Wage subsidy	16	9	3	4
c) Public works	23	5	15	3
d) Services/Sanctions	17	6	10	1
e) Other programmes	7	5	2	0
f) Migrant-specific programme	6	5	0	1
g) General programme	87	29	43	15
3) Programme duration				
a) Up to 4 months	25	3	18	4
b) 5 or more months	20	1	15	4
c) Mixed/unknown	48	30	10	8
4) Time evaluated programme ran				
a) 1980s	2	2	0	0
b) 1990s	31	15	0	16
c) 2000s	60	17	43	0
5) Method employed				
a) Matching	55	4	42	9
b) Duration	29	22	0	7
c) Other method	9	8	1	0
6) Publication status				
a) Working paper	65	14	37	14
b) Published	28	20	6	2
Number of estimates	93	34	43	16

Remarks: The table displays absolute numbers. Short-run estimates are defined as effect estimates based on outcomes observed up to two years after programme start. Where there are more than one such short-term estimates, the latest (most long-term) one is sampled. Nordic countries include Denmark, Finland, Norway and Sweden. Other countries include the Netherlands and Switzerland.

Table 3: Distribution of the estimated programme effects in the estimation sample

	Estimated effect is		
	significantly negative	insignificant	significantly positive
<u>1) ALMP type</u>			
a) Training	5	14	11
b) Wage subsidy	2	4	10
c) Public works	5	15	3
d) Services/Sanctions	0	9	8
e) Other programmes	1	6	0
f) Migrant-specific programme	0	4	2
g) General programme	13	44	30
<u>2) Method employed</u>			
a) Matching	4	38	13
b) Duration	8	6	15
c) Other method	1	4	4
<u>3) Time evaluated programme ran</u>			
a) 1980s	0	0	2
b) 1990s	9	11	11
c) 2000s	4	37	19
<u>4) Programme duration</u>			
a) Up to 4 months	2	10	13
b) 5 or more months	5	14	1
c) Mixed/unknown	6	24	18
<u>5) Publication status</u>			
a) Working paper	11	35	19
b) Published	2	13	13
Number of estimates	13	48	32

Remarks: The table displays absolute numbers. The numbers relate to short-run estimates which are defined as effect estimates based on outcomes observed up to two years after programme start. Where there are more than one such short-term estimates, the latest (most long-term) one is sampled.

Table 4: Estimation results

	(1)	(2)	(3)	(4)	(5)	(6)
ALMP type (baseline: training)						
Wage subsidy	0.5732	0.6987	1.0664**	1.0561**	1.1301**	0.9855*
	(0.4398)	(0.4890)	(0.5051)	(0.5279)	(0.5116)	(0.5400)
Public works	-0.5024	-0.4879	-0.1016	-0.1287	-0.2983	-0.1835
	(0.3170)	(0.3158)	(0.3279)	(0.4567)	(0.4579)	(0.4574)
Services/sanctions	0.5002	0.4474	0.342	0.2952	0.2386	0.2475
	(0.3303)	(0.3394)	(0.3480)	(0.3803)	(0.4049)	(0.3886)
Other programmes	-0.5985*	-0.6727*	-0.5925*	-0.6039*	-0.9004**	-0.7541*
	(0.3271)	(0.3678)	(0.3490)	(0.3627)	(0.3530)	(0.4108)
Study characteristics (baseline: matching, working paper)						
Duration analysis		-0.1694	-0.1524	-0.1252	0.5544	-0.1132
		(0.3715)	(0.3919)	(0.4024)	(0.4542)	(0.4054)
Other method		-0.3006	0.0912	0.1477	0.2542	0.1718
		(0.4948)	(0.5966)	(0.5858)	(0.5948)	(0.5695)
Published paper		0.6123**	0.7224**	0.7099*	0.9968***	0.6462*
		(0.2846)	(0.2997)	(0.3704)	(0.3727)	(0.3681)
Programme characteristics (baseline: regular ALMP, duration unknown or greater than four months)						
Migrant-specific pro-gramme			-0.0546	-0.0768	0.0452	-0.0954
			(0.5311)	(0.5421)	(0.5569)	(0.5224)
Short programme (up to 4 months)			0.8051**	0.7990**	0.7728*	0.7464*
			(0.3314)	(0.4036)	(0.4246)	(0.4143)
Sample characteristics (baseline: pooled estimation for men and women, 1980s or 1990s pro-gramme)						
Separate estimation for males				0.2034	-0.0472	0.2282
				(0.4502)	(0.4486)	(0.4762)
Separate estimation for females				-0.258	-0.5787	-0.244
				(0.4641)	(0.4812)	(0.4784)
2000s programme				0.1314	-0.3203	0.1842
				(0.3374)	(0.3653)	(0.3358)
Contextual controls						
Unemployment rate					0.2167***	
					(0.0790)	
GDP growth rate						0.0615
						(0.1074)
Number of observations	93	93	93	93	93	93
Pseudo R-squared	0.0708	0.0925	0.1217	0.1344	0.1777	0.1364
Akaike information crite-rion	181.9881	184.0131	182.6708	186.3419	180.4334	187.9807

The table displays estimated coefficients of ordered probit models. Heteroskedasticity-robust standard errors in parentheses. The dependent variable takes value 1 for significantly positive estimates, 0 for insignificant and -1 for significantly negative estimates. Unemployment and GDP growth rates are annual rates in % for the year the evaluated programme started. "Other countries" includes Netherlands and Switzerland. *** denotes $p < 0.01$, ** denotes $p < 0.05$ and * denotes $p < 0.1$.

Table 5: Sensitivity analysis

	Preferred	Variation
ALMP type (baseline: training)		
Wage subsidy	1.1301**	1.2955**
	(0.5116)	(0.6004)
Public works	-0.2983	-0.2629
	(0.4579)	(0.4553)
Services/sanctions	0.2386	0.2203
	(0.4049)	(0.4010)
Other programmes	-0.9004**	-1.0741**
	(0.3530)	(0.4343)
Method	yes	yes
Programme characteristics	yes	yes
Sample characteristics	yes	yes
Contextual controls	UE	UE
Number of observations	93	86
Pseudo R-squared	0.1777	0.1821
Akaike information criterion	180.4334	171.6131

The table displays the estimated coefficients of ordererd probit models. Heteroskedasticity-robust standard errors in parentheses. "Preferred" reproduces the results from column (5) in Table 4. "Variation" tightens the definition of ALMP and the short run, eliminating 6 studies. "Other programmes" includes aggregate ALMP effects and combined programmes. UE is for unemployment rate. *** denotes $p < 0.01$, ** denotes $p < 0.05$ and * denotes $p < 0.1$.

Appendix

Table A.1: Size and composition of the foreign-born population in European OECD countries, 2009/2010

| Country | Foreign-born individuals | | | | Born in: | | | | |
| | | | | | Africa | Asia | Latin America and the Caribean | United States, Canada and Oceania | Europe |
	Total number of individuals (thousands)	Percentage of the total population	Number of individuals aged 15-64 (thousands)	Percentage of the age group 15-64	(% of all foreign-born 15-64)				
Denmark	414	7.5	341	9.5	3.3	33.5	1.3	7.4	54.4
Finland	233	4.4	197	5.5	0.0	0.0	0.0	0.0	100.0
Germany	10 601	12.9	8271	15.1	3.3	15.6	1.3	1.6	78.2
Netherlands	1 833	11.1	1558	13.9	21.7	24.2	24.1	2.6	27.4
Norway	527	10.9	443	13.9	11.0	33.2	5.2	4.0	46.7
Sweden	1 338	14.4	1051	17.4	7.5	36.9	7.8	1.8	46.0
Switzerland	2 038	26.3	1649	31.3	5.8	8.2	6.1	2.8	77.1
7-country average (unweighted)		12.5		15.2	7.5	21.7	6.5	2.9	61.4
European OECD average (unweighted)		11.2		13.3	12.1	14.3	7.1	2.5	64.0

Source: OECD (2012) and own calculations

Table A.2: Sensitivity analysis, full results

	Preferred	Variation
ALMP type (baseline: training)		
Wage subsidy	1.1301**	1.2955**
	(0.5116)	(0.6004)
Public works	-0.2983	-0.2629
	(0.4579)	(0.4553)
Services/sanctions	0.2386	0.2203
	(0.4049)	(0.4010)
Other programmes	-0.9004**	-1.0741**
	(0.3530)	(0.4343)
Study characteristics (baseline: matching, working paper)		
Duration analysis	0.5544	0.5938
	(0.4542)	(0.4915)
Other method	0.2542	0.405
	(0.5948)	(0.6215)
Published paper	0.9968***	0.8905**
	(0.3727)	(0.4184)
Programme characteristics (baseline: regular ALMP, duration unknown or greater than four months)		
Migrant-specific programme	0.0452	0.1917
	(0.5569)	(0.5594)
Short programme (up to 4 months)	0.7728*	0.7713*
	(0.4246)	(0.4210)
Sample characteristics (baseline: pooled estimation for men and women, 1980s or 1990s programme)		
Separate estimation for males	-0.0472	-0.0039
	(0.4486)	(0.4888)
Separate estimation for females	-0.5787	-0.503
	(0.4812)	(0.5263)
2000s programme	-0.3203	-0.2076
	(0.3653)	(0.4511)
Contextual controls		
Unemployment rate	0.2167***	0.1947**
	(0.0790)	(0.0836)
Number of observations	93	86
Pseudo R-squared	0.1777	0.1821
Akaike information criterion	180.4334	171.6131

The table displays the estimated coefficients of ordererd probit models. Heteroskedasticity-robust standard errors in parentheses. "Preferred" reproduces the results from column (5) in Table 4. "Variation" tightens the definition of ALMP and the short run, eliminating 6 studies. "Other programmes" includes aggregate ALMP effects and combined programmes. *** denotes $p < 0.01$, ** denotes $p < 0.05$ and * denotes $p < 0.1$.

Table A.3: Estimation results with clustered standard errors

	(1)	(2)	(3)	(4)	(5)	(6)
ALMP type (baseline: training)						
Wage subsidy	0.5732	0.6987*	1.0664***	1.0561**	1.1301***	0.9855**
	(0.3832)	(0.3905)	(0.4109)	(0.4441)	(0.4156)	(0.4580)
Public works	-0.5024*	-0.4879	-0.1016	-0.1287	-0.2983	-0.1835
	(0.2672)	(0.3032)	(0.2662)	(0.3658)	(0.3835)	(0.3480)
Services/sanctions	0.5002	0.4474	0.342	0.2952	0.2386	0.2475
	(0.3823)	(0.3381)	(0.3608)	(0.4006)	(0.3973)	(0.4307)
Other programmes	-0.5985*	-0.6727*	-0.5925	-0.6039	-0.9004**	-0.7541*
	(0.3454)	(0.3880)	(0.3701)	(0.3780)	(0.3630)	(0.4337)
Study characteristics (baseline: matching, working paper)						
Duration analysis		-0.1694	-0.1524	-0.1252	0.5544	-0.1132
		(0.6052)	(0.6494)	(0.4951)	(0.5220)	(0.5019)
Other method		-0.3006	0.0912	0.1477	0.2542	0.1718
		(0.4608)	(0.4575)	(0.4055)	(0.4936)	(0.4028)
Published paper		0.6123	0.7224*	0.7099*	0.9968**	0.6462
		(0.4127)	(0.4282)	(0.3809)	(0.4033)	(0.3972)
Programme characteristics (baseline: regular ALMP, duration unknown or greater than four months)						
Migrant-specific pro-gramme			-0.0546	-0.0768	0.0452	-0.0954
			(0.4215)	(0.4801)	(0.5252)	(0.4634)
Short programme (up to 4 months)			0.8051**	0.7990**	0.7728**	0.7464*
			(0.3512)	(0.3653)	(0.3812)	(0.3948)
Sample characteristics (baseline: pooled estimation for men and women, 1980s or 1990s programme)						
Separate estimation for males				0.2034	-0.0472	0.2282
				(0.4847)	(0.4730)	(0.5173)
Separate estimation for females				-0.258	-0.5787	-0.244
				(0.5125)	(0.5277)	(0.5316)
2000s programme				0.1314	-0.3203	0.1842
				(0.4723)	(0.4942)	(0.4561)
Contextual controls						
Unemployment rate					0.2167***	
					(0.0783)	
GDP growth rate						0.0615
						(0.1226)
Number of observations	93	93	93	93	93	93
Pseudo R-squared	0.0708	0.0925	0.1217	0.1344	0.1777	0.1364
Akaike information crite-rion	181.9881	184.0131	182.6708	186.3419	180.4334	187.9807

The table displays estimated coefficients of ordered probit models. Standard errors (clustered by study) in parentheses. The dependent variable is an indicator taking value 1 for significantly positive estimates, 0 for insignificant and -1 for significantly negative estimates. Unemployment and GDP growth rates are annual rates in % for the year the evaluated programme started. "Other countries" includes Netherlands and Switzerland. *** denotes $p < 0.01$, ** denotes $p < 0.05$ and * denotes $p < 0.1$.

Table A.4: Sensitivity analysis for the estimation with clustered standard errors

	Preferred	Variation
ALMP type (baseline: training)		
Wage subsidy	1.1301***	1.2955***
	(0.4156)	(0.4753)
Public works	-0.2983	-0.2629
	(0.3835)	(0.3911)
Services/sanctions	0.2386	0.2203
	(0.3973)	(0.4005)
Other programmes	-0.9004**	-1.0741***
	(0.3630)	(0.3666)
Study characteristics (baseline: matching, working paper)		
Duration analysis	0.5544	0.5938
	(0.5220)	(0.5652)
Other method	0.2542	0.405
	(0.4936)	(0.5084)
Published paper	0.9968**	0.8905**
	(0.4033)	(0.4255)
Programme characteristics (baseline: regular ALMP, duration unknown or greater than four months)		
Migrant-specific programme	0.0452	0.1917
	(0.5252)	(0.5592)
Short programme (up to 4 months)	0.7728**	0.7713**
	(0.3812)	(0.3604)
Sample characteristics (baseline: pooled estimation for men and women, 1980s or 1990s programme)		
Separate estimation for males	-0.0472	-0.0039
	(0.4730)	(0.5862)
Separate estimation for females	-0.5787	-0.503
	(0.5277)	(0.6433)
2000s programme	-0.3203	-0.2076
	(0.4942)	(0.6543)
Contextual controls		
Unemployment rate	0.2167***	0.1947**
	(0.0783)	(0.0801)
Number of observations	93	86
Pseudo R-squared	0.1777	0.1821
Akaike information criterion	180.4334	171.6131

The table displays estimated coefficients of ordered probit models. Standard errors (clustered by study) in parentheses. "Preferred" reproduces results from column (5) in Table A.3. Variation tightens the definition of ALMP and the short run, eliminating 6 studies. "Other programmes" includes aggregate ALMP effects and combined programmes. *** denotes $p < 0.01$, ** denotes $p < 0.05$ and * denotes $p < 0.1$.