

What Active Labour Market Programmes Work for Immigrants in Europe?

Sebastian Butschek
(ZEW Mannheim)

Thomas Walter
(ZEW Mannheim)

This version: April 23, 2013

Abstract

There is a growing body of programme evaluation literature that recognises immigrants as a disadvantaged group on European labour markets and investigates the effects of Active Labour Market Programmes (ALMPs) on this subgroup. Using a meta-analysis, we condense the findings of 24 such studies on the relative effectiveness of five types of ALMPs employed across Europe to combat immigrant unemployment: training, public works, wage subsidies, services and sanctions, and combinations thereof. We find that only wage subsidies can be recommended with confidence to European policy-makers.

JEL-Codes: J15, J61, J68, I38

Key Words: immigrants, unemployment, labour market integration, ALMP, evaluation, meta-analysis

Acknowledgements: We thank Stephan L. Thomsen for insightful comments. All remaining errors are ours.

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 constitute a sizeable share of the working-age population in European OECD countries: in 2009/2010, foreign-born individuals made up 13.3 per cent of those aged 15 to 64. During this same period, the foreign-born did comparatively badly on the labour market: their employment rates were 4.6 per cent lower and their unemployment rates were 36.5 per cent higher than those of native workers (OECD, 2012). European governments have pursued two avenues toward improving immigrants' labour-market integration: on the one hand, they have offered programmes tailored to immigrants' specific needs, such as language and introduction courses, and on the other, they assign unemployed immigrants to regular Active Labour Market Programmes (ALMPs), including job search assistance, training and subsidised employment. Which of these activation measures work best is clearly of policy interest, a point also evidenced by the growing body of evaluation literature.

With respect to ALMPs' effects on all unemployed workers, recent analyses have strengthened a growing consensus: job-search assistance (services/sanctions) and, to some extent, subsidised work in the private sector (wage subsidies) are effective in the short run while training works in the longer run; subsidised public sector employment, 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, albeit with contrasting conclusions (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 introduction/language courses designed for immigrants and eight recent evaluations of 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 easily accessible quantitative summary of the existing empirical evidence. To this end, we start by defining a sample of relevant studies and then proceed to con-

dense the findings extracted from them in two steps: first, using simple descriptive analysis and second, performing a meta-analysis.

The key step in generating a data set for such an analysis is the selection of studies. We choose to focus on micro-econometric papers estimating an individual-level ALMP treatment effect on immigrants' labour market outcomes. For comparability, we require that the intervention evaluated fit into one of five ALMPs categories: training, public works, wage subsidies, services/sanctions or a combination. These criteria give us a sample of 79 estimates from 24 separate studies of ALMPs implemented in the Nordic countries, Germany, the Netherlands and Switzerland between 1984 and 2007. The restriction of both time period and geographic scope is an unintended consequence of the availability of literature meeting our criteria.

Our descriptive analysis looks at the distribution of impact estimates conditional on study characteristics with two goals: first, run a preliminary check that region, time period, or publication status do not completely explain which studies find significant effect estimates; and second, 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 all at the same time when investigating which ALMP types are associated with significant impact findings, simultaneously addressing such issues as publication bias or changes in programme effectiveness over time. Second, we can provide a summary measure for whether the evidence suggests that a type of ALMP works better than some other one, namely a significant coefficient among the independent variables of the ordered probit model we estimate.

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

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; section 5 concludes.

2 Immigrants in Europe and their labour market integration

On average, the share of foreign-born persons among the total population amounted to 11.2% in European OECD countries in 2009/2010¹ (see Table A.1 in the appendix; OECD, 2012). The number of immigrants is particularly high in Austria, Estonia, Ireland, Luxemburg and Switzerland, with a share of more than 15%. 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.

The vast majority of immigrants aged 15 to 64 is from Europe: on average, almost two thirds of the foreign-born. Yet, there is substantial heterogeneity in the origin of immigrants across countries. While in France more than half of all foreign-born persons aged 15 to 64 come from Africa, about 80% of all immigrants of working age in Austria, Germany and Switzerland were born in Europe.

Despite the 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 across European OECD countries in 2009/2010 and how they compare to the respective rates of the native populations. On average, the employment rate of immigrants amounts to 63.2%. It is 2.9 percentage points lower than the rate of natives. Differences in the employment rate are especially pronounced in Belgium, Denmark, the Netherlands, Poland and Sweden, with a difference of more than 10 percentage points. However, there are also countries in which the employment rate of immigrants is larger than the employment rate of natives (Czech Republic, Estonia, Greece, Hungary, Italy, Luxemburg, Portugal and Turkey).

The unemployment rate of immigrants is in almost all countries larger than the unemployment rate of natives. There are only two exceptions: Hungary and the Slovak Republic. The average unemployment rate of immigrants amounts to 12.6% across European OECD countries. It is, on average, 4.3 percentage points larger than the native unemployment rate. The differences between immigrants and natives are most pronounced in Belgium, Finland, Spain and Sweden.

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

Table 1: Employment and unemployment rates of immigrants in 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)
Austria	65.5	-7.5	8.9	5.1
Belgium	52.6	-10.8	16.7	9.9
Czech Republic	66.9	1.8	8.4	1.4
Denmark	65.6	-10.0	11.8	5.5
Estonia	63.5	1.4	18.7	3.5
Finland	62.1	-6.6	16.3	8.2
France	57.8	-7.1	14.5	6.0
Germany	63.8	-8.7	12.2	5.6
Greece	65.0	5.2	14.1	3.3
Hungary	65.5	10.2	8.3	-2.4
Iceland	75.9	-2.5	12.6	5.5
Ireland	60.8	-0.1	16.1	3.9
Italy	62.3	5.7	11.2	3.4
Luxembourg	70.0	8.7	6.4	3.3
Netherlands	65.5	-11.9	7.7	4.2
Norway	66.6	-9.8	9.9	7.0
Poland	47.9	-11.4	11.5	2.5
Portugal	69.5	3.9	14.0	3.7
Slovak Republic	58.8	-0.7	12.5	-0.7
Slovenia	65.6	-1.3	8.5	2.0
Spain	57.4	-2.1	28.1	11.1
Sweden	61.7	-12.9	15.8	8.7
Switzerland	75.1	-5.1	7.4	4.2
Turkey	48.4	3.2	13.9	2.2
United Kingdom	66.1	-4.2	8.9	1.3
Average (unweighted)	63.2	-2.9	12.6	4.3

Source: OECD (2012) and own calculations

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

General programmes comprise four types of interventions (see e.g. Card et al., 2010, or Kluve, 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*: This type of intervention aims at offering temporary job opportunities outside the private sector, mainly for community services. Job opportunities are not allowed to compete with regular employment. The key objective is to maintain the employability of participants.
- 4) *Job search assistance and sanctions*: This intervention type has the objective to make 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.

The four intervention types need not be offered separately. They can also be combined. For example, job search assistance and training could be offered in a combined programme providing counselling on job opportunities in a certain occupation and training of specific skills for that occupation at the same time.

Migrant-specific programmes can be grouped into three categories:

- 1) *Language training*: This intervention type aims at improving the participants' ability to communicate in the host country's language. Language courses often also provide 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*: This intervention type provides an individualised integration plan towards employment uptake. It is mainly targeted at newly arriving immigrants and combines language training with one or more general

activation programmes. Introduction programmes usually start with language training and then continue either with a training programme or subsidised employment or a combination thereof depending upon the participant's needs. Throughout the programme job search assistance is provided. See e.g. Andersson Joonas and Nekby (2012) and Sarvimäki and Hämäläinen (2012) for introduction programmes in Sweden and Finland, respectively.

- 3) *General programmes exclusively for immigrants*: This intervention type comprises all general ALMPs (training, subsidised private or public sector employment, job search assistance and sanctions) which are only targeted at immigrants (and not at natives) and which do not include language training. One example for such an intervention is intensified job search assistance programmes. In such programmes, immigrants are assigned to caseworkers whose caseload is reduced, i.e. 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 the latter to be more successful since they are specifically designed for the needs of immigrants whereas general programmes address the needs of average native participants. 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 including a dummy variable for whether or not an intervention was designed specifically for migrants. However, because only three of the studies in our sample evaluate migrant-specific programmes, this analysis is limited. We effectively focus on the effect of general ALMPs on immigrants.

² If the migrant-specific programs 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.

3 Description of the Data

3.1 The Estimation Sample

In a first step, we attempted to obtain an exhaustive sample of studies evaluating the effects of ALMPs on migrants' labour market outcomes. Our search strategy was as follows, yielding a total of 56 papers:

- 1) Collect studies on ALMPs surveyed by Nekby (2008), Rinne (2012), Greenberg, Kluve (2010), Card et al. (2010). We identified 27 papers considering migrants as a sub-group.
- 2) Perform a Google Scholar keyword search³ on 27 November 2012, yielding 27 more studies potentially meeting our criteria.

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

- 1) Studies that estimate ALMP treatment effects for immigrants. (Some studies analyse heterogeneous effects for this subgroup while others have a sample of only immigrants). We excluded 14 studies based on this criterion.
- 2) Studies that perform micro-econometric evaluation of the intervention's effect on individual labour market outcomes, outlining the identification strategy. We dropped 9 studies following this rule.
- 3) Studies that evaluate an intervention that fits into one of four ALMP categories (described more fully below) or a combination thereof: training, wage subsidy, public works, or services/sanctions. Based on this criterion, we excluded 7 studies.

Applying these criteria yielded a sample of 26 studies estimating ALMP effects on migrants' probability of or hazard to employment.⁴ 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 one year after pro-

³ Keywords, in different combinations: ALMP, labour market programmes, labor market programs, migrants, foreign, native, born, citizen, subgroup, sub-group, hetero.

⁴ See the appendix for the list of 24 studies analysed; instead of looking at employment, one study considers earnings and another, focusing on transitions to self-employment, uses being neither unemployed nor in receipt of unemployment benefits as an outcome variable.

gramme participation⁵. Where there are more than one such short-term estimates per gender-region-group combination, we choose the latest (most long-term) one. This gives 79 estimates in 24 study clusters.⁶

The four ALMP categories we use are as follows:

- 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 and 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). Like Card et al. (2010), we allow for a fifth group of programmes, namely combinations of the four ALMP types defined above. See also section 2 for a definition of these ALMP categories.

From our sample of 24 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 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.⁷

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 (41 estimates), followed by ones finding significantly positive effects (27), with significantly negative effect estimates less frequent still (11 estimates).

Next, lines 2a) to 2e) show that among ALMP types, public works feature prominently, contributing 22 data points; column 3 reveals that this pattern is driven by German evalua-

⁵ While we do have information on longer-run outcomes (38 estimates), there is not enough variation in them to permit a separate (ordered) probit analysis. That is, various programme types perfectly predict outcomes.

⁶ Two studies report only long-term estimates for 24 months after the programme and beyond (Caliendo and Künn, 2010, and Groß et al., 2006).

⁷ We categorised PhD dissertations as published studies because of the similarities between PhD supervision and the referee process.

tions. Training programmes dominate (27), with both wage-subsidy (12) and services/sanctions (13) interventions each providing about half as many observations as the largest category. There are only 5 estimates for programmes that combine several ALMP types in a single treatment.

Lines 3a) through 3c) show that those short-run programmes with standardised length for which we know their duration are evenly split between four months or shorter (21) and five months or longer (19). Most of our short-run estimates, however, are for programmes of unknown or mixed duration (39), reflecting some heterogeneity in the level of detail on interventions given in the papers.

Lines 5a) through 5c) illustrate that matching approaches were clearly the most popular method, contributing almost two thirds of the sample of estimates (50); most remaining estimates were based on duration analysis (21). 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) summarise studies' publication status to address the question whether there might be a tendency to publish only significant results ("publication bias"). Less than a third of the short-run estimates came from published papers (23), with Nordic evaluations accounting for disproportionately many publications (17)⁸, which may be in part explained by the fact that Nordic studies are older on average. This can be seen from lines 4a) through 4c), which also reveal that about two thirds of estimates are from the 2000s.

Comparing estimates by origin reveals that the largest contributor, Germany, differs markedly from Denmark, Finland 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. Programme duration is unknown or mixed for all of the Nordic interventions we have estimates for. German estimates are based on more recent data than either one of the other groups of countries. As pointed out before, German estimates are almost exclusively from matching analyses and account for the great number of estimates for public works programmes. In contrast, there is more variety in the methods used to evaluate Nordic and Swiss/Dutch programmes and training is the dominant ALMP type in those two country groups.

⁸ We do not intend to suggest that published papers meet different quality standards, 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).

Table 2: Characteristics of the estimation sample

	Short-run estimates	Nordic countries	Estimates for Germany	Other countries
<u>1) Estimated programme effect</u>				
a) Negative	11	2	1	8
b) Insignificant	41	7	29	5
c) Positive	27	13	12	2
<u>2) ALMP type</u>				
a) Training	27	8	12	7
b) Wage subsidy	12	5	3	4
c) Public works	22	4	15	3
d) Services/Sanctions	13	2	10	1
e) Combined programmes	5	3	2	0
<u>3) Programme duration</u>				
a) Up to 4 months	21	0	18	3
b) 5 or more months	19	0	15	4
c) Mixed/unknown	39	22	9	8
<u>4) Time evaluated programme ran</u>				
a) 1980s	2	2	0	0
b) 1990s	26	11	0	15
c) 2000s	51	9	42	0
<u>5) Method employed</u>				
a) Matching	50	1	41	8
b) Duration	21	14	0	7
c) Other method	8	7	1	0
<u>6) Publication status</u>				
a) Working paper	56	6	36	14
b) Published	23	16	6	1
Number of estimates	79	22	42	15

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

3.3 Descriptive Analysis

Here we discuss the distribution of the outcome variable (short-run effect: negative, insignificant or positive) conditional on the covariates we extracted from the studies. This subsection 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 eval-

uated. This is important because our meta-analysis, by virtue of its method, only allows conclusions about the relative effectiveness of different types of programme.

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 9 out of 12 estimates are positive. For training, almost half of the estimates are insignificant; for services/sanctions and public works, about two thirds are insignificant. For both training and services/sanctions, positive estimates are clearly more frequent than negative ones, while the converse is true for public works and combination treatments, though less obviously so.

The next three lines relate the method used to the distribution of effect estimates. Over two thirds of matching estimates are insignificant, with most of the remaining third positive. In contrast, both duration and other models find positive effects in about half of their estimations. The remaining duration results are more frequently negative than insignificant while the “other” category provides more insignificant estimates than negative ones. These statistics seem to suggest that duration analysis is more optimistic about programme effectiveness than matching approaches are; given, however, that the degree of correlation between, e.g., matching and Germany is quite high, assessment of potential methodological bias should await multivariate analysis.

Most estimates are for programmes implemented in the 1990s and 2000s, with only two data points from the 1980s. 1990s estimates are fairly evenly distributed across negative, insignificant and positive effects. Over two thirds of the estimates for the 2000s are insignificant, with the rest almost exclusively positive. This may point to a deterioration of ALMP quality over time. Yet, note that that the pattern in line 3c) closely mirrors that for German estimates, so that regional differences may be confounding the influence of time. Again, it will be necessary to hold other covariates constant to be able to judge whether programmes became more or less effective over time.

Conditional on programme duration, ALMP effect distributions are skewed to the right for short treatments (up to 4 months) and those with mixed or unknown duration, with negative estimates making up only about a tenth of the total. Programmes lasting over 5 months produce insignificant effect estimates in three out of four cases and no positive ones. Considering lock-in effects, it is of course possible that shorter interventions tend to be more effective than longer ones.

Finally, the conditional distribution of effect estimates is skewed to the right for both the working-paper and the published subsample. However, it appears that this is more pro-

nounced for published papers, where over two thirds provide positive estimates, than for working papers, out of which only less than one third estimates a positive effect and a more substantial fraction of estimates is negative. While this may be interpreted as an indication of publication bias, a substantial portion of working papers study more recent programmes and will probably still get published. Yet again, multivariate analysis should help clarify this.

Table 3: Distribution of the estimated programme effects

	significantly negative	Estimated effect is insignificant	significantly positive
<u>1) ALMP type</u>			
a) Training	4	13	10
b) Wage subsidy	2	1	9
c) Public works	4	15	3
d) Services/Sanctions	0	8	5
e) Combined programmes	1	4	0
<u>2) Method employed</u>			
a) Matching	4	34	12
b) Duration	6	4	11
c) Other method	1	3	4
<u>3) Time evaluated programme ran</u>			
a) 1980s	0	0	2
b) 1990s	9	7	10
c) 2000s	2	34	15
<u>4) Programme duration</u>			
a) Up to 4 months	2	10	9
b) 5 or more months	5	14	0
c) Mixed/unknown	4	17	18
<u>5) Publication status</u>			
a) Working paper	9	30	17
b) Published	2	11	10
Number of estimates	11	41	27

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 one year after programme participation. Where there are more than one such short-term estimates, the latest (most long-term) one is sampled.

The results of this descriptive analysis can be summarised as follows:

- Evaluations of private-sector subsidised employment programmes tend to reach positive conclusions on their short-run effects on migrants' employment.

- Insignificant effect estimates dominate the evaluation of the other ALMP types; those evaluations that do present significant findings are more positive about the short-run effects of training and services/sanctions than of public works and combination programmes.
- While there are some indications of publication bias, methodological bias and programme deterioration over time, we emphasise the univariate nature of this simple analysis and postpone assessment of potential biases to the probit analysis in section 4.2.

4 Empirical Analysis

4.1 Method

We perform an ordered probit analysis with sign and significance of the estimate as the outcome variable. It 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 dummies 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 combination.

While it would be desirable to take effect sizes into account (e.g. Greenberg et al., 2003), plurality of underlying econometric estimation techniques and a small sample size make this unfeasible. The simplified 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. Card et al. (2010) show 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 yield similar results. 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.

The index model underlying our estimation is as follows:

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

where *WS*, *PW*, *SE* and *CO* are dummy variables describing the programme type analysed in study *i* (wage subsidy, public works, services/sanctions, or combined programmes, with training being the omitted category), *X* is a vector of control variables (method employed, program characteristics, sample characteristics, institutional setting) and *u* is an error term.

4.2 Estimation results

We estimate 5 different specifications, gradually introducing groups of control variables. Specification 1 includes only the type of programme, omitting training. Specification 2 adds study characteristics: whether the method was duration analysis or some other econometric technique (omitted: matching), and whether the paper is published (baseline: working paper). Column 3 introduces two more programme characteristics, namely whether the intervention was designed for migrants and whether the treatment was short, that is, no longer than four months. For the latter, the omitted category lumps together programmes of longer and unknown/mixed duration. In specification 4, sample characteristics enter the equation: participant gender (baseline: mixed) and treatment in the 2000s (omitted: 1980s or 1990s). The final column, specification 5, adds the unemployment rate at the time that the evaluated programme started as a proxy for the macroeconomic context.⁹

Table 4 presents estimation results from the five specifications. Wage subsidies are estimated to be significantly more likely than training programmes to obtain a positive evaluation result. For combination programmes, the estimated probability of being evaluated negatively is significantly higher than for training in some specifications. Point estimates are consistently insignificant for services/sanctions. They are always negative but mostly insignificant for public works. Almost all additional controls have insignificant coefficient estimates.

In specification 1, wage subsidy and services/sanctions are not significant, whereas public works and combinations are. When we add study characteristics in specification 2, the coefficients on ALMP types remain similar in magnitude but only wage subsidy is now significant. The additional controls are insignificant. Specification 3's programme features cause the wage subsidy coefficient to grow, while retaining significance; all other ALMP type coefficients remain insignificant.

⁹ The unemployment rate was obtained from the Online OECD Employment database; see <http://www.oecd.org/employment/employmentpoliciesanddata/onlineoecdemploymentdatabase.htm> (accessed on 7 January, 2013).

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). However, the combinations coefficient increases in magnitude and becomes significant.

Table 4: Estimation results

	(1)	(2)	(3)	(4)	(5)
<u>ALMP type (baseline: training)</u>					
Wage subsidy	0.7245 (0.4565)	0.8071** (0.3957)	1.0599*** (0.3984)	0.9985*** (0.3848)	1.0195*** (0.3624)
Public Works	-0.4673* (0.2834)	-0.428 (0.3102)	-0.1359 (0.2602)	-0.2709 (0.3647)	-0.4604 (0.3772)
Services/Sanctions	0.2914 (0.3537)	0.2518 (0.3009)	0.2238 (0.3187)	0.0278 (0.3522)	-0.0643 (0.2964)
Combined programmes	-0.7342* (0.3927)	-0.7638 (0.4930)	-0.7815 (0.5000)	-1.1203** (0.4657)	-1.2536*** (0.4469)
<u>Method employed and publication status (baseline: matching, working paper)</u>					
Duration analysis		-0.0934 (0.8802)	-0.058 (0.9283)	0.4794 (0.7360)	0.9347 (0.6608)
Other method		0.0698 (0.5503)	0.3339 (0.5779)	0.9647* (0.5329)	0.9504* (0.5662)
Published		0.3885 (0.5564)	0.4964 (0.5490)	0.0576 (0.4676)	0.4331 (0.5219)
<u>Programme characteristics (baseline: programme also available to natives, programme of long or mixed/unknown duration)</u>					
Migrant-specific programme			-0.0322 (0.4694)	-0.263 (0.5658)	-0.3033 (0.5607)
Duration up to 4 months			0.5824 (0.3950)	0.4928 (0.3298)	0.3947 (0.3301)
<u>Sample characteristics (baseline: pooled estimation for men and women, programme ran in the 1980s or 1990s)</u>					
Separate estimation for men				0.0439 (0.6130)	-0.1134 (0.5951)
Separate estimation for women				-0.5214 (0.6071)	-0.6931 (0.6008)
Programme ran in the 2000s				0.8098 (0.5507)	0.2005 (0.5623)
<u>Macroeconomic context</u>					
Unemployment rate at programme start					0.2088** (0.0865)
Number of observations	79	79	79	79	79
Pseudo-R ²	0.0686	0.0806	0.0961	0.1337	0.1707
Akaike information criterion	156.4835	160.6299	162.2255	162.3942	158.6513

Remarks: The table displays the estimated coefficients of ordered probit models. Standard errors (clustered by study) are in parantheses. The dependent variable is an indicator for the sign and significance of the sampled estiamtes of programme effects. It is 1 for significantly positive estimates, 0 for insignificant estimates and -1 for significantly negative estimates.

*** denotes $p < 0.01$, ** denotes $p < 0.05$ and * denotes $p < 0.1$.

Specification 5 is our preferred specification, as it also includes a contextual control variable. The coefficient on wage subsidy remains positive, highly significant and similar in size. Similarly, that on combinations remains negative and significant. The coefficient esti-

mates on public works and services/sanctions are still insignificant. The coefficient on the unemployment rate is positive and significant, suggesting that inferior macroeconomic conditions at the time of treatment predict a higher probability of a positive evaluation result.¹⁰

As the services/sanctions coefficient is consistently insignificant, we do not worry about the variation in the sign. The public works coefficient is insignificant almost throughout so we cannot fully confirm for migrants the negative verdict of other meta-analyses on the short-run effect of this type of ALMP. There are only five data points for combination programmes; four out of these five studies estimate an insignificant short-run effect. We therefore lean toward attributing the fluctuations in the coefficient's magnitude to the shortage of variation and do not put much weight on this result. Finally, the wage subsidy coefficient is fairly robust to different specifications of the model.

It is worth reiterating 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 and that combinations might do worse. Because the corresponding coefficients are insignificant, our results do not provide a clear summary of what the underlying evaluation studies have found on the effectiveness of public works and services/sanctions as compared to training.

Our meta-analytic conclusions broadly confirm what section 3's descriptive analysis suggests. They add confidence that formally taking into account contextual and methodical factors does not alter the positive verdict on wage subsidies; but they fail to clarify the muddled picture the descriptive analysis gives on the other ALMP types.

These findings are based on a smaller sample of studies and on a more specific group of programme participants than the meta-analyses of Kluge(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 *Sensitivity checks*

Table A.2 in the appendix presents the results of four alternatives to our preferred specification (now in column 1).

In column 2, we replace the unemployment rate at the start of the intervention with the GDP growth rate for the year in which the programme started. The coefficient estimate on this

¹⁰ 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.

control variable is insignificant and the ALMP type estimates change size but their sign and significance remains the same.

In columns 3 and 4, we introduce region dummies in order to control for systematic differences not already captured by sample characteristics and contextual controls. We do this both with and without the method dummies to address the collinearity between the matching and Germany dummies. Comparing these columns with the preferred specification in (1) shows differences in coefficient magnitude but the same picture in terms of the sign and significance of ALMP types.

The final column presents the results of a probit model classifying significantly positive estimates as one and insignificant and significantly negative ones as zero. The combined-programme coefficient cannot be estimated as the variable perfectly predicts failure (now that insignificant and negative estimates are grouped together). However, wage subsidy remains the only ALMP type with a significant coefficient estimate, confirming the picture from our preferred specification: wage subsidy programmes are significantly more likely to have a positive estimated effect.

5 Conclusion

In this paper, we present a meta-analysis of studies estimating the treatment effect of ALMPs on migrants. We first show that migrants constitute an important group on European labour markets in terms of both the risks they face and the potential they harbour: migrants are numerous and over-represented in unemployment on the one hand but younger than the native population on the other. This provides a rationale for investigating the effectiveness of policy measures for the labour market integration of unemployed immigrants.

On the one hand, our study contributes to the ALMP evaluation literature (see, e.g. Kluge (2010) and Card et al. (2010) for recent meta-analyses) by focusing on a vulnerable subgroup of unemployed workers. On the other hand, our paper adds a quantitative analysis of the existing literature to the qualitative reviews of labour market policy targeting migrants by Nekby (2008) and Rinne (2010).

Using the 79 effect estimates extracted from 24 relevant studies we found, we perform a meta-analysis of the evaluation results. An ordered probit analysis based on the sign and significance of the effect estimate suggests that wage subsidy programmes work better for migrants than do training programmes. Public works and combined programmes tend to be

less effective than training, but estimated coefficients are not significant in all the specifications. We find no significant differences between training and services and sanctions. To help interpret these relative statements, consider the results of the descriptive analysis: there, our baseline category, training, has mostly insignificant effect estimates, suggesting that there is no short-run effect.

Taken together, the descriptive analysis of our sample of 79 estimates and the meta-analytic results suggest that, in the short run, wage subsidy programmes work best for European migrants. Many of the underlying estimates for the remaining programme types are insignificant. However, both descriptive and the meta-analysis seem to suggest that public works and combined programmes perform worse than training and services and sanctions.

At this point, only wage subsidies can be confidently recommended to policymakers seeking to improve the labour market integration of migrants through general ALMPs. Yet, there might be potential in migrant-specific interventions on which the evidence is still scarce; further research in this area is warranted.

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

Studies used in meta-analysis:

- 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 Economics* 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.
- 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): “Microeconomic 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.
- 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.
- 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.
- Sarvimäki, Matti and Kari Hämäläinen (2012): “Assimilating immigrants. The impact of an integration program.” Updated version of HECER Discussion Paper No. 306.

- 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*.

Appendix

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

Country	Foreign-born persons				Born in:				
	Total number of persons (thousands)	Percentage of the total population	Number of persons aged 15-64 (thousands)	Percentage of the age group 15-64	Africa	Asia	Latin America and the Caribbean	United States, Canada and Oceania	Europe
					(% of all foreign-born 15-64)				
Austria	1 293	15.5	1051	18.5	4.0	12.1	2.1	1.2	80.6
Belgium	1 376	12.7	1074	15.0	33.7	9.0	3.0	1.4	52.9
Czech Republic	676	6.4	492	6.6	0.9	8.9	0.5	0.8	88.8
Denmark	414	7.5	341	9.5	3.3	33.5	1.3	7.4	54.4
Estonia	222	16.6	135	14.9	0.0	5.4	0.0	0.1	94.6
Finland	233	4.4	197	5.5
France	7 235	11.6	5448	13.3	54.5	10.1	3.5	1.3	30.5
Germany	10 601	12.9	8271	15.1	3.3	15.6	1.3	1.6	78.2
Greece	858	7.9	769	10.6	2.2	20.7	0.2	2.3	74.5
Hungary	407	4.1	297	4.3	2.8	10.7	2.1	2.6	81.9
Iceland	35	11.1	29	13.5	2.5	13.7	2.4	5.6	75.9
Ireland	767	17.2	632	20.8	8.1	11.1	1.7	4.2	74.9
Italy	4 730	7.9	4136	10.5	22.6	17.1	14.6	2.2	43.5
Luxembourg	182	36.9	150	44.5	5.1	2.7	2.3	1.4	88.4
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
Poland	307	0.8	88	0.3	2.4	10.6	0.8	4.6	81.7
Portugal	673	6.3	588	8.3	44.7	2.2	23.4	1.8	27.9
Slovak Republic	38	0.7	28	0.7	0.5	1.9	0.6	0.4	96.7
Slovenia	161	7.9	129	9.1	0.4	0.1	0.5	0.6	98.5
Spain	6 567	14.3	5833	18.5	20.8	5.1	54.5	0.5	19.1
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
Turkey	2 066	2.9	1598	3.4
United Kingdom	6 899	11.3	5589	13.8	20.5	35.7	4.8	6.8	32.2
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

	(1)	(2)	(3)	(4)	(5)
<u>ALMP type (baseline: training)</u>					
Wage subsidy	1.0195*** (0.3624)	0.8301** (0.3815)	1.4860*** (0.5482)	1.6130*** (0.5698)	1.3860*** (0.4321)
Public Works	-0.4604 (0.3772)	-0.4407 (0.3592)	-0.3363 (0.4399)	-0.1184 (0.5212)	-0.1966 (0.5184)
Services/Sanctions	-0.0643 (0.2964)	-0.1808 (0.4524)	0.4753 (0.5995)	0.4783 (0.5799)	-0.029 (0.4569)
Combined programmes	-1.2536*** (0.4469)	-1.5971*** (0.4624)	-1.6836*** (0.3947)	-1.6028*** (0.3872)	
<u>Method employed and publication status (baseline: matching, working paper)</u>					
Duration analysis	0.9347 (0.6608)	0.5296 (0.7780)	0.3681 (0.9506)		1.6470** (0.6613)
Other method	0.9504* (0.5662)	1.1243** (0.5550)	-0.7851 (0.9607)		2.0432** (0.9860)
Published	0.4331 (0.5219)	-0.2396 (0.5389)	-0.3143 (0.6327)	-0.3888 (0.6290)	-0.2208 (0.7921)
<u>Programme characteristics (baseline: programme also available to natives, programme of long or mixed/unknown duration)</u>					
Migrant-specific programme	-0.3033 (0.5607)	-0.4581 (0.5165)	0.2423 (1.2932)	-0.6586 (0.9167)	-0.5723 (1.0311)
Duration up to 4 months	0.3947 (0.3301)	0.303 (0.3916)	0.8473** (0.4176)	1.0187** (0.4690)	1.0336*** (0.3548)
<u>Sample characteristics (baseline: pooled estimation for men and women, programme ran in the 1980s or 1990s)</u>					
Separate estimation for men	-0.1134 (0.5951)	0.045 (0.6518)	-0.6513 (0.6531)	-0.4834 (0.5793)	-0.3266 (0.6324)
Separate estimation for women	-0.6931 (0.6008)	-0.54 (0.6301)	-1.4018** (0.6818)	-1.2175* (0.6226)	-1.0904** (0.5426)
Programme ran in the 2000s	0.2005 (0.5623)	1.0204* (0.5914)	-1.5942* (0.8184)	-0.9213 (0.5701)	0.07 (0.7506)
<u>Macroeconomic context</u>					
Unemployment rate at programme start	0.2088** (0.0865)		0.0918 (0.1800)	0.1323 (0.1943)	0.152 (0.1244)
GDP growth rate (year of programme start)		0.1737 (0.1372)			
<u>Country/Region (baseline: Nordic countries)</u>					
Germany			-0.0717 (2.0983)	-1.0105 (1.3781)	
Other countries			-3.6567*** (1.0702)	-3.3486*** (0.9520)	
Constant					-2.3534*** (0.8404)
Number of observations	79	79	79	79	74
Pseudo-R ²	0.1707	0.1474	0.349	0.3371	0.2905
Akaike information criterion	158.6513	162.257	134.9905	132.8333	94.9055

Remarks: In columns (1) to (4), the table displays the estimated coefficients of ordered probit models. The dependent variable is an indicator for the sign and significance of the sampled estimates of programme effects. It is 1 for significantly positive estimates, 0 for insignificant estimates and -1 for significantly negative estimates. Column (1) is the preferred specification which is also displayed in Table 4. Column (5) shows the estimated coefficients of a probit model, in which the dependent variable is 1 for significantly positive estimates and 0 otherwise. Standard errors (clustered by study) are in parentheses. *** denotes $p < 0.01$, ** denotes $p < 0.05$ and * denotes $p < 0.1$. Nordic countries include Denmark, Finland and Sweden. Other countries include the Netherlands and Switzerland.