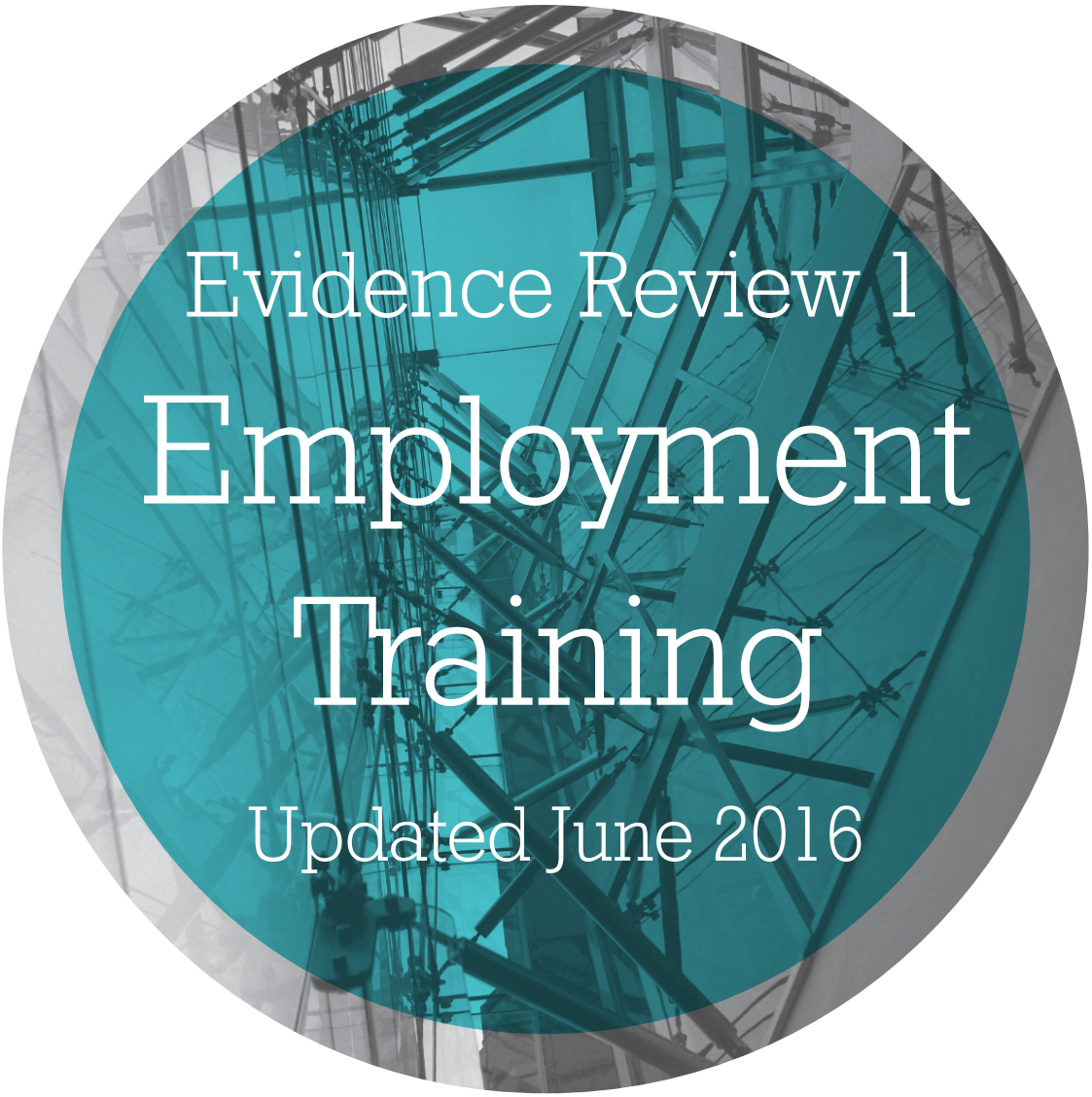


Evidence Review 1

Employment Training

Updated June 2016

what works centre for
local economic growth

00

Contents

01

Preface

This report presents findings from a systematic review of evaluations of employment training programmes aimed at improving labour market outcomes.

It is the first of a series of reviews that will be produced by the What Works Centre for Local Economic Growth. The What Works Centre is a collaboration between the London School of Economics and Political Science, Centre for Cities and Arup and is funded by the Economic & Social Research Council, The Department for Communities and Local Government and The Department for Business Innovation and Skills.

These reviews consider a specific type of evidence – **impact evaluation** – that seeks to understand the causal effect of policy interventions and to establish their cost-effectiveness. To put it another way they ask 'did the policy work' and 'did it represent good value for money'?

By looking at the details of the policies evaluated we can also start to answer questions about delivery issues – for example, whether policies offering in-firm training perform better or worse than those that offer classroom-based training.

Evidence on impact and effectiveness is clearly a crucial input to good policy making. Process evaluation – looking in detail at *how* programmes operate day to day – provides a valuable complement to impact evaluation, but we deliberately do not focus on this. We recognise that may sometimes cause frustration for practitioners and decision-makers who are responsible for the delivery of policy.

However, we see these impact-focused reviews as an essential part of more effective policy making. We often simply do not know the answers to many of the questions that might reasonably be asked when designing a new policy – not least, does it work? Figuring out what we do know allows us to better design policies and undertake further evaluations to start filling the gaps in our knowledge. **This also helps us to have more informed discussions about process and delivery issues and to improve policy making.**

These reviews therefore represent a first step in improving our understanding of what works for local economic growth. In the months ahead, we will be working with local decision-makers and practitioners, using these findings to help them generate better policy.

Henry Overman
Director, What Works Centre for Local Economic Growth

02

Executive Summary

This report presents findings from a systematic review of evaluations of training programmes aimed at improving adult skills and labour market outcomes.

It is the first of a series of reviews that will be produced by the What Works Centre for Local Economic Growth.

The review considered almost **1,000** policy evaluations, evidence reviews and meta-analyses from the UK and other OECD countries.

It found **71** impact evaluations which met the Centre's minimum standards. This represents a relatively large evidence base compared to many other local economic growth policies. But it is a small base relative to that available for some other policy areas (e.g. medicine, aspects of international development, education and social policy).

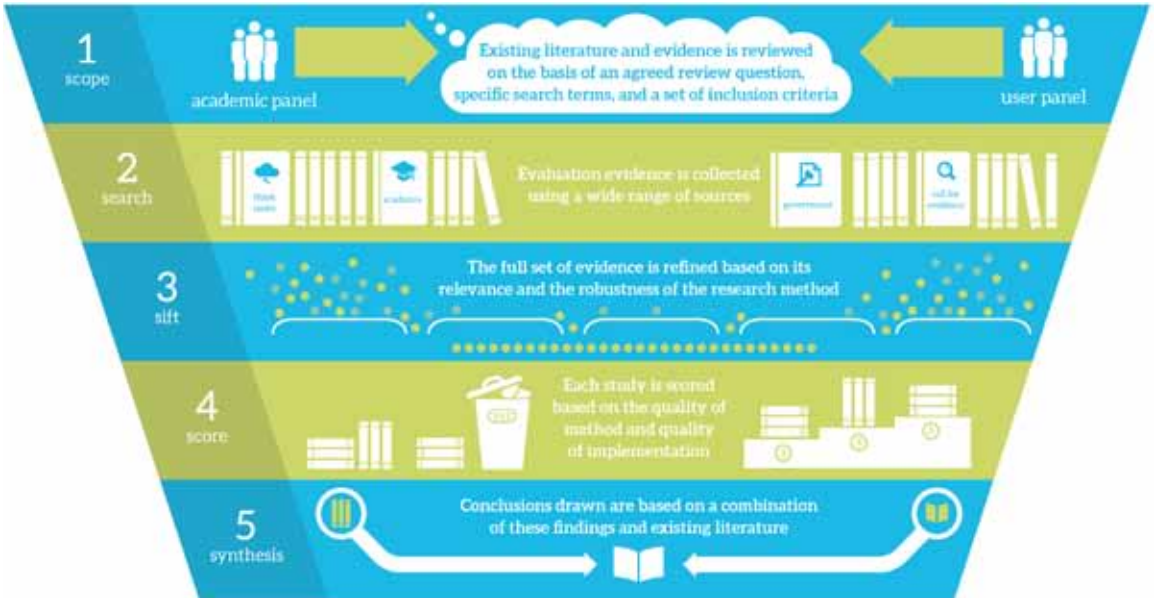
We define 'employment training' programmes as:

- ✓ including training targeted at the over 18s
- ✓ including day-release and short courses, and retraining
- × excluding training in schools, HE and apprenticeships
- × excluding specifically targeted training e.g. for those with mental health problems, ex-convicts, or particular ethnic groups

Approach

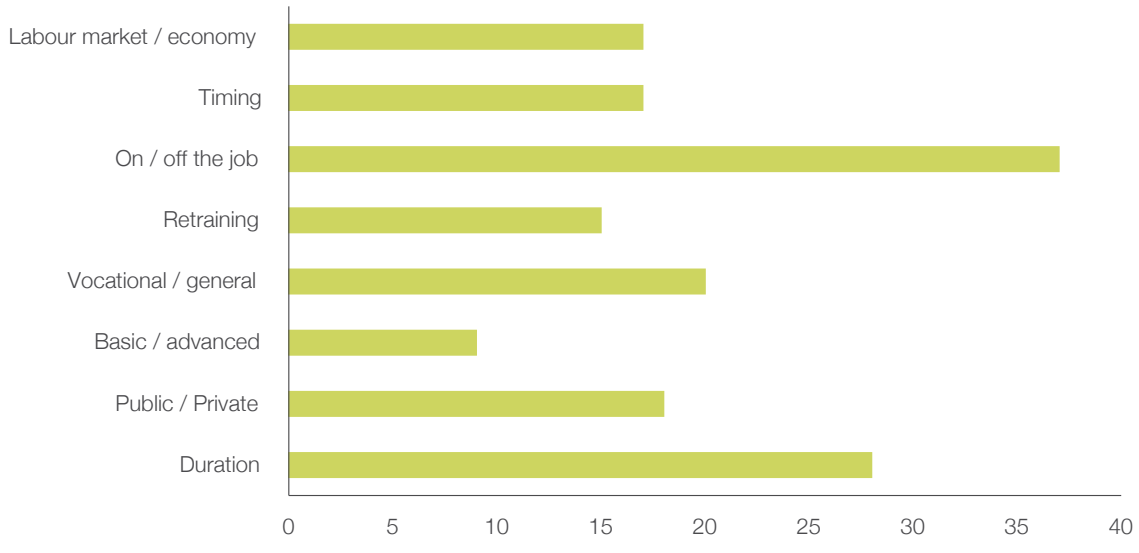
To identify what works, each evidence review sifts and assesses the evidence to find evaluations which are robust and clearly identify policy impact. We do this using a 5 stage process:

Figure 1: Methodology



This particular review considers whether there is any evidence of a link between specific programme features and the impact of training on labour market outcomes. Figure 2 provides a summary of the number of evaluations that look at different programme features.

Figure 2: Number of shortlisted evaluations by programme feature and context



Findings

What the evidence shows

- Training has a positive impact on participants' employment or earnings in around half of the evaluations reviewed.
- Shorter programmes (below six months, and probably below four months) are more effective for less formal training activity. Longer programmes generate employment gains when the content is skill-intensive.
- **In-firm / on the job training programmes tend to outperform classroom-based training programmes. Employer co-design and activities that closely mirror actual jobs appear to be key design elements.**
- The state of the economy is not a major factor in the performance of training programmes; **programme design features appear to be more important than macroeconomic factors.**

Where the evidence is inconclusive

- **Comparing different skill content training – such as 'basic' versus 'advanced' interventions – is extremely difficult**: finding suitable comparators (i.e. policies that target similar groups using different types of training) is challenging, and skill content usually reflects real participant differences.
- **Training programmes that respond to structural shocks in the local economy are usually highly tailored to a given local context**. This means that pulling out generalisable findings on impact is difficult.
- **It is hard to reach any strong conclusions on private-led versus public-led delivery on the basis of the (limited) available evidence.**

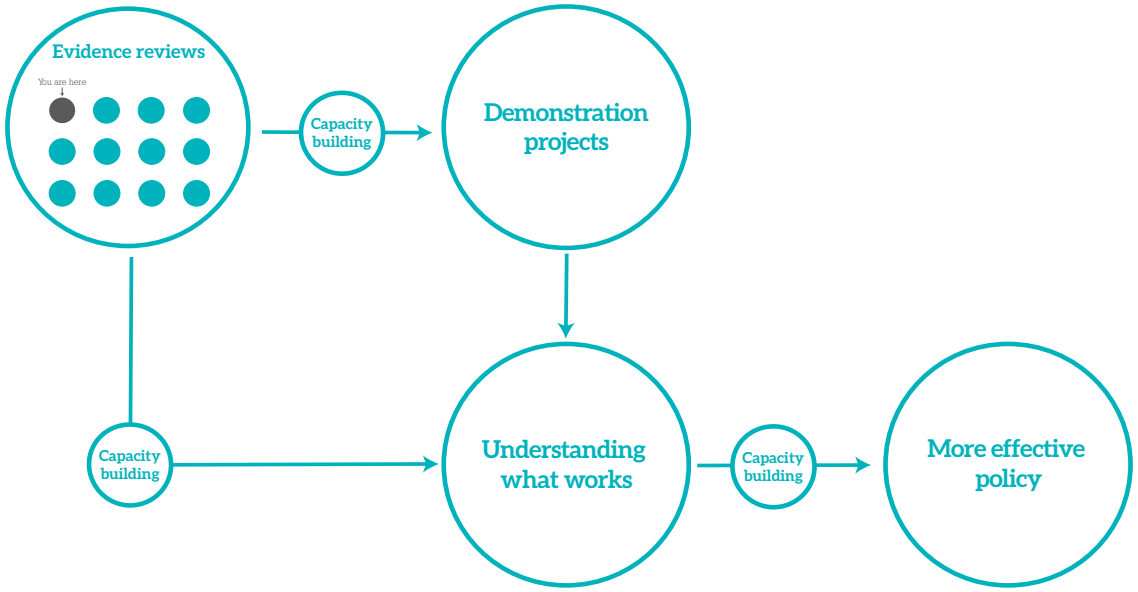
Where there is a lack of evidence

- **We have found little evidence which provides robust, consistent insight into the relative value for money of different approaches**. Most assessments of 'cost per outcome' fail to provide a control group for comparison.
- **We found no evidence that would suggest local delivery is more or less effective than national delivery.**

How to use these reviews

The Centre's reviews consider a specific type of evidence – impact evaluation – that seeks to understand the causal effect of policy interventions and to establish their cost-effectiveness. In the longer term, the Centre will produce a range of evidence reviews that will help local decision-makers decide the broad policy areas on which to spend limited resources. Figure 3 illustrates how the reviews relate to the other work streams of the Centre.

Figure 3: What Works Centre work programme



Supporting and complementing local knowledge

The evidence review sets out a number of 'Best Bets' which outline what tends to work in the employment training policy field based on the best available impact evaluations.

The 'Best Bets' do not generally address the specifics of 'what works where' or 'what will work for a particular individual'. In some cases evaluations do break out results by area type or different groups. But even when they do, detailed local knowledge and context remain crucial.

Any policy intervention focused on employment training will need to be tailored and targeted. And an accurate diagnosis of the specific local employment and skills challenges this policy seeks to address needs to be the first step to understanding how the overall evidence applies in any given situation.

Providing general guidance on what works

The 'Best Bets' highlight the common characteristics of employment training programmes and projects that have positive effects.

Whilst the 'Best Bets' cannot provide definitive guidance as to what will or won't work in any specific context, they do provide useful overall guidance to policy-makers to use when designing an employment training programme. They also raise a note of caution for policy-makers if they decide to try out a programme which has not worked so well elsewhere.

Providing detailed evidence on specific programmes

The 71 evaluations offer a rich source of material for policy-makers to use in designing specific employment training policies. In particular the evaluations will be of use to policy-makers at two key stages in the policy design process: determining the policy options, and then selecting the preferred option.

For both stages, the policy-makers should ensure that their understanding of their specific situation and the different policy options available is as detailed and comprehensive as possible.

Filling the Evidence Gaps

This review has not found answers to some of the questions which will be foremost in policy-makers' minds.

These gaps highlight the need for improved evaluation and greater experimentation, specifically experiments that focus on:

- identifying how different elements of employment training programme design contribute to better or worse outcomes; and,
- the value for money of different approaches.

This requires evaluation to be embedded in policy design, and thinking differently about the policy cycle as a whole.

03

Introduction

Why are we interested in training policy as a tool of local economic growth? The short answer is that we hope training improves skills, because higher skills are linked to better economic performance.

For individuals, higher skills are associated with better labour market outcomes. For local areas, there is a clear link between skills and economic growth and labour market outcomes. That connection holds nationally, across countries and at a local level for towns and cities (see figure 1).[1]

Figure 1: The positive link between UK city skill levels and economic growth

Source: Swinney, Larkin and Webber 2010. Firm Intentions: Cities, private sector jobs and the coalition. London: Centre for Cities

1 For cross-country evidence, see Barro (2001) or Cohen and Soto (2007). For recent reviews of the local-level evidence see Duranton and Puga (2014), Henderson (2007) or Moretti (2004b).

Of course faster-growing towns and cities may be able to invest more in skills and may also attract people with greater knowledge and experience, thus reversing the direction of causality between economic performance and skills; but recent studies have established a causal connection *from* the local skills base *to* local earnings and employment growth.[2]

For all of these reasons, policy-makers in the UK and elsewhere are keen to raise skill levels. Policy-makers have been particularly keen on retraining or 'upskilling' low skilled workers and the unemployed, especially those in long term unemployment. In the UK these training policies have often been delivered as part of active labour market programmes such as the New Deal and the Work Programme.[3]

Economists use the idea of 'human capital' to talk about the stock of competencies, knowledge, social and personal attributes ('skills') that help people produce economic output. We can usefully break human capital down into three main components: education, soft skills and experience.[4] Formal education might happen in the classroom or on the job, and cover basic skills such as literacy and numeracy; academic knowledge (of a given subject) or vocational skills (for a given job or trade). Soft skills such as team-working or communication tend to be acquired along the way. Experience helps us develop formal understanding and soft skills, but also gives us broader types of knowledge about a given firm or industry.

Our review of adult employment training programmes has turned up a range of policies designed to enhance these different components of human capital – from formal qualifications to courses on soft skills and advice on job search; from classroom teaching to on-the-job internships and in-firm training; from short-term intensive courses of a few weeks, to long term retraining leading to a degree, and lasting two years or more.[5]

In most countries, employment training programmes are devised and delivered through national governments but often have some degree of local flexibility (for instance, case workers or personal advisers may hold individual 'budgets' on behalf of programme participants). And in some cases – especially the US – states and cities have designed their own devolved employment training interventions.

Given the level of policy attention and spending, it is important to understand 'what works'. To understand the impact of the policy, we need answers to four questions:

- Who are we aiming to help?
- What does the policy aim to achieve?
- How is it delivered?
- Does it work?

To understand whether policy is cost-effective, we then need to compare the benefits of any policy impacts to the cost of achieving those impacts. In the area of employment training there is reasonable evidence on the impact of policy, although it is hard to compare the magnitude of effects across different studies (we focus on whether effects are positive, zero or negative). There is very little evidence that compares benefits to costs and thus allows us to assess cost-effectiveness.

2 Some of the key studies at local level are Glaeser and Maré (2001), Glaeser and Saiz (2004), Moretti (2004a), Moretti (2004c) and Shapiro (2006). For cross-country evidence see Hanushek and Woessmann (2012).
3 OECD (2009), Payne and Keep (2011).
4 Becker (1962).
5 This report does not cover evidence on apprenticeships.

04

Impact evaluation

Governments around the world increasingly have strong systems to monitor policy inputs (such as spending on a training programme) and outputs (such as the number of people who have gone through the programme). However, they are less good at identifying policy *outcomes* (such as the effect of a training programme on employment or wages). In particular, many government sponsored evaluations that look at outcomes do not use credible strategies to assess the **causal impact** of policy interventions.

By causal impact, the evaluation literature means an estimate of the difference that can be expected between the outcome for individuals 'treated' in a programme, and the average outcome they would have experienced without it. Pinning down causality is a crucially important part of impact evaluation. **Estimates of the benefits of a programme are of limited use to policy-makers unless those benefits can be attributed, with a reasonable degree of certainty, to that programme.**

The credibility with which evaluations establish causality is the criterion on which this review assesses the literature.

Using Counterfactuals

Establishing causality requires the construction of a valid counterfactual – i.e. what would have happened to programme participants had they not been treated under the programme. That outcome is fundamentally unobservable, so researchers spend a great deal of time trying to rebuild it. The way in which this counterfactual is (re)constructed is the key element of impact evaluation design.

A standard approach is to create a counterfactual group of similar individuals not participating in the programme being evaluated. Changes in outcomes can then be compared between the 'treatment group' (those affected by the policy) and the 'control group' (similar individuals not exposed to the policy).

A key issue in creating the counterfactual group is dealing with the 'selection into treatment' problem. Selection into treatment occurs when individuals participating in the programme differ from those who do not participate in the programme.

An example of this problem in training programmes is skimming, where providers choose participants with the greatest chance of getting a job. If this happens, estimates of policy impact may be biased upwards because we incorrectly attribute better labour market outcomes to the policy, rather than to the fact that the participants would have done better even without the programme.

Selection problems may also lead to downward bias. For example, people may participate to extend benefit entitlement and such people may be less likely to get a job independent of any training they receive. These factors are often unobservable to researchers.

So the challenge for good programme evaluation is to deal with these issues, and to demonstrate that the control group is plausible. If the construction of plausible counterfactuals is central to good policy evaluation, then the crucial question becomes: **how do we design counterfactuals?** Box 1 provides some examples.

Box 1: Impact evaluation techniques

One way to identify causal impacts of a programme is to randomly assign participants to treatment and control groups. For researchers, such Randomised Control Trials (RCTs) are often considered the 'gold standard' of evaluation. Properly implemented, randomisation ensures that treatment and control groups are comparable both in terms of observed and unobserved attributes, thus identifying the causal impact of policy. However, implementation of these 'real world' experiments is challenging and can be problematic. RCTs may not always be feasible for local economic growth policies – for example, policy-makers may be unwilling to randomise.6 And small-scale trials may have limited wider applicability.

Where randomised control trials are not an option, 'quasi-experimental' approaches of randomisation can help. These strategies can deal with selection on unobservables, by (for example) exploiting institutional rules and processes that result in some people quasi-randomly receiving treatment.

Even using these strategies, though, the treatment and control groups may not be fully comparable in terms of observables. Statistical techniques such as Ordinary Least Squares (OLS) and matching can be used to address this problem.

Note that higher quality impact evaluation first uses identification strategies to construct a control group and deal with selection on unobservables. Then it tries to control for remaining differences in observable characteristics. It is the combination that is particularly powerful: OLS or matching alone raise concerns about the extent to which unobservable characteristics determine both treatment and outcomes and thus bias the evaluation.

6 Gibbons, Nathan and Overman (2014).

Evidence included in the review

We include any evaluation that compares outcomes for people receiving treatment (the treated group) after an intervention with outcomes in the treated group before the intervention, relative to a comparison group used to provide a counterfactual of what would have happened to these outcomes in the absence of treatment.

This means we look at evaluations that do a reasonable job of estimating the impact of treatment using either randomised control trials, quasi-random variation or statistical techniques (such as OLS and matching) that help make treatment and control groups comparable. We view these evaluations as providing credible impact evaluation in the sense that they identify effects which can be attributed, with a reasonable degree of certainty, to the implementation of the programme in question.

Evidence excluded from the review

We exclude evaluations that provide a simple before and after comparison only for those receiving the treatment because we cannot be reasonably sure that changes for the treated group can be attributed to the effect of the programme.

We also exclude case studies or evaluations that focus on process (how the policy is implemented) rather than impact (what was the effect of the policy). Such studies have a role to play in helping formulate better policy but they are not the focus of our evidence reviews.

05

Methodology

To identify robust evaluation evidence on the causal impact of training programmes, we conducted a systematic review of the evidence from the UK and across the world. Our reviews followed a five-stage process: scope, search, sift, score and synthesise.

Stage 1: Scope of Review

Working with our User Panel and a member of our Academic Panel, we agreed the review question, key terms and inclusion criteria. We also used existing literature reviews and meta-analyses to inform our thinking.

Stage 2: Searching for Evaluations

We searched for evaluation evidence across a wide range of sources, from peer-reviewed academic research to government evaluations and think tank reports. Specifically, we looked at academic databases (such as EconLit, Web of Science and Google Scholar), specialist research institutes (such as CEPR and IZA), UK central and local government departments, and work done by think tanks (such as the OECD, ILO, IPPR and Policy Exchange.) We also issued a call for evidence via our mailing list and social media. **This search found close to 1,000 books, articles and reports.** Appendix B provides a full list of sources and search terms.

Stage 3: Sifting Evaluations

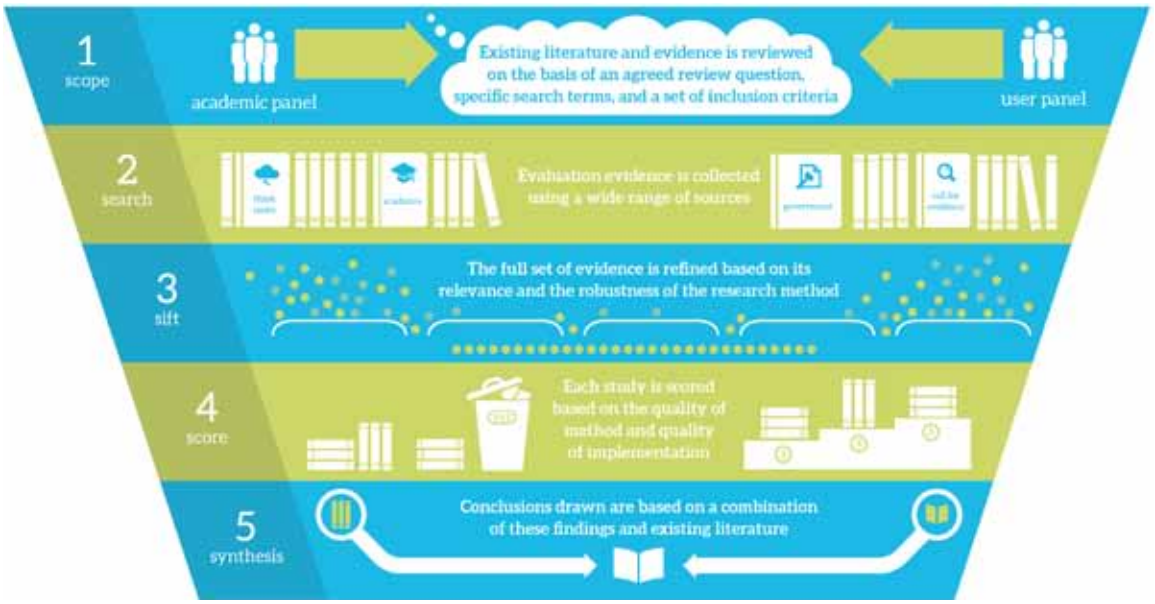
We screened our long-list on relevance, geography, language and methods, keeping impact evaluations from the UK and other OECD countries, with no time restrictions on when the evaluation was done. We focussed on English-language studies, but would consider key evidence if it was in other languages. We then screened the remaining evaluations on the robustness of their research methods, keeping only the more robust impact evaluations. We used the Scientific Maryland Scale (SMS) to do this.[7] The SMS is a five-point scale ranging from 1, for evaluations based on simple cross sectional correlations, to 5 for randomised control trials (see Box 2). We shortlisted all those impact evaluations that could potentially score 3 or above on the SMS. Find examples of employment training evaluations that score 3, 4 and 5 of the SMS scale on whatworksgrowth.org.

7 Sherman, Gottfredson, MacKenzie, Eck,Reuter, and Bushway (1998).

Stage 4: Scoring Evaluations

We conducted a full appraisal of each evaluation on the shortlist, collecting key results and using the SMS to give a final score for evaluations that reflected both the quality of methods chosen and quality of implementation (which can be lower than claimed by some authors). Scoring and shortlisting decisions were cross-checked with the academic panel member and the core team at LSE. The final list of included studies and their reference numbers (used in the rest of this report) can be found in Appendix A.

Figure 2: Methodology



Stage 5: Synthesising Evaluations

We drew together our findings, combining material from our evaluations and the existing literature.

Box 2: Our robustness scores (based on adjusted Maryland Scientific Methods Scale)

Level 1: Either (a) a cross-sectional comparison of treated groups with untreated groups, or (b) a before-and-after comparison of treated group, without an untreated comparison group. No use of control variables in statistical analysis to adjust for differences between treated and untreated groups or periods.

Level 2: Use of adequate control variables and either (a) a cross-sectional comparison of treated groups with untreated groups, or (b) a before-and-after comparison of treated group, without an untreated comparison group. In (a), control variables or matching techniques used to account for cross-sectional differences between treated and controls groups. In (b), control variables are used to account for before-and-after changes in macro level factors.

Level 3: Comparison of outcomes in treated group after an intervention, with outcomes in the treated group before the intervention, and a comparison group used to provide a counterfactual (e.g. difference in difference). Justification given to choice of comparator group that is argued to be similar to the treatment group. Evidence presented on comparability of treatment and control groups. Techniques such as regression and (propensity score) matching may be used to adjust for difference between treated and untreated groups, but there are likely to be important unobserved differences remaining.

Level 4: Quasi-randomness in treatment is exploited, so that it can be credibly held that treatment and control groups differ only in their exposure to the random allocation of treatment. This often entails the use of an instrument or discontinuity in treatment, the suitability of which should be adequately demonstrated and defended.

Level 5: Reserved for research designs that involve explicit randomisation into treatment and control groups, with Randomised Control Trials (RCTs) providing the definitive example. Extensive evidence provided on comparability of treatment and control groups, showing no significant differences in terms of levels or trends. Control variables may be used to adjust for treatment and control group differences, but this adjustment should not have a large impact on the main results. Attention paid to problems of selective attrition from randomly assigned groups, which is shown to be of negligible importance. There should be limited or, ideally, no occurrence of 'contamination' of the control group with the treatment.

Note: These levels are based on but not identical to the original Maryland SMS. The levels here are generally a little stricter than the original scale to help to clearly separate levels 3, 4 and 5 which form the basis for our evidence reviews.

06

Definition

We define employment training programmes as:

Including:

- training targeted at people aged over 18
- short courses and day-release courses
- retraining initiatives (e.g. for the over 50s)
- training that is in some way publicly subsidised (either directly or indirectly); where private sector provision is contracted out from the public sector, or on payment-by-results basis

Excluding:

- fully corporate/commercial providers on the assumption that this part of the sector operates as a competitive market
- 'A Level' education (even where offered to adults; because this provision is primarily targeted at under-18s and is not usually publicly funded)
- academic and Higher Education (which is a topic in its own right)
- apprenticeships (because the primary relationship is with the employer, and there may or may not be public subsidy)

07

Findings

This section sets out the review's findings. We begin with a discussion of the evidence base then explore the overall pattern of results. After this we consider specific outcomes as well as programme features in more detail.

Quantity and quality of the evidence base

From an initial long-list of 1,000 studies, we found 71 impact evaluations which met our minimum standards.[8] This represents a relatively large evidence base compared to many other local economic growth policies; but a small base relative to that available for some other policy areas (e.g. medicine, aspects of international development, education and social policy).

Table 1 shows the distribution of the studies ranked according to the SMS. There are five randomised control trials,[9] two of which score 5 for implementation, two score 4 (one due to attrition issues, the other due to poor correction for self-selection bias) and the fifth scores 3 (due to the treatment being imprecise and coverage imperfect).

For the studies which do randomise the allocation into either treatment or control group is done after individuals have already volunteered, or been selected, for training. This means these studies tell us whether the training is effective for the type of people likely to volunteer or be selected into training. It does not, therefore, tell us whether extending training to individuals who differ from those volunteered or selected will reap similar benefits.

10 evaluations score 4 on the Maryland Scale (including the two RTCs discussed above) and eight evaluations use credible quasi-random sources of variation, including one that conducts new analysis based on previous randomised control trial,[10] four that use a timing-of-events approach,[11] one that exploits a discontinuity in the policy eligibility criteria,[12] one that uses a factor control

8 Many of these 1,000 studies provided case studies or process evaluations which are not the focus of our review. See Methodology section for further discussion. We initially shortlisted 77 evaluations. However, six evaluations (97, 103, 131, 148, 224 and 241) were given level 2 for implementation even though they used a method that could have scored a level 3. These studies were downgraded if they failed to sufficiently control for observed characteristics such as by omitting factors that are considered important in the literature.

9 Studies 79, 134, 152, 209, 223.

10 Study 140.

11 Studies 231, 234, 236 and 252.

12 Study 239.

function to model unobservables[13] and one that uses Heckman's two-step approach with credible instrumental variables.[14]

59 evaluations (the majority) score 3 on the Maryland scale. This includes some studies which utilize less robust implementations of RCT and quasi-random techniques, though most use variations on OLS or matching techniques. By and large, this means that we can be confident that the study has done a good job of controlling for all observable or recorded factors (for example: race, gender, previous education, strength of the labour market) which might explain better labour market outcomes following training. However for these studies, it is likely that 'unobservable' characteristics such as motivation or aptitude may still be affecting the results raising concerns that the evaluation incorrectly attributes employment or wage effects to the programme rather than to these individual characteristics.

The sections below set out the findings with respect of some key programme characteristics, based on the evidence presented in these 71 evaluations.

Table 1: Ranking studies by quality of implementation

Scientific Maryland Scale score	Number by implementation
5	2
4	10
3	59
Total	**71**

Type and Focus of Training Programmes

Most of the shortlisted evaluations considered training programmes as part of wider Active Labour Market Policies (ALMPs), whose overall aim is to reduce welfare dependency with a broad package of training and employment measures.

ALMPs tend to involve a 'two-pronged' approach, targeting both increased employment opportunities[15] and increased wage potential[16] (since many countries have some element of benefit paid to those in work but on a low wage, and therefore an interest in increasing wages to reduce the social security bill).

Other programmes were more explicitly focused on addressing skill mismatches in the economy[17] and increasing human capital.[18] In the case of the former, most of the programmes assessed aimed to address large-scale structural unemployment (for example, in the former East Germany after reunification, or in cases where large local industries or employers have been lost).

As expected, some programmes were focused on achieving 'quick-wins', i.e. short term statistical targets (such as increased employment), whilst others looked to secure long term outcomes, such as more sustainable, secure employment opportunities.[19]

13 Study 127.
14 Study 219.
15 Studies 127,140, 152.
16 Studies 140 and 123.
17 Studies 79, 145, 111, 149.
18 Studies 96 and116.
19 Study 138.

Findings

This section of the report considers the impact of employment training programmes on specific outcomes, before examining whether there is any evidence of a link between specific programme features and programme success.

The comparison of the effects of programme features is not straightforward because even if we find (say) that private-led delivery seems more effective in terms of delivering positive outcomes a number of other 'confounding factors' may be in play which could explain that relationship. It might be, for example, that the public-led programmes focused only on the hard to reach.

To help deal with this issue we use the following approach. First, we look for any general pattern / correlation between the feature and the outcome we're interested in (say delivery type and employment). Second, we look for studies that make explicit comparisons. We group these into two types. 'Type 1' comparisons compare different programmes (say, one public-led and one private-led) but for broadly similar types of people. 'Type 2' comparisons look at different strands of the same programme, where there is (say) one private-led and one public-led element.[20] Both Type 1 and Type 2 comparisons help control for confounding factors, but Type 2 does this more cleanly and we put more weight on these results. Where appropriate we also use SMS scores to quality-weight individual studies when reaching our overall conclusions.

Individual Outcomes: Employment

Training has a positive impact on participants' employment in nearly half the evaluations reviewed.

Table 2: Employment

Result	No of studies	Study reference numbers
Positive	29	84, 87, 94, 109, 110, 111, 116, 132, 140, 145, 151, 218, 219, 220, 222, 225, 227, 232, 233, 234, 240, 242, 244, 246, 247, 249, 251, 254, 256.
Zero	12	115, 123, 130, 147, 150, 152, 223, 226, 231, 237, 238, 245.
Negative	3	120, 128, 228.
Mixed	19	79, 85, 88, 90, 99, 117, 126, 127, 146, 149, 221, 229, 230, 235, 236, 239, 243, 250, 252.

Unsurprisingly, given one of the primarily rationales of employment training programmes is to increase the employment levels of participants, the majority of studies, 63 out of 71, considered impacts on employment (see Table 2). Of these nearly half, 29 out of 63, find positive programme impacts, and a further 19 find mixed results (some interventions worked better than others; some groups gained more than others; different evaluation approaches yielded different results i.e. a mixture of positive, negative and zero effects).

12 studies found that training interventions did not increase employment, (i.e. they found no statistically significant evidence of positive outcomes) and a further three found significant negative impacts. Note that negative impacts do not necessarily mean that training makes people less employable in the longer term: a number of studies which find mixed or negative effects

20 In a few cases this distinction is not clear; we consider these separately.

attribute this to the so-called 'lock-in' effect whereby those who are participating in training are usually not actively job-seeking, or spending less time doing so than their counterparts who are not in training.

Of the 29 studies which find positive programme impacts, a small number were higher quality (SMS 4) studies using quasi-random sources of variation in their methodologies.[21] We consider the results of these in a little more detail. Study 140 examined the impacts of the Greater Avenues for Independence (GAIN), the state of California's official Job Opportunities and Basic Skills Training programme after 1989, finding positive impacts across all programmes. Initially, programmes that focussed on job searches achieved better results, but those that worked to improve basic skills became more successful in the long term. Study 219 considered both classroom-based education and 'on the job' training programmes, finding positive outcomes for both, though it noted that 'on the job' was the more effective of the two. Additionally, it noted that for private sector employees who had successfully found a job after one year, the probability of remaining in that job two years after is 80.2%. In the case of study 234 for Norway, a 'lock-in' effect is noted during participation in active labour market programmes, but following this the impacts on employment are positive. The paper finds that the programmes reduce the total number of participants' unemployment months by 6.42%.

A further 5 high quality (SMS 4 studies)[22] find mixed results with respect to employment, with variation in results relating to a number of factors, including:

- Specification – e.g. Study 239, which considers various UK labour market training initiatives aimed at young people, finds some positive coefficients, but variation in effects across different specifications (with the preferred IV method finding no statistically significant results);
- Time period – e.g. Study 252, which looks at the Swedish AMU employment training programme, identifies a significant improvement in terms of transition to employment, but only in the short term;
- Beneficiary – e.g. Study 79, which considers the US JTPA Title II-A training programmes, finds positive impacts for adults but not for youths;
- Type of effect – e.g. Study 236 finds that for state-sponsored training programmes in France, there is no acceleration in exit rate from unemployment, but programmes do have a significant and positive effect on the duration of subsequent employment spells.

Of the 12 studies which find no statistically significant results with respect to employment, one is a high quality (SMS 5) RCT (Study 223). This study examines the effect of an adult education voucher scheme in Switzerland, and finds no statistically significant average effects of the voucher program on earnings and employment probabilities 1 year after treatment.

Interestingly, of the three studies which find programmes to have negative effects on employment, two focus on labour market interventions in East Germany.[23] For example, study 228 concluded that subsidised apprentices need on average 5 months for entering their first employment, approximately two months longer than comparable graduates of regular training. However, given the unique factors impacting the labour market in East Germany following re-unification, these studies should be treated with caution (see discussion in Weak v Strong Labour Markets).

21 Studies 140, 219 and 234.
22 Studies 127, 236, 239, 252 (SMS 4) and Study 79 (an RCT scored at SMS 4).
23 Studies 120 and 228.

Individual Outcomes: Wages

Training has a positive impact on participants' wages in just over half the evaluations reviewed, though in several cases it is noted that wages do not increase in the immediate term but only after a number of years.

Table 3: Wages

Result	No of studies	Study reference numbers
Positive	11	80, 89, 125, 137, 138, 149, 209, 249, 250, 251, 256
Zero	5	120, 152, 223, 237, 245
Negative	3	99, 134, 238
Mixed	2	96, 128

A smaller number of studies, 21 out of 71, consider the impact of employment training programmes on the wages of participants. In over half of these evaluations, 11 out of 21, there were positive impacts, and mixed impacts in a further two (see Table 3).

In cases where programmes are successful in increasing wages, these gains are often substantial. For example, Study 209, an SMS 5 RCT which examines the impact of the US federal Jobs Corps programme, found that after 4 years, participants on average earned $1,150 more than similar non-participants (a 12% increase). This study also illustrates a finding which is mirrored in many others; wage increases are often not immediate, instead taking a number of years to materialize. This may explain the results of Study 223, a high quality RCT and one of five studies which found that employment training had no significant effect on wages but only looked at the impact on wages one year after participation.

While only three of the 21 studies found that employment training was harmful to wages, it is notable that one of these is a higher quality (SMS 4) study that used quasi-random techniques.[24] However, while the study concluded that Jobs Corps programme participants lagged behind their control counterparts in the labour market by as much as 8 months, translating into as much as a 5.8% wage disadvantage even four years, this is attributed to the 'lock in' effect described previously, whereby participants are disadvantaged by lost labour market experience during time in the programme (when they are not actively job seeking). Furthermore, the authors emphasise the primary focus of Jobs Corps on raising human capital, rather than solely encouraging work.

Programme Design Elements

Local vs national delivery

We found no evidence that would suggest one level of delivery is more effective than others.

The 71 studies involve delivery models at various scales, from highly centralised programmes (for example, in France and Sweden) to the more devolved systems prevalent in the US (where States have substantial freedom to design and deliver policy). The majority of programmes covered in the

evaluation literature are national level or state level programmes, with few examples of significant local flexibility.[25]

Surprisingly, none of the evaluations in this review look directly at the question of whether programmes are more successful when they are locally, regionally or centrally administered. When we classified evaluations according to the level at which the programme is delivered, we found no evidence that would suggest one level of delivery is more effective than others.

Public- vs private-led delivery

Overall, it is difficult to reach any strong conclusions on the effectiveness of private-led versus public led-delivery. Results appear to be more mixed for public-led than private-led but this may be explained by other characteristics of the programme.

While the review focuses on evaluations of publicly-funded programmes, the management and delivery of such programmes is often divided and shared between public and private sector organisations.

A frequently-used hybrid structure involves national public funding, public sector agencies providing overall programme management, with specific courses and sub-programmes contracted out to a mixture of other public and private delivery groups. Different agencies might also be involved at various delivery stages, for example state-run employment offices providing initial client-facing services with these customers then being referred on to private companies once training needs are established.

Based on the evaluations we can classify programme delivery models as hybrid, public-led, private-led, or not stated. As Table 4 shows, a hybrid delivery model is the most common across OECD countries. We can see that very few programmes use an exclusively or predominantly private sector-led delivery model.

Table 4: Delivery models

Delivery model	No. of studies	Study reference numbers
Public-led	9	109, 120, 123, 126, 130, 226, 230, 231, 246
Private-led	6	84, 94, 134, 137, 147, 256
Hybrid	26	89, 96, 99, 132, 209, 218 , 219, 220, 221, 225, 228, 229, 233, 234, 235, 236, 237, 238, 240, 242, 243, 244, 245, 249, 251, 252
Not stated	30	79, 80, 85, 87, 88, 90, 110, 111, 115, 116, 117, 125, 127, 128, 138, 140, 145, 146, 147, 149, 150, 151, 152, 222, 223, 227, 232, 239, 247, 250
Compare public vs private-led	3	219, 225, 242

25 The only evaluation we have found which covers a local programme and is of sufficient methodological quality is Study 94 which looks at a programme delivered by New York City.

The six programmes delivered only by the private sector predominantly record positive effects on participants' labour market outcomes (employment and wages), the exception being Study 147, which found no statistically significant result.

Nine evaluations looked at programmes delivered largely or solely through public sector bodies. These programmes showed more varied results in terms of the overall effect on participants' employment and wages. Two of these studies are from East Germany which, as we note elsewhere, is a special case.[26] An evaluation of a Danish programme highlights poor policy design as a possible explanation, with participants being poorly trained for the set of available jobs (resulting in both lower employment probabilities and more unstable work patterns).[27]

Just three studies compared private and public delivery models, with the former performing better than the latter in all of these studies. However, none make direct, like for like (Type 2) comparisons. The more robust, a Danish study, found on-the-job training to be more effective when delivered through the private sector as opposed to public sector. However, participants in private sector strands had substantially more work experience, suggesting some skimming; at the same time, private sector programmes had closer links to the 'real' labour market than their public sector counterparts.[28] Similar outcomes are found in the case of a Danish study, which considers two training programmes aimed at increasing IT competence: an active labour market programme run by Swedish industry (private); and a traditional active labour market programme for the unemployed run by the Swedish National Labour Market Board (public).[29] In this case, the private sector led programme increased participants' employment by 20% versus the public sector programme. The final evaluation compared on the job training (delivered by the private sector) with community jobs and courses (delivered by the public sector) in France. As we discuss below, in-firm training delivers significantly better labour market effects for programme participants.[30]

26 studies looked at programmes which involved both public and private components. Findings for these programmes are, again, highly varied.

Weak v Strong Labour Markets

Overall, it is difficult to reach any strong conclusions on the extent to which the effectiveness of training programmes is affected by the strength of the labour market. Other programme design features appear to dominate macro factors.

Many of the programmes evaluated have taken place in response to weak labour markets: Table 5 summarises the macro conditions at the time of study.

Several evaluations suggest that economic cycles affect the outcomes of training programmes. How strong is the evidence for this link?

The overall link between macro conditions and programme performance is weaker. 14 evaluations acknowledged the presence of a weak labour market. In a further 11 evaluations, the labour market is in a state of transition. There is no clear pattern of the performance of the programmes covered by these evaluations. At least one existing meta-analysis suggests that once programme features are taken into account, there is little systematic link between the macro environment and the effectiveness of training.[31] Other programme features, then, seem more important to performance.

26 Studies 120, 130.
27 Study 231.
28 Study 219.
29 Study 242.
30 Study 225.
31 Kluve 2010 covers 139 European training and welfare to work programmes.

Table 5: Labour market context to the studies

Labour market	No. of studies	Study reference numbers
Strong labour market	2	137, 219,
Weak labour market	14	96, 109, 110, 127, 130, 147, 229, 237, 238, 240, 243, 244, 245, 246
Changing labour market over evaluation period	11	99, 111, 120, 125, 126, 128, 228, 234, 249, 252, 254
Multiple areas w/ differential labour markets	12	89, 116, 117, 138, 140, 146, 149, 151, 218, 220, 226, 235
Not stated	32	79, 80, 74, 85, 87, 88, 90, 94, 115, 123, 132, 134, 145, 150, 152, 209, 221, 222, 223, 225, 227, 230, 231, 232, 233, 236, 239, 242, 247, 250, 251, 256
Studies with a focus on East Germany	10	116, 120, 130, 146, 147, 149, 218, 221, 228, 245

However, in a few cases there appears to be a clear link. In a Swedish evaluation, for instance, programmes that commenced during a weak phase of the business cycle, in the early 1990s, were less successful in reducing unemployment than those that commenced in the subsequent years[32], whilst a Norwegian study also suggested, in response to evidence of heterogeneity of treatment effects in terms of stages of the business cycle, that training effects are pro-cyclical.[33] One evaluation even noted that, due to fluctuations in the labour market that occurred during the study, the results for the treated group were no longer comparable with those in the control group. In this case, programme participants re-entered the labour market at a time when the market was weaker than for the control group and thus suffered outcomes that were significantly worse.[34]

A particular feature of our shortlisted evaluations which should be noted is the presence of 10 studies that considered the impact of training programmes in Germany at the time of reunification, with a particular focus on the former East German regions. While many of the evaluations themselves were carried out robustly, lessons from these evaluations may not be fully transferable as a result of the unique economic restructuring which the region was undergoing at the time. East Germany was said to be suffering from particularly acute skill mismatches because of the poor judgements made about future labour market requirements[35], while in some cases the training programmes themselves were deemed to be poor quality in the immediate period after reunification.[36] It may therefore be that insignificant or negative findings across several German studies may reflect the poor design and implementation of those programmes rather than any inherent characteristic of training.

32 Study 128.
33 Study 234.
34 Study 99.
35 Study 149.
36 Study 116.

Shorter v longer programmes

Overall, short programmes (below six months, and probably below four months) are more effective for less formal training activity. Longer programmes generate employment gains when the content is skill-intensive, but benefits to the individual typically play out over a longer time frame.

The training programmes reviewed vary considerably in length, ranging from 10 days up to three years. Programmes tend to be classified as 'short' and 'long' on their own terms, and programme lengths are not always explicitly stated in the literature. Many programmes do not have a fixed length, incorporating a mixture of different training elements that vary in duration. Programmes may also be tailored to the needs of individual people. In such cases, evaluations generally acknowledge this by providing a range for programme length which may vary significantly.[37]

For the purposes of our review, we classified short programmes as those stated to be six months or less, with long term programmes over six months. Table 6 presents the breakdown across types for those evaluations which reported a fixed programme length.

Table 6: Programme lengths

Programme length	No. of studies	Study reference numbers
Short	10	84, 90, 94, 96, 146, 235, 238, 239, 244, 254
Long	6	99, 109, 228, 229, 251, 256
Both	31	80, 85, 87, 88, 89, 111, 116, 117, 120, 126, 128, 130, 134, 140, 149, 209, 218, 219, 220, 221, 225, 230, 233, 234, 236, 237, 240, 242, 243, 246, 247, 257
Compare short vs long	10	85, 88, 149, 218, 220, 221, 236, 240, 243, 246
Not stated	24	79, 110, 115, 123, 125, 127, 132, 137, 138, 145, 147, 150, 151, 152, 222, 223, 226, 227, 231, 232, 245, 249, 250, 252

10 evaluations looked solely at short programmes. These generally found positive or mixed impacts on employment and wages. These programmes tended to be focused on upskilling, either partially or entirely classroom-based. One evaluation found negative coefficients, but questioned the reliability of the sample selection methods utilised and thus the reliability of the results in general.[38]

Six evaluations just looked at long programmes, with the majority identifying positive impacts. Notably, two programmes were found to have negative effects; in one case, graduates of subsidised training were found to be in slightly less qualitifed and worse paid jobs than participants in the control group;[39] in the second case though, it was acknowledged that

37 Study 85 provides a good example of this, programme length varying from 1 week up to 13 months.
38 Study 238.
39 Study 228.

programme effects might have been evaluated too early due to the severe recession that participants faced on exiting the training.[40]

Part of the problem around long term programmes, is an observed lock-in effect, whereby participants reduce their job-searching activity while undertaking training.[41] This means that the longer the training course, the longer the benefits take to materialise for individuals. Studies often find that the negative "lock-in" effect on employment and earnings is more prolonged for long programmes.

Some evaluations considered similar types of training but provide comparisons of the effectiveness of different programme lengths. That is, they seek to isolate the effect of programme length on effectiveness. For basic skills or interventions aimed at raising general employability – shorter programmes have a larger, stronger effect on participants' employment. [42] We found one study that explicitly looked at the link between different programme lengths and the effect on employment. According to this study, increasing training length up to around 150 days has a positive effect on outcomes, but there is no additional effect beyond 150 days. If this result is generalised, it suggests that the largest employment effects might be for courses that are short, but not too short.

Comparative results for more intensive programmes, that involve longer periods of training, are more mixed. The strongest comparative study, an evaluation of French public training programmes, found that longer courses lead to longer employment spells, and explain this in terms of improved job matching.[43] A less robust study for Ireland suggests that for more skill-intensive training, longer courses have better employment outcomes than shorter interventions.[44] However, two German studies find zero positive effects of long term retraining (for a vocational degree) versus shorter, non-degree programmes.[45]

One reason for these mixed results may be that long training programmes, especially those which result in formal qualifications, may only have detectable effects some years later – longer than the data points of many of these studies. [46] One recent review finds that longer evaluations (two years plus) of a training programme often find positive employment effects, shorter evaluations of the same programme are often negative. [47]

One German evaluation which was carried out over a longer period found evidence to suggest that long term programmes have more persistent, long term effects on employment levels after the initial locking-in period.[48] Another notable exception tracks the effects of West German programmes up to seven years post-treatment.[49] This suggests that retraining programmes, which run for two years or more and lead to a degree, have substantial positive employment effects over the seven year period that are larger than those from shorter training programmes (12 months or less).

40 Study 99.
41 Study 126.
42 Studies 218, 221 (Type 1); studies 85, 88, 149, 236 (Type 2). Studies 240 and 246 fall between Type 1 and Type 2.
43 Study 236.
44 Study 246.
45 Studies 218, 221.
46 In one case (study 149) negative effects of long term non-degree programmes was also found, but poor policy design explained a large part of this.
47 Card et al 2010. In our shortlist Study 218 indicates that this might have caused bias in their results.
48 Study 220.
49 Study 126.

Classroom v in-firm training

Overall, in-firm / on-the-job training programmes tend to outperform classroom-based training programmes. Employer co-design and activities that closely mirror actual jobs appear to be key design elements.

This suggests that hybrid programmes that combine off-the-job and on-the-job strands should ensure on-the-job elements follow these guidelines. Where participants are among the 'hardest to help', however, classroom training may pay dividends (e.g. covering basic skills).

Training can be provided in different contexts, and many programmes blend different elements of off-the-job' training (usually classroom based but sometimes practical or vocational) with 'in-firm' or 'on-the job' training (often employer based[50] and sometimes involving subsidized employment with a training element).

Table 7 shows the breakdown of programme types. As with programme lengths, this is one of the few areas where there are several comparisons of the same programme delivered in different contexts.

Table 7: Training programmes by site of learning

Site of learning	No. of studies	Study reference numbers
Classroom-focused	10	89, 90, 94, 96, 125, 134, 223, 231, 245, 250
In-work-focused	2	109, 228
Both	37	79, 80, 85, 87, 88, 111, 116, 120, 123, 126, 130, 132, 145, 146, 149, 209, 218, 219, 220, 221, 225, 228, 229, 232, 233, 234, 235, 237, 238, 240, 242, 243, 247, 249, 251, 254, 256
Compare classroom vs in work-focused	14	132, 145, 146, 218, 219, 221, 225, 233, 235, 237, 238, 240, 243, 251
Not stated	22	84, 99, 110, 115, 117, 127, 128, 137, 138, 140, 147, 150, 151, 152, 222, 226, 230, 236, 239, 244, 246, 252

Ten evaluations considered the effects of programmes which were undertaken entirely in a classroom environment. These programmes included training across a broad spectrum of different levels, from basic through to advanced. The majority were undertaken in the setting of a formal, educational institution and none had any element of on-the-job experience. Only one study found that classroom training had a positive impact on employment prospects, with a 15% increase in employment likelihood following programme participation,[51] whilst three conclude that such programmes have positive impacts on wages.[52]

In other cases, impacts are mixed, dependent on the statistical methods used[53] or the point at which the intervention is evaluated,[54] or often insignificant; in one case, the programmes had no statistically significant impacts upon employment prospects,[55] while a further two find similar

50 Study 240.
51 Study 94.
52 Studies 89, 125, 250.
53 Study 90.
54 Study 96.
55 Study 231.

outcomes for both employment and wages.[56] In one case, this was because the evaluation was carried out so shortly after the end of the programme,[57] another because of a poor uptake of the training vouchers under evaluation.[58] A further study attributed negative results to external, non-observable factors unrelated to the nature of training.[59]

Given the general programme design features discussed above, very few evaluations look at solely in-work programmes: we found only two of these.

Of those studies which look explicitly at the relative performance of on-the-job versus off-the-job training, the weight of evidence suggests that in-firm training, and specifically those training programmes which closely align with specific jobs and have high levels of employer contact, are most successful in raising the probability of employment. Of the 14 studies that compare programmes, five confirmed that in-work-focused training programmes have bigger employment effects; the bulk of these score highly on the SMS.

The superior performance of workplace-focused training may arise from a combination of factors. First, when employers engage directly with developing course content and with delivery, that content may be more relevant and useful in the marketplace. Second, in-firm training allows trainees to gain firm-specific skills and knowledge that is sufficiently valuable to increase their chance of employment in that firm at the end of the training period.[60] Third, in-firm training tends to have a shorter lock-in effect since participants are already likely to have forged close contact with firms directly,[61] thus reducing the time between programme end and employment.

There are a few variations to this general finding. Two East German studies found that classroom-based training was more effective than placing participants in 'simulated firms' – results which in effect confirm our main finding.[62] One Finnish study found that classroom based training outperforms practical in-firm training in programmes for unemployed young people.[63] Finally, in Germany a long term study of outcomes eight to nine years after the programme found classroom training to be the more effective.[64] There is a plausible explanation here that long term studies are required to estimate the full benefit of classroom training over on-the-job training.

Basic vs advanced skills training

Comparing fundamentally different types of training is difficult because finding suitable comparators (i.e. policies that target similar groups using different types of training) is challenging. Increasing understanding of what works for different groups needs to be a specific focus of policy design and evaluation.

Employment training is often adapted to cater for different groups, skill sets and situations. As in the case of short and long training programmes, many of the evaluations are quite loose in defining the type of training provided, with overlap between different types of training.

For the purposes of our review we classified 'basic training' as primarily aimed at those who have achieved only a basic education, but may also be carried out by skilled people in long term unemployment and even graduates with little experience in the labour market. We included

56 Studies 223, 245.
57 Study 245.
58 Study 223.
59 Study 231.
60 Study 249.
61 Study 146.
62 Study 221; Study 228.
63 Study 237.
64 Study 221.

programmes that featured job search assistance, interview training, basic IT training, literacy and numeracy, and readiness for work training in our definition of basic skills. 'Advanced training' is much more difficult to define. For the purposes of this review, we included on evaluations that solely discuss training designed to further develop a pre-existing skill set.[65]

Table 8: Skill content of training programmes

Skill content	No. of studies	Study reference numbers
Vocational	10	96, 132, 140, 147, 149, 220, 227, 228, 239, 256
General	1	94
Basic	3	94, 140, 243
Advanced	2	149, 239
Comparisons amongst various skill level characteristics (basic vs advanced or vocational vs general)	9	111, 128, 146, 221, 225, 243, 246, 251, 252

Training programmes are also frequently described as either vocational or general. These are again terms that are loosely defined, and individual evaluations usually do not give a great deal of information on course content and structure. For ease of comparison, we have again simplified for the purposes of our review. Broadly, we defined vocational training as aimed at entering a specific sector or profession, encompassing elements of classroom education and on-the-job experience, whereas general training might involve more general up-skilling which does not focus on a specific career.[66]

There is little evidence on the performance of basic and advanced programmes, as few evaluations treat them in isolation. However, the three evaluations that consider basic skills tend to find a positive impact on employment and salaries. According to one evaluation, soft skills such as CV writing and interview techniques can assist people into employment.[67] Of the two studies which specifically look at advanced skills content, both find mixed results. One finds overall positive effects of training on employment probability, but notes negative impacts for men who participate in long-term training or retraining.[68] The other, which uses an SMS 4 approach, finds some positive coefficients with respect to employment outcomes, but notes that the overall picture across different time periods/specifications is that the effect varies.[69]

Our review included 10 evaluations which looked specifically at vocational training. On balance, it appears that vocational training has largely positive effects on employment and wage outcomes. However, it must be acknowledged that results are mixed. For example, one evaluation suggests that outcomes depend upon the type of vocational training undertaken,[70] and that firms may build up different prejudices and biases to youths in vocational training depending on how closely linked the programme is to the firm. Meanwhile, another evaluation suggests that there is an

65 Study 149 provides a good example of a programme which provides 'Further' training.
66 For example, Study 111 discusses both a vocational programme that trains participants in a new occupation, and a Professional Skills and Techniques programme which provides more general skills training.
67 Study 94.
68 Study 149.
69 Study 239.
70 Study 228.

optimal time for participants to have been out of work before undertaking vocational training, depending also on age.[71]

All comparative evaluations for programme type can be described as 'indirect' (type 1) involving comparisons of independent programmes with different characteristics. This presents significant difficulties as the programmes being evaluated are fundamentally different in terms of target participants, goals and general characteristics.

On balance, the three comparative evaluations that compared basic and advance training in our review tended to find advanced training programmes to be more effective than basic training, though it is perhaps unsurprising that those with advanced skills are likely to experience better labour market outcomes than those with only basic skills. Comparisons of vocational and general training suggest general training tends to show more positive coefficients, though the results here again are mixed. These examples illustrate the problems of comparing fundamentally different types of training; finding suitable comparators is challenging.

Retraining in response to structural change

Training programmes that respond to structural shocks in the local economy are usually highly tailored to a given local context. This means that pulling out generalisable findings on impact is difficult.

In areas where the local labour market is strongly dependent on a single employer that relocates or goes out of business,[72] or where the labour market has been badly affected by economic restructuring, it is not unusual to see the public sector develop interventions aimed at retraining unemployed people to take up new jobs in different sectors entirely. We found 13 of these studies (see Table 9).

Table 9: Breakdown of 'structural adjustment' programmes

Programme	No. of studies	Study reference numbers
Retraining	2	110, 240
Upskilling	10	89, 94, 116, 117, 134, 146, 147, 209, 223, 252
Compare retraining vs upskilling	1	111

Characteristically, these 'structural adjustment' programmes tend to be of much longer duration than other interventions. This makes it harder to assess their impacts, because of lock-in effects and difficulties in picking up long term effects.

This is not to suggest that retraining programmes do not work. Three studies[73] (including some which discuss both retraining and upskilling programmes) find *some* positive impacts of retraining programmes, with workers between 30 and 50 and with a relatively long labour market history experiencing the greatest benefits. One of these studies looked at the longevity of earnings impacts and found the benefits could still be seen up to five years after training.[74] In one case, outcomes for employment vary by the length of retraining, with short term programmes found to be more effective, though wages consistently increase for participants of all retraining by between 100 and 200 EUR.[75]

71 Study 239.
72 Study 80.
73 Studies 80, 110, 149.
74 Study 80.
75 Study 149

08

Summary of findings

What the evidence shows

- Training has a positive impact on participants' employment or earnings in around half the evaluations reviewed.
- **Short programmes (below six months, and probably below four months) are more effective for less formal training activity**. **Longer programmes are more effective when the content is skill-intensive**, but benefits typically play out over a longer time frame.
- **In-firm / on the job training programmes tend to outperform classroom-based training programmes. Employer co-design and activities that closely mirror actual jobs appear to be key programme design elements.**

This suggests that hybrid programmes that combine off the job and on the job strands should ensure on the job elements follow these guidelines. Where participants are among the 'hardest to help', however, classroom training may pay dividends (e.g. covering basic skills).

- Programme design features appear to be more important for effectiveness than macroeconomic factors.

It is hard to reach any strong conclusions on whether the effectiveness of training programmes is affected by the strength or weakness of the labour market.

Where the evidence is inconclusive

Comparing different skill content training – such as 'basic' versus 'advanced' interventions – is extremely difficult; finding suitable comparators (i.e. policies that target similar groups using different types of training) is challenging, and skill content usually reflects real participant differences.

Training programmes that respond to structural shocks in the local economy are usually highly tailored to a given local context. This means that pulling out generalisable findings on impact is difficult.

It is hard to reach any strong conclusions on private-led versus public-led delivery on the basis of the (limited) available evidence.

The vast majority of the programmes reviewed use hybrid (public-private) delivery models. And the results appear to be more mixed for public-led than private-led but this may be explained by other characteristics of the programme.

Where there is a lack of evidence

- **We have found little evidence which provides robust, consistent insight into the relative value for money of different approaches.**

Most assessments of 'cost per outcome' fail to provide a control group for comparison. To address this we need to develop and incorporate stronger cost-benefit metrics in future evaluations.

- We found no evidence that would suggest local delivery is more or less effective than national delivery.

Surprisingly, we did not find impact evaluations which consider whether programmes are more successful when they are locally, regionally or centrally delivered. When we classify evaluations according to the level at which the programme is delivered, we find no clear pattern that would suggest one level of delivery is more effective than others.

This suggests that improved evaluation and greater local experimentation (on programme targeting, design and delivery) will be crucial to understanding whether greater local flexibility could improve policy effectiveness.

09

How to use this review

This review considers a specific type of evidence – **impact evaluation**. This type of evidence seeks to identify and understand the causal effect of policy interventions and to establish their cost-effectiveness. To put it another way they ask 'did the policy work' and 'did it represent good value for money'?

The focus on impact reflects the fact that we often do not know the answers to these and other basic questions that might reasonably be asked when designing a new policy. Being clearer about what is known will enable policy-makers to better design policies and undertake further evaluations to start filling the gaps in knowledge.

Supporting and complementing local knowledge

The evidence review sets out a number of 'Best Bets' which outline what tends to work in the employment training policy field based on the best available impact evaluations.

The 'Best Bets' do not generally address the specifics of 'what works where' or 'what will work for a particular individual'. In some cases evaluations *do* break out results by area type or different groups. But even when they do, detailed local knowledge and context remain crucial.

Reflecting this, the overall findings from the evaluations should be regarded as a complement, not a substitute, for local, on-the-ground knowledge.

Any policy intervention focused on employment training will need to be tailored and targeted. And an accurate diagnosis of the specific local employment and skills challenges this policy seeks to address needs to be the first step to understanding how the overall evidence applies in any given situation.

Providing general guidance on what works

The 'Best Bets' highlight the common characteristics of employment training programmes and projects that have positive effects. For example:

- If an employment training programme wants to improve the employment prospects for an individual, it's probably a good idea to involve employers in the design of the programme and through providing on-the-job training.

- The greater the length of a training programme, the greater the need for extra support to participants to counteract the 'lock-in' effects of being taken away from job searching.
- Extending shorter length training programmes to a wider group of recipients is likely to produce a better success rate than providing longer training to a smaller group.

While the 'Best Bets' cannot provide definitive guidance as to what will or won't work in any specific context, they do provide useful overall guidance to policy-makers to use when designing an employment training programme. They also raise a note of caution for policy-makers if they decide to try out a programme which has not worked so well elsewhere.

Providing detailed evidence on specific programmes

The 71 evaluations offer a rich source of material for policy-makers to use in designing specific employment training policies. In particular the evaluations will be of use to policy-makers at two key stages in the policy design process: determining the policy options, and then selecting the preferred option. For both stages, the policy-makers should ensure that their understanding of their specific situation and the different policy options available is as detailed and comprehensive as possible.

The source evaluations can be used to compare different policy interventions, for example longer and shorter training programmes (see evaluations 85, 88, 149, 218, 220, 221, 236, 240, 243, 246), or in-firm and classroom based programmes (see evaluations 132, 145, 146, 218, 219, 221, 225, 233, 234, 235, 237, 238, 240, 243, 251).

They can also provide detailed insights once the policy option has been selected. For example, for policy-makers wanting to design short training programmes, evaluations 84, 90, 94, 96, 146, 235, 238, 239, 244 and 254 will provide detailed insights about these types of programmes. Whereas, for policy-makers interested in designing employment training policies that address the closure of a major employer evaluations 89, 94, 110, 111, 116, 117, 134, 146, 147, 209, 223, 240 and 252 will be particularly useful.

Helping to fill the evidence gaps

The employment training evidence base is nowhere near complete. For example, this review has not found answers to some of the questions which are foremost in policy-makers' minds:

- We have found no evidence which provides robust, consistent insight into relative value for money of different approaches.
- We have found no evidence which supports the use of local over national delivery models.

The gaps highlight the need for more policy experimentation, and specifically for experiments that identify how different elements of employment training policy design contribute to better (or worse) employment and wages outcomes; and the value for money of different approaches.

More experimentation also requires evaluation to be embedded in the employment training policy design process, rather than being thought about as a last minute add-on. This means policy-makers need to think differently about the policy-making cycle as a whole.

The Centre's longer term objectives are to ensure that robust evidence is embedded in the development of policy, that these polices are effectively evaluated and that feedback is used to improve them. To achieve these objectives we want to:

- work with local decision-makers to improve evaluation standards so that we can learn more about what policies work, where.
- set up a series of ‘demonstration projects’ to show how effective evaluation can work in practice.

Interested policy-makers please get in touch.

10

Bibliography

Barro (2001) Human Capital and Growth, American Economic Review 91: 12-17

Becker (1962) Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education, Chicago: University of Chicago Press.

Card, Kluve and Weber (2010). Active Labour Market Policy Evaluations: A Meta-Analysis. The Economic Journal 120, F452-F477.

Cohen and Soto (2007) Growth and human capital: good data, good results, Journal of Economic Growth 12: 51-76.

Duranton and Puga (2014) The Growth of Cities, Handbook of Economic Growth. 781-853

Gibbons, Nathan and Overman (2014), 'Evaluating Spatial Policies', SERC Policy Paper SERCPPXXX.

Glaeser and Maré (2001) Cities and Skills, Journal of Labor Economics 19: 316-342

Glaeser and Saiz (2004) The Rise of the Skilled City, Brookings-Wharton Papers on Urban Affairs 5: 47-94

Hanushek and Woessmann (2012) Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation, Journal of Economic Growth 17: 267-321.

Henderson (2007) Understanding knowledge spillovers, Regional Science and Urban Economics 37: 497-508

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

Moretti (2004a) Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data, Journal of Econometrics 121: 175-212

Moretti (2004b) Human capital externalities in cities, in Henderson JV and Thisse J-F (eds) Handbook of Regional and Urban Economics. Elsevier, 2243-2291.

Moretti (2004c) Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions, The American Economic Review 94: 656-690

OECD (2009) Designing Local Skills Strategies, Paris: OECD

Payne and Keep (2011) 'One Step Forward, Two Steps Back? Skills Policy in England under the Coalition Government', SKOPE Research Paper 102, Cardiff and Oxford Universities

Shapiro (2006) Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital, Review of Economics and Statistics 88: 324-335.

Sherman, Gottfredson, MacKenzie, Eck,Reuter, and Bushway (1998) Preventing Crime: What Works, What Doesn't, What's Promising, National Institute of Justice Research Brief, US Department of Justice.

Appendix A: Findings by Outcome

Outcome type	Positive	Zero	Negative	Mixed	Share positive
Employment	84, 87, 94, 109, 110, 111, 116, 132, 140, 145, 151, 218, 219, 220, 222, 225, 227, 232, 233, 234, 240, 242, 244, 246, 247, 249, 251, 254, 256,	115, 123, 130, 147, 150, 152, 223, 226, 231, 237, 238, 245	120, 128, 228.	79, 85, 88, 90, 99, 117, 126, 127, 146, 149, 221, 229, 230, 235, 236, 239, 243, 250, 252.	29/63
Wages	80, 89, 125, 137, 138, 149, 209, 249, 250, 251, 256	120, 152, 223, 237, 245	99, 134, 238	96, 128	11/21

Appendix B: Shortlisted Evaluations

Number	Reference
79	Bloom, H.S. (1997). The Benefits and Costs of JTPA Title II-A Programs: Key Findings from the National Job Training Partnership Act Study. Journal of Human Resources *32*, 549–576.
80	Winter-Ebmer, R. (2006). Coping with a Structural Crisis: Evaluating an Innovative Redundancy-Retraining Project. International Journal of Manpower *27*, 700–721.
84	Kluve, J., Lehmann, H., and Schmidt, C. (2008). Disentangling Treatment Effects of Active Labor Market Policies: The Role of Labor Force Status Sequences. Labour Economics 15, 1270–1295.
85	Kluve, J., Schneider, H., Uhlendorff, A., and Zhao, Z. (2012). Evaluating continuous training programmes by using the generalized propensity score. Journal of the Royal Statistical Society Series a-Statistics in Society 175, 587–617.
87	Firth, D., Payne, C., and Payne, J. (1999). Efficacy of programmes for the unemployed: discrete time modelling of duration data from a matched comparison study. Journal of the Royal Statistical Society Series a-Statistics in Society 162, 111–120.
88	Kluve, J., Rinne, U., Uhlendorff, A., and Zhao, Z. (2013). The impact of training duration on employment outcomes: Evidence from LATE estimates. Economics Letters 120, 487–490.
89	Flores, C., and Flores-Lagunes, A. (2010). Partial Identification of Local Average Treatment Effects with an Invalid Instrument.
90	Lalive, R., van Ours, J., and Zweimuller, J. (2008). The impact of active labour market programmes on the duration of unemployment in Switzerland. Economic Journal 118, 235–257.
94	Ifcher, J. (2010). Identifying the Effect of a Welfare-to-Work Program Using Program Capacity Constraints: A New York City Quasi-experiment. Eastern Economic Journal 36, 299–316.
96	Raaum, O., and Torp, H. (2002). Labour market training in Norway - effect on earnings. Labour Economics 9, 207–247.
99	The U.S. Department of Labor Employment and Training Administration Office of Policy Development and Research (2012). Evaluation of the Trade Adjustment Assistance Program.

Number	Reference
109	Eichler, M., and Lechner, M. (2002). An evaluation of public employment programmes in the East German State of Sachsen-Anhalt. Labour Economics 9, 143–186.
110	Cavaco, S., Fougere, D., and Pouget, J. (2013). Estimating the effect of a retraining program on the re-employment rate of displaced workers. Empirical Economics 44, 261–287.
111	Fitzenberger, B., Osikominu, A., and Volter, R. (2008). Get Training or Wait? Long-Run Employment Effects of Training Programs for the Unemployed in West Germany. Annales d'Economie et de Statistique 321–355.
115	Conniffe, D., Gash, V., and O'Connell, P. (2000). Evaluating state programmes: "Natural experiments" and Propensity Scores. Economic and Social Review 31, 283–308.
116	Fitzenberger, B., and Besser, S. (2007). Employment effects of the provision of specific professional skills and techniques in Germany. Empirical Economics 32, 529–573.
117	Gerfin, M., and Lechner, M. (2002). A microeconometric evaluation of the active labour market policy in Switzerland. Economic Journal 112, 854–893.
120	Lechner, M., and Wunsch, C. (2009). Active labour market policy in East Germany. Economics of Transition 17, 661–702.
123	Leigh, D. (1995). Can a voluntary workfare program change the behavior of welfare recipients - new evidence from Washington state's Family Independence Program (FIP). Journal of Policy Analysis and Management 14, 567–589.
125	Dahl, E., and Lorentzen, T. (2005). What works for whom? An analysis of active labour market programmes in Norway. International Journal of Social Welfare 14, 86–98.
126	Huber, M., Lechner, M., Wunsch, C., and Walter, T. (2011). Do German Welfare-to-Work Programmes Reduce Welfare Dependency and Increase Employment? German Economic Review 12, 182–204.
127	Andren, T. & Andren, D. (2006). Assessing the employment effects of vocational training using a one-factor model, Applied Economics, 2006, 38, 2469–2486.

Number	Reference
128	Larsson, L. (2003). Evaluation of Swedish youth labor market programs. Journal of Human Resources 38, 891–927.
130	Lechner, M. (2000). An evaluation of public-sector-sponsored continuous vocational training programs in East Germany. Journal of Human Resources 35, 347–375.
132	Neubaumer, R. (2012). Bringing the unemployed back to work in Germany: training programs or wage subsidies? International Journal of Manpower 33, 159–177.
134	Lee, D. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. Review of Economic Studies 76, 1071–1102.
137	Raphael, S., and Stoll, M. (2006). Evaluating the effectiveness of the Massachusetts workforce development system using no-shows as a nonexperimental comparison group. Evaluation Review 30, 379–429.
138	Mueser, Peter ; Heinrich, Carolyn J. ; Troske, Kenneth R. International, LLC. (2009). Workforce Investment Act Non-Experimental Net Impact Evaluation.
140	Hotz, V.J., Imbens, G.W., and Klerman, J.A. (2000a). The Long-Term Gains from GAIN: A Re-Analysis of the Impacts of the California GAIN Program.
145	Kraus, F., Puhani, P., and V, S. (1999). Employment effects of publicly financed training programs - The East German experience. Jahrbucher Fur Nationalokonomie Und Statistik 219, 216–248.
146	Kopf, E. (2013). Short training for welfare recipients in Germany: which types work? International Journal of Manpower 34, 486–516.
147	Lechner, M. (1999). Earnings and employment effects of continuous off-the-job training in East Germany after unification. Journal of Business & Economic Statistics 17, 74–90.
149	Lechner, M., Miquel, R., and Wunsch, C. (2007). The curse and blessing of training the unemployed in a changing economy: The case of East Germany after unification. German Economic Review 8, 468–509.

Number	Reference
150	Magnac, T. (2000). Subsidised training and youth employment: Distinguishing unobserved heterogeneity from state dependence in labour market histories. Economic Journal 110, 805–837.
151	Perry, G., and Maloney, T. (2007). Evaluating active labour market programmes in New Zealand. International Journal of Manpower 28, 7–29.
152	Puma, M., and Burstein, N. (1994). The National Evaluation of Food Stamp Employment and Training-program. Journal of Policy Analysis and Management *13*, 311–330.
209	Schochet, Peter Z. Burghardt, John and McConnell, Sheena (2008) Does Job Corps Work? Impact Findings from the National Job Corps Study
218	Biewen, M., Fitzenberger, B., Osikominu, A., and Waller, M. (2007). Which Program for Whom? Evidence on the Comparative Effectiveness of Public Sponsored Training Programs in Germany (ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research).
219	Blache, G. (2011). Active labour market policies in Denmark: A comparative analysis of post-program effects (Université Panthéon-Sorbonne (Paris 1), Centre d'Economie de la Sorbonne).
220	Fitzenberger, B., Osikominu, A., and Paul, M. (2010). The heterogeneous effects of training incidence and duration on labor market transitions (ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research).
221	Fitzenberger, B., and Volter, R. (2007). Long-Run Effects of Training Programs for the Unemployed in East Germany. Labour Economics *14*, 730–755.
222	Rotar, L.J. (2012). How Effective Is the Slovenian Institutional Training Program in Improving Youth's Chances of Reemployment? Eastern European Economics 50, 94–106.
223	Schwerdt, G., Messer, D., Woessmann, L., and Wolter, S.C. (2012). The impact of an adult education voucher program: Evidence from a randomized field experiment. Journal of Public Economics *96*, 569–583.
225	Brodaty, T., Crépon, B., and Fougère, D. (2001). Using Matching Estimators to Evaluate Alternative Youth Employment Programs: Evidence from France, 1986-1988.

Number	Reference
226	Caroleo, F.E., and Pastore, F. (2002). Training Policy for Youth Unemployed in a Sample of European Countries (CELPE (Centre of Labour Economics and Economic Policy), University of Salerno, Italy).
227	Koning, J. de, Koss, M., and Verkaik, A. (1991). A quasi-experimental evaluation of the Vocational Training Centre for Adults. Environment and Planning C: Government and Policy *9*, 143–153.
228	Dettmann, E., and Gunther, J. (2010). Subsidized Vocational Training: Stepping Stone or Trap? An Evaluation Study for East Germany.
229	Ehlert, C., Kluve, J., and Schaffner, S. (2011). Training + Temp Work = Stepping Stone? – Evaluating an Innovative Activation Program for Disadvantaged Youths (Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Ruhr-Universität Bochum, Universität Dortmund, Universität Duisburg-Essen).
230	Ferracci, M., Jolivet, G., and van den Berg, G.J. (2010). Treatment evaluation in the case of interactions within markets (IFAU - Institute for Evaluation of Labour Market and Education Policy).
231	Bolvig, I., Jensen, P., and Rosholm, M. (2003). The Employment Effects of Active Social Policy (Institute for the Study of Labor (IZA)).
232	Blasco, S., Crépon, B., and Kamionka, T. (2012). The Effects of On-the-job and Out-of-Employment Training Programmes on Labor Market Histories (CEPREMAP).
233	Carling, K., and Richardson, K. (2001). The relative efficiency of labor market programs: Swedish experience from the 1990's (IFAU - Institute for Evaluation of Labour Market and Education Policy).
234	Zhang, T. (2003). Identifying treatment effects of active labour market programmes for Norwegian adults (Oslo University, Department of Economics).
235	Wolff, J., and Jozwiak, E. (2007). Does short-term training activate means-tested unemployment benefit recipients in Germany? (Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg, Germany]).
236	Crépon, B., Ferracci, M., and Fougère, D. (2007). Training the Unemployed in France: How Does It Affect Unemployment Duration and Recurrence? (Institute for the Study of Labor (IZA)).

Number	Reference
237	Hämäläinen, K., and Ollikainen, V. (2004). Differential effects of active labour market programmes in the early stages of a young person's unemployment. Government Institute for Economic Research, Helsinki.
238	Wodon, Q., and Minowa, M. (2001). Training for the Urban Unemployed: A Reevaluation of Mexico's Training Program, Probecat (University Library of Munich, Germany).
239	Hämäläinen, K., and Tuomala, J. (2007). Vocational Labour Market Training in Promoting Youth Employment (Government Institute for Economic Research Finland (VATT)).
240	Stephan, G., and Pahnke, A. (2008). The Relative Effectiveness of Selected Active Labour Market Programmes and the Common Support Problem (Institute for the Study of Labor (IZA)).
242	Johansson, P. (2006). Using internal replication to establish a treatment effect.
243	Sianesi, B. (2001). Differential effects of Swedish active labour market programmes for unemployed adults during the 1990s.
244	Kluve, J., Lehmann, H., and Schmidt, C.M. (1999). Active Labour Market Policies in Poland: Human Capital Enhancement, Stigmatization or Benefit Churning? (C.E.P.R. Discussion Papers).
245	Lechner, M. (1995). Effects of continuous off-the-job training in East Germany after unification (ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research).
246	McGuinness, S., O'Connell, P., and Kelly, E. (2011). One Dummy Won't Get it: The Impact of Training Programme Type and Duration on the Employment Chances of the Unemployed in Ireland.
247	McVicar, D., and Podivinsky, J.M. (2003). Unemployment Duration Before and After New Deal (Royal Economic Society).
249	Schiller, B.R. (1978). Lessons from WIN: A Manpower Evaluation. The Journal of Human Resources *13*, 502–523.

Number	Reference
250	Rinne, U., Uhlendorff, A., and Zhao, Z. (2008). Vouchers and Caseworkers in Public Training Programs: Evidence from the Hartz Reform in Germany (Institute for the Study of Labor (IZA)).
251	Rinne, U., Schneider, M., and Uhlendorff, A. (2007). Too Bad to Benefit? Effect Heterogeneity of Public Training Programs (Institute for the Study of Labor (IZA)).
252	Richardson, K., and van den Berg, G.J. (2008). Duration dependence versus unobserved heterogeneity in treatment effects: Swedish labor market training and the transition rate to employment (IFAU - Institute for Evaluation of Labour Market and Education Policy).
254	Puhani, P.A. (2002). Advantage through Training in Poland? A Microeconometric Evaluation of the Employment Effects of Training and Job Subsidy Programmes. LABOUR *16*, 569–608.
256	Pessoa e Costa, S., Robin, S. (2009) An Illustration of the Returns to Evaluation of the "Qualifying Contract" in France. Discussion Paper 2009-38, Institut de Recherches Economiques et Sociales de l'Universite catholique de Louvain (UCL).

For a full list of search terms, please visit:
http://www.whatworksgrowth.org/policy-reviews/employment-training/search-terms/

The What Works Centre for Local Economic Growth is a collaboration between the London School of Economics and Political Science (LSE), Centre for Cities and Arup.

www.whatworksgrowth.org

what works centre for
local economic growth

This work is published by the What Works Centre for Local Economic Growth, which is funded by a grant from the Economic and Social Research Council, the Department for Business, Innovation and Skills and the Department of Communities and Local Government. The support of the Funders is acknowledged. The views expressed are those of the Centre and do not represent the views of the Funders.

HM Government

June 2016

What Works Centre for Local
Economic Growth

info@whatworksgrowth.org
@whatworksgrowth

www.whatworksgrowth.org