Study on the Return on ESF Investment in Human Capital

Final Report
August 2010

Evaluation for the European Commission
Directorate General for Employment, Social Affairs and Equal Opportunities
This study was commissioned by the European Commission, DG Employment, under the Multiple Framework Contract ‘Provision of evaluation and evaluation related services to DG EMPL, including support for Impact Assessment Activities’ (Lot N°4 - VT/2009/060 - N°007) signed between DG EMPL and Euréval in association with Ecorys and Rambøll Management (hereafter referred to as The Consortium).

The study team was led by Mr Jacques Toulemonde (toulemonde@eureval.fr).

The study was managed by Mr Dirk Reyntjens. Its progress was monitored by a steering group composed of members of DG Employment and representatives of several concerned Member States.

The opinions expressed in this document represent the authors’ points of view which are not necessarily shared by the European Commission.
Executive summary

This study applies to human capital investments understood as public expenditures in: (1) education and training of employees, job-seekers and/or inactive people, (2) internships and temporary work placement associated with education or training, and (3) strengthening the capacity of education and training systems and structures. Overall, such investments account for over 50% of the ESF resources over the current programming period.

The study included three clusters of tasks as follow:

- A synthesis of available research findings related to the return on human capital investments, with a focus on: (1) the impact of training programmes targeting employed people, and (2) the impact of return-to-work programmes targeting job seekers. Based on a first extensive screening of peer-reviewed publications, the research synthesis was improved with the assistance of a panel of seven experts from the EU and the USA.
- A series of impact analyses based on individual data provided by four volunteering Authorities respectively in Belgium (Flanders), Italy (Northern regions), Hungary, and Poland. Through these analyses, the study team tested a series of impact analysis techniques, all including control groups. Due to the nature of the available data, this part of the study applied to unemployed people only.
- An in-depth investigation into four success stories of impact analyses that had been carried out in the same four countries between 2005 and 2009. This investigation focused on the process of learning from impact evaluations. The study team discussed the recommendations stemming from this investigation in a focus group meeting attended by several managing authorities.

Accumulated knowledge

Building upon an extensive review of the literature and the assistance of a panel of international experts, the study shows that:

- Training of employed people is socially profitable but not rewarding enough at the level of the firm. For this reason there is an under-investment of private funds in this area.

- Training of unemployed people is socially profitable, but not for all target groups, and under the condition that it is associated with job-search assistance. Training benefits are positive in the medium term only. Impact in the short term is negative because people stop seeking jobs when they participate in a training programme (lock-in effect).

- There are considerable knowledge gaps in areas that are fundamental for justifying ESF investments in human capital, such as the impact of training ‘at-risk’ workers, or the fact that a trainee may find a job at the expense of someone else (substitution).
Training of employed people (see Box 1)

The short term return on training investment is not clearly established for an individual trainee or a company. However recent studies have shown a **positive impact on productivity at industry level** and this impact is sufficiently high to argue that human capital investments is profitable and competitive with other investments.

The problem is that only part of the benefits goes to the company that pays for the training. The major part may go to other companies in the industry and to the employees who received the training, especially when they move to other companies. This raises the question of underinvestment by the employers who do not have sufficient incentives to invest in the training of their staff. Since the larger part of the returns to human capital investments actually goes to the industry rather than to the individual firms, there is a rationale for **public investments in systems** and structures in charge of training at the industry level.

Off-the-job training is known to be more effective than on-the-job training, provided that trainees have the opportunity to apply their newly acquired skills when returning to the work place. This is something that calls for a close connection between training and human resource management.

Training of unemployed people (see Box 3)

The impact of return-to-work training is **positive in the medium term**, only after balancing out the negative lock-in effects, i.e. the fact that people stop seeking jobs when they participate in a training programme. Impact is quite limited in the case of young job seekers.

Moreover, significant impact is achieved **only if training is associated** with job-search assistance and/or a policy framework that creates incentives for seeking jobs. This calls for a closer connection between ESF investment in return-to-work training and the reform of national labour-market policies.

Knowledge gaps (see Box 2 and Box 4)

There is a specific knowledge gap as regards **at-risk industries and firms**, where human capital investments might be particularly valuable, but where shortages of resources may generate under-investments.

**Skill obsolescence** and the impact of training on aged staff members is another field where more knowledge should be accumulated. This is particularly relevant in the case of the development of greying economies in Europe.

Very little research exists on the economic returns on training unemployed people, because: (1) costs and benefits tend to be studied separately, and (2) limited efforts have been made at valuing **social benefits** resulting from inclusion.

There is insufficient evidence on the important issue of **quality of training**, in terms both of how to assess quality and of whether the benefits of higher quality are worth the cost.

Similarly, little is known about the patterns of employment impacts in relation to **specific target groups**, except that young job seekers are systematically assessed as less successful with labour market training.

No clear conclusion can be drawn concerning the possible pattern of training effectiveness over the **business cycle**. There are some indications that the training effect is pro-cyclical, in the sense that the impact of return-to-work training is higher in periods with good job opportunities and low unemployment, but the picture is not clear.
Finally, the whole issue of secondary impacts remains a black box. The main concern is about substitution effect, i.e. the fact that an unemployed trainee finds a job at the expense of someone else. Substitution is assumed to be limited where a training programme is designed for matching some skill gap. Virtuous circles are also assumed to emerge through improved productivity and innovation, reinforced competitiveness, large market shares, and secondary impact on employment.

**Return on public investment** (see 4.5.2)

There is a rationale for investing public money in systems and structures at the level of industry, but this study could not assess the social return on such investments.

Building on a fragile evidence base, the study concludes that training the unemployed may be profitable under strict conditions. In order to justify public investments in this area, policy-makers and programme managers should demonstrate that they control the factors that ensure a positive rate of return on training investment, i.e. not targeting young people, combining job-search assistance and training, keeping the cost of training reasonable, ensuring high enough quality, and paying due attention to the issue of substitution.

This study could not conclude on the social profitability of investing public money in training employed people. It is clear that there is a *rationale for taking public action at the level of industries* rather than firms. How should such an action be structured remains an open question, and especially the following sub-questions:

- Should public policies be based on regulations (obligation to train) or public investment through financial incentives?
- Under which assumptions would public investments be socially profitable?
- What is the relevant level for designing and implementing such policies (regional, national, European)?
- If the relevant level is not that of the EU as a whole, then is there a residual need for taking action at EU level? ... and if yes, what should be the instruments (soft coordination, financial incentives, or both)?

As regards the training of unemployed people, this study suggests that there is a cluster of reasonable assumptions under which public investment is socially profitable. These assumptions are summarised as follows:

- The success rate (additional probability of finding a job) should be in a range of 10-20%. The smallest figure is a profitability threshold and the largest figure is the most optimistic assumption that can be reasonably made.
- The training programme should be combined with job-search assistance and its duration should be in a range of 2 to 4 months. Other types of training might also be profitable, provided that their cost does not exceed that of a 2-4 month programme.
- The duration of benefits should be 8 years or more, an assumption which seems to be consistent with recent and credible research works.
- No more than 50% of the participants having found a job should have succeeded at the expense of someone else (substitution), a proportion of 20% being a very optimistic assumption.
Social benefits should amount to at least 50% of individual benefits (i.e. gross salary of the job found), and this threshold may have to be increased to 100% or more if other assumptions are pessimistic.

Despite the shaky nature of the available evidence, the study team concludes that **training the unemployed may be profitable under strict conditions**. In order to justify public investments in this area, policy-makers and programme managers should demonstrate that they control the factors that make training programmes socially profitable. In particular, they should:

- Avoid targeting young unemployed people;
- Combine training and job-search assistance;
- Keep the cost of training reasonable, i.e. equivalent to that of a 2 to 4 month programme;
- Ensure that subsidised programmes have a high enough quality so that they can be assumed to reach a success threshold of 10%;
- Ensure that the issue of substitution is paid sufficient attention so that it can be assumed to remain under 50% (proportion of successful participants who find their job at the expense of someone else).

Even if such principles are applied, the **profitability remains conditional to some factors which cannot be controlled**:

- The duration of benefits should be 8 years or more, which means that there should be a long enough time period before ageing, skill obsolescence, and business cycles reduce the employability of successful trainees.
- Social benefits should amount to at least 50% of individual benefits (i.e. gross salary of the job found).

In the case of the two above conditions, there is a severe lack of knowledge. As a consequence, it would be difficult to justify public investments in training the unemployed if no effort is made in parallel towards bridging these knowledge gaps.

**Lessons from the methodological tests** (see 5.2.1)

From the impact analyses undertaken in cooperation with four Member States, the study concludes that sufficient attention, skills, and resources should be devoted to data collection and impact analysis. Impact analyses should be designed on a case by case basis since there is nothing like a gold standard approach.

It takes years for deriving lessons from an impact analysis. This means that impacts cannot be included in a performance monitoring system.

More lessons can be learned from narrow studies focusing on specific training approaches, specific target groups, and/or specific territories, especially if such studies cut across regions or countries.

For several reasons, this study casts serious doubts about the relevance of conditioning annual payments to results, at least as far as performance indicators are not complemented with sound impact analyses.

The study confirms that there is a large discrepancy between actual impact and what is measured with impact indicators such as "proportion of trainees having found a job after one year". The four country studies show that where control groups are used for analysing causes and effects, impact indicators have to be divided by a factor ranging from 2 to 9 (see Table 39).
Moreover, the discrepancy between impact indicators and the estimates of actual impacts do not follow a standard pattern across publics and countries. This casts serious doubts about the relevance of conditioning annual payments to results, at least as far as performance measurement systems are not complemented with sound impact analysis. However, impact analyses cannot be implemented every year and everywhere.

The study has shown considerable cross-country differences in terms of impact estimates. Such differences call for a cautious interpretation, but they could create opportunities for learning, for instance if they can be attributed to variations in the quality of training or to differences in the economic environment. However, there are at least two instances where exceptionally high or low impacts had to be attributed to unreliable data or inappropriate control groups. The lesson is that data collection and evaluation design matter, and that sufficient attention, skills, and resources should therefore be devoted to them.

Depending on the analysis technique chosen, the evaluation should be launched up to one year before training. The study has shown that impact should be analysed over a period of one to two years after the training. An additional year should be added for the analysis itself, drawing conclusions, and learning lessons. Overall, the duration of the learning cycle may be in a range of 2.5 to 4 years. This shows that impact analyses can feed into long learning cycles only. Again, this is not consistent with a system that would focus on annual result-based conditionalities.

Many findings of the study called for interpretation and explanations. This showed that impact estimates seldom speak for themselves. Most often, they need to be included in some evaluation process where a common understanding can be developed in an interactive way.

Finally, the findings of the study were difficult to interpret in cases where a wide variety of training approaches were analysed together. More lessons could be learned from narrow studies focusing on specific training approaches, specific target groups, and/or specific regions, especially if the study had a comparative perspective. The other side of the coin is that impact analyses should not cover a programme as a whole, or even a component of the programme such as a priority axis. This is one more reason why actual impact could not feed into a management system focusing on result-based conditionalities.

As regards impact analysis techniques, the study does not make any general recommendations. There is nothing like a gold standard approach to evaluating impact. On the contrary, a methodological approach is good as far as it has a good fit with the purpose and context of the evaluation. The study includes a series of recommendations for taking the main methodological options (see 5.2.2).
Learning about the learning processes (see 7.1 and 7.3)

Building upon an in-depth investigation into four success stories of impact analyses, the study concludes that they all delivered new, credible, and relevant lessons, but that actual learning from these lessons was limited.

The study team makes a series of proposals for bringing Managing Authorities to learning more and faster from impact evaluations. These proposals build upon the idea that learning should be part of an EU-wide permanent process, and no longer be restricted to the perimeter and time frame of individual programmes.

These proposals involve obligations, incentives and advices related to:

- Systematising multi-annual evaluation plans, and follow up of recommendations;
- Applying the EC practice of impact assessment at the level of programmes;
- Making actual learning from impact analyses a financial conditionality;
- Earmarking funds for impact evaluations bridging the identified knowledge gaps;
- Encouraging the evaluation of the impact of pilot training approaches;
- Disseminating technical guidance on impact analysis approaches
- Fostering mutual learning about evidence-based policy-making

The study team investigated four success stories of impact analyses and came up with mixed findings. On the positive side, it was found that the impact analyses delivered new, credible, and relevant lessons. On the negative side, it must be said that actual learning from these lessons tended to be limited. This is explained as follows:

- Impact evaluations were pushed by information suppliers much more than they were pulled by policy-makers. Good evaluation practices were mainly driven by administrative capacity, good monitoring and statistical databases, and institutionalisation of evaluation. The need for acquiring highly credible knowledge was not the primary reason why these evaluations were launched. Limited attention was paid to matching the political agendas.
- Releasing the evaluation findings to policy-makers contributed very little to the uptake of such findings, especially because technical issues tended to be given too much attention at the expense of substantial conclusions.
- The newly acquired knowledge was often conveyed to policy makers in an indirect way, by feeding into streams of successive evaluations and studies before being ready for use in the policy-making process.
- Policy-makers had a limited capacity to absorb evaluation findings, unless they had a lasting proximity with a stable group of technical and administrative experts.

In addition to the above conclusion, this part of the study confirms that, even in exceptionally favourable circumstances, cost and feasibility constraints make it impossible to undertake impact analyses in a comprehensive and/or...
systematic manner. In other words, impact cannot be known every year and everywhere.

The study team makes a series of proposals for developing a culture of learning among those involved in managing the ESF. This set of proposals builds upon the idea that **learning from impact analyses should be part of an EU-wide permanent process**, and no longer be restricted to the perimeter and time frame of individual programmes.

Implementing these recommendations would represent a significant additional effort on the side of Managing Authorities. Inevitably, such efforts would be perceived as a supplementary burden, unless **other obligations are downsized in parallel**. This could be achieved by:

- Alleviating management and audit activities through approaches like that of budget support used in the area of development aid;
- Lightening the demands in terms of reporting, especially as regards monitoring indicators.

**Making policy-makers accountable for learning**

The set of regulations applying to the next programming cycle might include a series of provisions aimed at bridging the gap between evaluation, evidence, and knowledge on the one hand, and policy-making on the other:

- Every year, Managing Authorities would have to approve a rolling multi-annual evaluation plan. Over a five year period, each plan should make at least one contribution to bridging one of the identified knowledge gaps through a sound impact analysis.
- Every year at the same time, Managing Authorities should approve an annual follow-up report identifying the main lessons learned from previous evaluations, and the actions taken.
- A synthesis of relevant knowledge (evidence base) should be included in: (1) all evaluation reports, and (2) all decisions having a significant financial impact. This would mirror the practice of impact assessment in the European Commission.
- By implementing the above three rules, Managing Authorities should demonstrate actual progress towards learning from impact analyses and towards evidence-based policy-making. A part of the disbursements should be contingent on demonstrating such progress.

**Managing knowledge through incentives**

Specific funds should be earmarked for evaluation and allocated to, and only to:

- Designing and evaluating pilot training approaches;
- Carrying out sound impact analyses aimed at bridging one of the identified knowledge gaps;
- Clustered evaluations addressing the same cause-and-effect question in several regions or several countries;
- Syntheses, quality assessments, peer reviews, and discussion of research and evaluation findings within knowledge communities (e.g. networks, conferences) involving practitioners.

**Providing technical guidance**

The Commission or any relevant body with the Commission’s support should create publicly available guidance on:

- Identified knowledge gaps and sound impact analysis techniques that are relevant for evaluations aimed at bridging such gaps;
• Accessing data bases and dealing with privacy issues;
• Assessing the quality of impact analyses and dealing with peer reviews;
• Effective approaches to translating impact analyses into usable knowledge.

**Mutual learning on evidence-based policy-making**

During the events gathering Managing Authorities, sufficient time should be devoted to learning from one another about evidence-based policy-making.
Cette étude porte sur les investissements en capital humain, compris ici comme les dépenses publiques dans le domaine de : (1) l'instruction et la formation des personnes employées, demandeuses d'emploi et/ou inactives (2) les stages et contrats en alternance associés à un enseignement ou une formation, et (3) le soutien apporté aux systèmes et structures d'enseignement et de formation. Pour la période actuelle de programmation, ces investissements représentent 50 % des ressources des FSE.

L'étude a été structurée en trois ensembles de tâches :

- Une synthèse des résultats d'études disponibles en lien avec le retour sur l'investissement en capital humain, en se concentrant sur : (1) l'impact des programmes de formation pour les personnes employées, et (2) l'impact des programmes de retour à l'emploi pour les demandeurs d'emploi. S'appuyant sur une revue approfondie des publications validées par les pairs, cette synthèse des études a été consolidée par les conseils et avis d'un panel de sept experts européens et américains.
- Une série d'analyses d'impact basées sur des données individuelles mises à disposition par quatre autorités de gestion volontaires en Belgique(Flandres), Italie (régions du Nord), Hongrie et Pologne. Ces analyses ont permis de tester une variété de techniques d'analyse d'impact, comprenant toutes des groupes de contrôle. Cette partie de l'étude ne porte que sur les demandeurs d'emploi puisque les données disponibles concernaient ce groupe uniquement.
- Une étude approfondie de quatre d'analyses d’impact, considérées comme des succès, effectuées dans les quatre mêmes pays entre 2005-2009. Cette recherche se concentre sur les processus d'apprentissage à partir de l'évaluation des impacts. L'équipe de recherche a présenté les recommandations issues de cette étude dans le cadre d'un groupe de travail réunissant plusieurs autorités de gestion.

**Connaissances acquises**

S'appuyant sur une large revue de la littérature et un panel d'experts internationaux, l'étude révèle que :

- La formation des personnes employées profite à la société dans son ensemble, mais n'a pas assez de retombées positives au niveau de l'entreprise. Ceci explique le sous-investissement des fonds privés dans ce domaine.

- La formation pour les demandeurs d'emploi profite à la société en général, mais pas à tous les publics cibles, et seulement si elle est associée à un accompagnement à la recherche d'emploi. La formation a un impact positif uniquement à moyen terme. Les impacts à court terme sont négatifs étant donné que les demandeurs d'emploi arrêtent de chercher du travail pendant leur participation à des programmes de formation (*Lock-in-effect*).
Il y a un manque de connaissances solides dans un grand nombre de domaines qui sont essentiels à la justification des investissements du FSE en capital humain, tel que l’impact de la formation sur les travailleurs « à risque » ou la capacité des personnes ayant suivi une formation à trouver un travail aux dépens d’un autre demandeur d’emploi (effet de substitution).

La formation des personnes employées (voir encadré Box 1)

À court terme, le retour sur investissement n’est pas clairement établi dans la cas d’une personne employée ou de son employeur. Néanmoins, des études récentes ont montré un impact positif sur la productivité au niveau des secteurs d’activité et cet impact est suffisamment fort pour considérer que les investissements en capital humain sont profitables et compétitifs par rapport à d’autres types d’investissements.

Le problème est qu’une partie seulement des avantages revient à l’entreprise qui finance la formation. La majeure partie peut profiter à d’autres entreprises de la même branche et aux personnes formées, particulièrement lorsqu’elles changent d’employeur après la formation. Cela pose la question du sous-investissement de la part des entreprises qui ne sont pas assez incitées à investir dans la formation de leur main-d’œuvre. Puisque la majeure partie du retour sur l’investissement en capital humain profite au secteur d’activité en général plutôt qu’aux entreprises en particulier, cela justifie un investissement public dans les systèmes et structures en charge de la formation au niveau du secteur industriel.

La formation en dehors de l’entreprise est plus efficace que la formation sur site, à condition que les personnes formées aient l’opportunité de mettre en pratique leurs nouvelles compétences quand ils retournent sur leur lieu de travail. Cet élément appelle au rapprochement entre la gestion de la formation et celle des ressources humaines.

La formation des demandeurs d’emploi (voir encadré 3)

L’impact des formations visant au retour à l’emploi est positif à moyen terme seulement après qu’ayant été compensés l’effet de verrouillage, c'est-à-dire le fait que les bénéficiaires arrêtent leur recherche d’emploi durant leur formation. L’impact de la formation est par contre très limité pour les jeunes demandeurs d’emploi.

De plus, un impact significatif est seulement atteint si la formation est associée à un accompagnement à la recherche d’emploi et/ou un cadre réglementaire encourageant à chercher du travail. Ce constat suggère la pertinence d’un rapprochement entre les investissements FSE en formation pour le retour à l’emploi et les reformes nationales des politiques relatives au marché du travail.

Défauts de connaissance (voir encadrés 2 et 4)

Les connaissances font particulièrement défaut en ce qui concerne les secteurs et entreprises à risque, dans lesquels un investissement en capital humain pourrait être très efficace, mais dont le manque de ressources pourrait provoquer un sous-investissement.

L’obsolescence des compétences et l’impact de la formation sur les travailleurs senior sont un autre domaine dans lequel il y aurait lieu d’accumuler plus de connaissances. Ce serait particulièrement justifié au vu du vieillissement affectant les économies européennes.

Il y a très peu de travaux de recherche concernant le retour sur investissement dans la formation des demandeurs d’emploi, notamment parce que : (1)
les coûts et avantages sont souvent étudiés séparément, et (2) peu d’efforts ont été faits pour valoriser les avantages sociaux résultants de l’insertion.

Il n’y a pas assez d’information probante sur l’importante question de la qualité de la formation, en ce qui concerne aussi bien l’appréciation de cette qualité que les avantages qui en découlent et le coût qu’elle engendre.

De même, très peu de connaissances existent sur le profil des impacts sur l’emploi en fonction des publics spécifiquement visés, mis à part que les jeunes chômeurs ont systématiquement moins de réussite après une formation visant un accès à l’emploi.

Aucune conclusion claire n’émerge à propos de la relation entre l’efficacité de la formation et les cycles économiques. Certains résultats indiquent que les effets de la formation sont pro-cycliques, c’est-à-dire que le taux de retour à l’emploi après une formation est plus important en période de création d’emploi et faible taux de chômage, mais l’état des connaissances n’est pas clair.

Enfin, le problème des impacts secondaires reste une boîte noire. Le problème principal est celui de l’effet de substitution, c’est-à-dire le fait qu’un chômeur trouve un travail aux dépens de quelqu’un d’autre. On suppose que l’effet de substitution est limité lorsque la formation concerne une compétence qui fait défaut. On suppose aussi que des cercles vertueux émergent grâce à une meilleure productivité, plus d’innovation, une compétitivité renforcée, des parts de marché gagnées, et donc des effets secondaires sur l’emploi.

**Retour sur investissement public** (voir 4.5.2)

Il y a une justification à l’investissement de fonds publics dans les systèmes et structures au niveau d’un secteur d’activité, mais cette étude n’a pu évaluer les retombées sociales de tels investissements.

Sur la base de preuves fragiles, l’étude conclut que la formation des demandeurs d’emploi peut être efficace, mais sous de strictes conditions. Pour légitimer l’investissement public dans ce domaine, les décideurs politiques et les gestionnaires de programmes devraient prouver qu’ils peuvent contrôler les facteurs garantissant un taux positif de retour sur investissement en formation, c’est-à-dire ne pas cibler les jeunes chômeurs, combiner la formation et l’accompagnement à la recherche d’emploi, limiter le coût de la formation tout en garantissant une qualité suffisante, et porter attention aux effets de substitution.

Cette étude n’a pu conclure sur l’utilité sociale des investissements publics dans la formation des personnes employées. Il est clair qu’il y a une raison d’être à l’action publique au niveau des branches plutôt qu’au niveau des entreprises. La question de savoir comment une telle action devrait être structurée reste ouverte, notamment dans les formes qui suivent :

- Est-ce que les politiques publiques devraient être basées sur des réglementations (obligation de former) ou des investissements publics sous forme d’incitations financières ?
- À quelles conditions l’utilité sociale des investissements publics serait-elle suffisante ?
- Quel est le niveau de gouvernement approprié pour concevoir et mettre en œuvre ces politiques (régional, national, européen) ?
Si le niveau pertinent de gouvernement n’est pas le niveau européen, reste-t-il des points nécessitant une action européenne ?...Si oui, quels devraient être les instruments (coordination ouverte, incitations financières, ou les deux) ?

Pour ce qui est de la formation des demandeurs d’emploi, cette étude suggère qu’il existe un ensemble assez large d’hypothèses sous lesquelles l’investissement public serait socialement profitable. Ces hypothèses peuvent se résumer à :

- Un taux de succès (probabilité additionnelle de trouver un emploi) compris entre 10-20%. Le plus petit taux représentant le seuil de profitabilité et le plus haut l’estimation raisonnable la plus optimiste pouvant être formulée.
- Le programme de formation devrait aller de pair avec un accompagnement à la recherche d’emploi et sa durée devrait s’étendre de 2 à 4 mois. D’autres types de formation peuvent aussi être profitables à condition que leur coût n’excède pas celui d’un programme de 2 à 4 mois.
- Les avantages retirés de la formation devraient se prolonger sur une période de 8 ans où plus, une hypothèse qui semble acceptable au vu de travaux de recherche récents et crédibles.
- Pas plus de 50 % des participants ne devraient trouver un emploi aux dépens de quelqu’un d’autre (effet de substitution), un taux de 20% étant une estimation très optimiste.
- Les avantages sociaux devraient s’élèver à au moins 50 % des bénéfices individuels (salaire brut de l’emploi trouvé), et ce seuil devrait être porté à 100 % ou plus si les autres hypothèses sont pessimistes.

Mis à part la nature encore peu solide des preuves disponibles, l’équipe de travail conclut que la formation des demandeurs d’emploi devrait être profitable sous de strictes conditions. Afin de justifier les investissements publics dans ce domaine, les décideurs publics et les gestionnaires de programme devraient démontrer qu’ils contrôlent les facteurs qui rendent les programmes de formation profitables. En particulier ils devraient :

- Éviter de cibler des jeunes demandeurs d’emploi ;
- Allier la formation et l’accompagnement à la recherche d’emploi ;
- Garder le coût de la formation à un niveau raisonnable, c’est-à-dire équivalent à celui d’un programme de 2 à 4 mois ;
- Garantir que les programmes subventionnés soient d’assez bonne qualité pour considérer qu’ils atteindront un seuil de succès de 10 % ;
- Garantir qu’assez d’attention est consacrée à l’effet de substitution pour envisager qu’il reste en dessous des 50 % (proportion de participants trouvant un emploi aux dépens de quelqu’un d’autre).

Même si ces principes sont appliqués, la profitabilité reste dépendante de facteurs ne pouvant être contrôlés :

- Les avantages devraient s’étendre sur une période de 8 ans ou plus, ce qui signifie qu’il devrait y avoir une période assez longue avant que le vieillissement, l’obsolescence des compétences, et les cycles économiques réduisent l’employabilité des personnes formées.
- Les avantages sociaux devraient atteindre un minimum de 50 % des bénéfices individuels (salaire brut de l’emploi trouvé).

Les deux conditions ci-dessus se caractérisent par une très grande faiblesse des connaissances. Par conséquent, il serait difficile de justifier l’investissement public dans la formation pour les demandeurs d’emploi si aucun effort n’est fait en parallèle pour acquérir des connaissances dans ce domaine.
Enseignements des tests méthodologiques (voir 5.2.1)

En s’appuyant sur les analyses d’impact effectuées en collaboration avec quatre États membres, l’étude conclut que davantage d’attention, de compétences et de ressources devraient être consacrées à la récolte de données et l’analyse d’impact. Les analyses d’impacts devraient être conçues au cas par cas étant donné qu’il n’y a pas « d’étalon or » parmi les différentes approches possibles.

Sachant qu’il faut des années avant de pouvoir tirer les leçons de l’analyse des impacts, il est impossible d’inclure de telles analyses dans les systèmes de suivi des performances.

Davantage de leçons peuvent être apprises si les analyses se concentrent sur des approches de formation spécifiques, des publics cible spécifiques et/ou des territoires particuliers, surtout de telles analyses ont un caractère trans-régional ou transnational.

Pour de nombreuses raisons, cette étude jette un doute sur la pertinence de conditionner les paiements annuels aux résultats, à moins que les indicateurs de performance ne soient complétés par de solides analyses d’impact.

L’étude confirme qu’il y a un décalage important entre l’impact réel et ce qui est mesuré avec des indicateurs d’impact tel que « la proportion de personnes formées ayant trouvé un emploi au bout d’un an ». Les quatre études nationales montrent que lorsque les groupes de contrôle sont utilisés pour analyser des relations de causes à effets, les indicateurs d’impact doivent être divisés par un facteur de 2 à 9 (voir tableau 39).

De plus, le décalage entre indicateurs d’impact et estimations de l’impact réel ne suivent pas un schéma identique selon le type de public et les pays. Cela porte à émettre de sérieux doutes sur la pertinence de conditionner les paiements annuels aux résultats, à moins que les indicateurs de performance ne soient complétés par de solides analyses d’impact. Néanmoins, des analyses d’impact ne peuvent être menées chaque année pour tous les programmes.

L’étude a mis en valeur des différences considérables entre pays en termes d’estimation d’impacts. De telles différences appellent à prendre des précautions lors de leur interprétation, mais elles pourraient constituer des opportunités d’apprentissage, par exemple si ces différences peuvent être attribuées à des variations de qualité de formation ou à différents environnements économiques. Cependant, il y a au moins deux exemples pour lesquels des impacts exceptionnellement forts ou faibles doivent être attribués à des données peu fiables et des groupes de contrôle peu pertinents. L’enseignement à tirer de ce constat est que la collecte de données et la structure de l’évaluation ont leur importance, et qu’il faut leur consacrer assez d’attention, de compétences et de ressources.

Le choix de certaines techniques d’analyse impose de lancer l’évaluation à l’avance, jusqu’à un an avant le début de la formation. L’étude a par ailleurs montré que l’impact devrait être analysé sur une période d’un ou deux ans après la formation. Une année de plus devrait être ajoutée pour l’analyse elle-même, pour formuler les conclusions et pour en tirer les enseignements. Au total, la durée du cycle d’apprentissage peut s’étendre de 2,5 à 4 ans. Ainsi, les analyses d’impact peuvent uniquement s’inscrire dans de longs cycles d’apprentissage. De fait, ces cycles ne sont pas adaptés à un système conditionnant les financements aux résultats annuels.
De nombreux résultats de l’étude requièrent une interprétation et des explications. Cela confirme le fait que les estimations d’impact parlent rarement d’elles-mêmes. En général, elles nécessitent d’être incluses dans un processus d’évaluation pour lequel une compréhension commune peut être développée de façon interactive.

Enfin, les résultats de l’étude étaient difficiles à interpréter dans les cas où une grande variété de démarches de formation étaient analysées ensemble. Davantage d’enseignements pourraient être tirés d’études plus restreintes portant sur des approches particulières de formation, des groupes cibles spécifiques et/ou des territoires particuliers, surtout dans le cadre d’une étude à perspective comparative. La réciproque de ce constat est que les analyses d’impact ne devraient pas porter sur l’intégralité d’un programme, ou même sur d’une composante comme un axe prioritaire par exemple. C’est une raison de plus pour dire que l’impact réel ne peut pas faire partie d’un système de gestion fondé sur des conditionnalités liées aux résultats.

Concernant les techniques d’analyses d’impact, l’étude ne porte pas de recommandations générales. Il n’y a pas d’étalon or pour les techniques d’évaluation des impacts. Au contraire, une approche méthodologique doit être considérée comme satisfaisante si elle correspond à la finalité et au contexte de l’évaluation. Cette étude comprend un ensemble de recommandations pour choisir entre les principales options méthodologiques en fonction du contexte (voir 5.2.2).
L’équipe d’étude s’est penchée sur quatre analyses d’impact considérées comme des succès et a abouti sur des résultats mitigés. Certes les analyses d’impact ont permis d’apporter des enseignements nouveaux, crédibles et pertinents, mais le réel apprentissage à partir de ces leçons a été assez limité. Cela s’explique par le fait que:

- La volonté de conduire des évaluations d’impact venait davantage des services d’étude que des décideurs politiques. Les bonnes pratiques d’évaluations ont surtout été liées à l’existence d’une capacité administrative, d’un bon système de suivi, de bases de données statistiques, et d’une forte institutionnalisation de l’évaluation. Le besoin d’acquérir une connaissance solide n’était pas la raison principale de la mise en œuvre de ces évaluations. **Peu d’attention a été portée à faire rencontrer le temps de l’évaluation et celui de la politique.**
- La présentation des résultats aux décideurs politiques a peu contribué à l’utilisation de ceux-ci, surtout parce que beaucoup d’attention était accordée aux questions techniques aux dépens de la substance des conclusions.
- Les nouvelles connaissances acquises ont souvent été acheminées aux décideurs politiques par des chemins indirects, en s’inscrivant dans une **succession d’évaluations** et d’études avant d’être utilisées par les décideurs.
Les décideurs politiques avaient une capacité limitée à absorber les résultats des évaluations, sauf dans le cas d'une proximité durable avec un groupe stable d'experts techniques ou administratifs.

En outre, cette partie de l'étude confirme que, même lors de circonstances favorables exceptionnelles, les contraintes de coûts et de faisabilité rendent impossible la mise en place d'analyses d'impact de façon systématique et/ou exhaustive. En d'autres termes, les impacts ne peuvent être connus partout et tout le temps.

L'équipe d'étude fait une série de propositions pour développer une culture d'apprentissage parmi les personnes participant à la gestion des FSE. Ces propositions s'appuient sur l'idée que l'apprentissage à partir des analyses d'impact devrait s'inscrire dans un processus permanent à l'échelle européenne, et ne plus se restreindre au périmètre et à la durée des programmes.

La mise en œuvre de ces recommandations représenterait un effort significatif de la part des autorités de gestion. Inévitablement, de tels efforts seraient perçus comme un fardeau supplémentaire, sauf si d'autres obligations sont réduites en contrepartie. Ceci pourrait être réalisé par :

- L’allègement des activités de gestion et d’audit par le choix d’approche telle que l’appui budgétaire utilisé dans le domaine de l’aide au développement ;
- L’allègement des demandes de rapports, particulièrement pour les indicateurs de suivi.

**Rendre les décideurs politiques responsables de l’apprentissage**

L'ensemble de réglementations s'appliquant au prochain cycle de programmation pourrait inclure des dispositions pour réduire l'écart entre l'évaluation, la preuve et la connaissance d'un côté et la décision politique de l'autre :

- Chaque année, les autorités de gestion devraient approuver un plan d'évaluation multi annuel. Sur une période de 5 ans, chaque plan devrait contribuer à combler, par une analyse d'impact de qualité, au moins un des points faibles identifiés dans les connaissances existantes.
- Au même moment, les autorités de gestion devraient approuver un rapport annuel de suivi identifiant les principales leçons tirées des évaluations précédentes, et les actions entreprises.
- Une synthèse des connaissances pertinentes (evidence base) devrait être incluse dans : (1) tous les rapports d'évaluation, et (2) toutes les décisions qui ont un impact financier important. Cela ferait écho à la pratique des études d’impact (impact assessment) de la Commission européenne.
- En mettant en œuvre les règles ci-dessus, les autorités de gestion devraient améliorer leur capacité d'apprentissage à partir des analyses d'impact et se rapprocher d'une politique publique fondée sur la preuve. Une partie du financement devrait être conditionnée à l'existence de tels progrès.

**Promouvoir la connaissance par des incitations financières**

Des fonds spécifiques devraient être consacrés à l'évaluation et alloués spécifiquement à :

- La conception et l’évaluation d’approches de formations innovantes ;
- La pratique d’analyses d’impact solides visant à combler les lacunes identifiées dans les connaissances existantes ;
• La coordination d'évaluations traitant des mêmes questions d'impact dans plusieurs régions ou plusieurs pays ;
• L'implication des patriciens dans des « communautés de connaissances » discutant les résultats d'études et d'évaluations, et réalisant des synthèses et des contrôles de qualité (réseaux, conférences, etc.)

Apporter un accompagnement technique
La Commission ou d'autres institutions compétentes avec le soutien de la Commission devraient créer des guides disponibles au public sur :
• Les lacunes identifiées dans les connaissances actuelles et les techniques d'analyses d'impact pertinentes pour des évaluations visant à combler ces lacunes ;
• L'accès à des bases de données et le traitement des questions de confidentialité ;
• L'évaluation de la qualité des analyses d'impact et les processus de revue par les pairs ;
• Les approches les plus efficaces pour traduire les analyses d'impact en connaissances utilisables.

Apprentissage mutuel sur l’usage de la preuve dans la conception des politiques
Lors des événements réunissant les autorités de gestion, une partie du temps devrait être consacrée à des échanges d'expérience sur l'usage de la preuve dans la conception des politiques.
Zusammenfassung


Die Studie umfasste die folgenden drei Aufgabenbereiche:

- Eine Synthese der verfügbaren Forschungsergebnisse, die sich auf die Rendite von Investitionen in Humankapital beziehen. Folgendes waren hierbei die Schwerpunkte: (1) die Auswirkungen der Schulungsprogramme, die auf Angestellte abzielen, und (2) die Auswirkungen von Programmen, die sich mit einem Wiedereinstieg in die Arbeit beschäftigen, und somit auf Arbeitssuchende abzielen. Die Meta-Analyse, die auf einer ersten umfangreichen Vorauswahl von referierten Publikationen basierte, wurde mit Hilfe einer Gruppe von sieben Experten aus der EU und den USA verbessert.


**Erarbeitetes Wissen**

Beruhend auf einer umfangreichen Literaturrecherche und der Hilfe eines Gremiums aus internationalen Experten zeigt die Studie, dass:

- Die Schulung von Erwerbstätigen gesellschaftlich profitabel, aber auf Unternehmensebene nicht ausreichend lohnend ist. Aus diesem Grund ist das Volumen der Privat-Investitionen zu gering.


- Es gibt erhebliche Wissenslücken in Bereichen, die fundamental für die Rechtfertigung von ESF-Investitionen in Humankapital sind. Beispiele hierfür sind die Auswirkung der Schulung von „gefaehrdeten“ Arbeitnehmern, oder die Tatsache, dass ein Schulungsteilnehmer auf Kosten von jemand anderem eine Arbeitsstelle finden könnte (Substitution).

**Schulung von Erwerbstätigen** (siehe Kasten 1)

Die Frage der Kurzzeit-Rendite von Investitionen in Schulungen ist im Bezug auf einen einzelnen Schulungsteilnehmer oder ein Unternehmen noch nicht eindeutig geklärt. Jedoch haben neue Studien eine positiven Einfluss auf die Produktivität auf Branchenebene aufgezeigt, wobei dieser Einfluss gross genug ist um behaupten zu können, dass Investitionen in Humankapital profitabel und mit anderen Investitionen konkurrierend sind.

Das Problem ist, dass die Firma, die fuer die Schulung zahlt, nur einen Teil des Nutzens erhaelt. Andere Unternehmen in der Branche und die Angestellten, die geschult werden, ziehen daraus möglicherweise den grössten Nutzen, vor allem dann, wenn letztere zu anderen Unternehmen wechseln. Dies wirft die Frage eines mangelnden Volumens an Investition seitens Arbeitgeber auf, bei denen der Anreiz in die Schulung ihrer Mitarbeiter zu investieren, nicht gross genug ist. Dass eher die Branche von den Vorteilen der Investitionen in Humankapital profitieren als einzelne Unternehmen, zieht die logische Grundlage für öffentliche Investitionen in Schulungssysteme und -strukturen auf Branchenebene mit sich.

Außerbetriebliche Schulung ist unter der Voraussetzung, dass die Auszubildenden die Möglichkeit haben, ihre neu erworbenen Fähigkeiten nach der Rückkehr zum Arbeitsplatz anzuwenden, bekanntlich wirksamer als betriebliche Schulung. Das verlangt nach einer engen Beziehung zwischen den Bereichen Schulung und Personalwesen.

**Schulung von Arbeitslosen** (siehe Kasten 3)

Die Auswirkungen der Schulungen, die auf den beruflichen Wiedereinstieg bezogen sind, sind erst dann mittelfristig positiv, wenn die negativen Lock-in-Effekte, d.h. die Tatsache, dass die Leute die Arbeitssuche beenden waehrend sie an einer Schulung teilnehmen, ausgeglichen wurden. Im Falle von jungen Arbeitssuchenden würde sich der Impact ziemlich in Grenzen halten.
Erhebliche Impacts werden darüber hinaus nur erzielt, wenn die Schulung mit Unterstützung bei der Arbeitssuche und / oder Rahmenbedingungen gekoppelt ist, die einen Ansporn für die Arbeitssuche schaffen. Dies erfordert eine engere Verbindung zwischen ESF Investitionen in Schulungen die sich auf den beruflichen Wiedereinstieg beziehen, und den Reformen der nationalen Arbeitsmarktpolitik.

**Wissenslücken** (siehe Kasten 2 und Kasten 4)

Es besteht ein gewisser Wissensmangel im Blick auf die gefährdeten Branchen und Unternehmen, für die Investitionen in Humankapital besonders wertvoll sein könnten, aber bei welchen ein Mangel an Ressourcen zu einem niedrigen Invesitionsvolumen führen kann.

Auch eingerostete Fähigkeiten und die Auswirkungen von Schulungen auf ältere Mitarbeiter sind Bereiche, über die mehr Wissen erlangt werden sollte. Dies ist besonders relevant angesichts der Entwicklung der alternden Wirtschaftssysteme in Europa.

Über den wirtschaftlichen Nutzen der Schulung von Arbeitslosen gibt es nur äußerst wenige Forschungsergebnisse, da: (1) Kosten und Nutzen gesondert untersucht werden, und (2) nur begrenzt Bemühungen gemacht wurden, die sozialen Vorteile, die aus der Integration resultieren, zu bewerten.

Es gibt nicht genügend Belege bezüglich des wichtigen Themas der Qualität der Ausbildung, und auch nicht darüber, wie die Qualität zu bewerten ist, oder ob die Vorteile, die aus einer besseren Qualität resultieren, die Kosten wert sind.

Ebenso ist nur wenig über das Verhaltensmustern der beschäftigungsspezifischen Auswirkungen im Bezug auf bestimmte Zielgruppen bekannt, außerdem, dass junge Arbeitssuchende generell als weniger erfolgreich in arbeitsmarktorientierter Schulung bewertet werden.

Es kann keine klare Schlussfolgerung über ein mögliches Verhaltensmuster der Effektivität von Schulungen bezüglich der Konjunktur gezogen werden. Es gibt einige Anzeichen dafür, dass der Effekt der Schulungen in dem Sinne prozyklisch ist, dass die Auswirkungen der Schulungen für den beruflichen Wiedereinstieg in Zeiten, in denen gute Beschäftigungsmöglichkeiten und niedrige Arbeitslosigkeit herrschen, höher sind. Jedoch ist dies unklar.

Das Thema der sekundären Auswirkungen bleibt letztendlich undurchsichtig. Hauptsächlich geht es um die Substitutionseffekte, d.h. die Tatsache, dass ein arbeitsloser Schulungsteilnehmer einen Arbeitsplatz auf Kosten von jemand anderem findet. Es wird davon ausgegangen, dass Substitution vermindert ist, wenn ein Schulungsprogramm nach einem Mangel an bestimmtem Fachkönnen gerichtet ist. Es wird auch davon ausgegangen, dass sich aus verbesserter Produktivität, Innovation, Konkurrenzfähigkeit, sowie hoher Marktanteile und sekundärer Auswirkungen auf Beschäftigung ein positiver Kreislauf entwickelt.
**Rendite öffentlicher Investitionen** (siehe 4.5.2)

Es gibt eine logische Grundlage für die Investitionen öffentlicher Gelder in Systeme und Strukturen auf Industriebene, doch diese Studie konnte die soziale Rendite solcher Investitionen nicht bewerten.

Aufbauend auf eine schwache Beweisgrundlage schließt die Studie darauf, dass die Schulung von Arbeitslosen unter strengen Auflagen gewinnbringend sein kann. Um öffentliche Investitionen in diesen Bereich zu rechtfertigen sollten die politischen Entscheidungsträger und Programm-Manager zeigen, dass sie die Faktoren kontrollieren, die eine positive Rendite durch die Investition in Schulungen gewährleisten. Das heisst, junge Leute nicht als Zielgruppe zu betrachten, Schulung und Unterstützung bei der Arbeitssuche zu verbinden, die Kosten der Schulungen im Rahmen zu halten, ausreichend hohe Qualität zu gewährleisten, und das Problem der Substitution ausreichend zu berücksichtigen.

Die Studie konnte bezüglich der sozialen Rendite öffentlicher Investitionen in die Schulung von Angestellten auf kein Ergebnis kommen. Natürlich herrscht eine logische Grundlage dafür, allgemeine Maßnahmen auf Branchenebene anstatt auf Unternehmensbasis zu ergreifen. Die Frage, wie solch eine Maßnahme strukturiert werden sollte, bleibt offen, so wie auch insbesondere die folgenden Teilfragen:

- Sollten allgemeine Maßnahmen auf Vorschriften (die Verpflichtung, Schulungen durchzuführen) oder auf öffentlichen Investitionen mittels finanzieller Anreize basieren?
- Unter welchen Annahmen wären öffentliche Investitionen sozial profitabel?
- Welche Ebene ist zuständig für die Gestaltung solcher Maßnahmen (die regionale, nationale, europäische)?
- Falls die zuständige Ebene nicht aus der gesamten EU besteht, sind verbleibende Maßnahmen auf EU-Ebene notwendig? ... Und falls ja, was sollten die Instrumente dafür sein („weicher“ Koordinierungsmethode oder „soft coordination“, finanzielle Anreize, oder beides)?

In Bezug auf die Schulung von Arbeitslosen schlägt die Studie vor, dass es einige vernünftige Annahmen gibt, denen nach öffentliche Investitionen als sozial gewinnbringend gelten. Diese Annahmen können wie folgt zusammengefasst werden:

- Die Erfolgsrate (erhöhte Wahrscheinlichkeit, eine Arbeitsstelle zu finden) sollte zwischen 10-20% liegen. Die geringste Zahl repräsentiert eine Rentabilitätsschwelle und die größte Zahl die optimistischste Annahme, die auf vernünftige Weise gemacht werden kann.
- Die Schulung sollte mit Unterstützung bei der Arbeitssuche kombiniert werden und die Dauer zwischen 2 bis 4 Monaten betragen. Andere Arten von Schulungen könnten auch profitabel sein, vorausgesetzt, dass ihre Kosten nicht die eines 2- bis 4-monatigen Programmes überschreiten.
Die größte Zahl repräsentiert eine Rentabilitätschwelle und die geringste Zahl eine sehr optimistische Annahme.

Soziale Vorteile sollten mindestens 50% der individuellen Vorzüge ausmachen (d.h. des Bruttogehalts der gefundenen Arbeitsstelle), wobei diese Rentabilitätschwelle könnte bis zu 100% oder mehr erhöht werden, wenn die anderen Annahmen pessimistisch sind.

Trotz des unsicheren Charakters der verfügbaren Beweise kommt das Studienteam zu dem Schluss, dass Schulung von Arbeitslosen unter strengen Auflagen profitabel sein könnte. Um öffentliche Investitionen in diesen Bereich zu rechtfertigen, sollten die politischen Entscheidungsträger und Programm-Manager zeigen, dass sie die Faktoren kontrollieren, die Schulungsprogramme sozial profitabel machen. Insbesondere sollten sie:

- Vermeiden, junge Arbeitslose als Zielgruppe zu sehen;
- Schulung mit Unterstützung bei der Arbeitssuche verbinden;
- Die Kosten der Schulung angemessen halten, d.h. entsprechend den Kosten eines 2- bis 4-monatigen Programms;
- Sicher stellen, dass die subventionierten Programme eine ausreichend hohe Qualität vorweisen, sodass davon ausgegangen werden kann, dass sie eine Erfolgsschwelle von 10% erreichen können;
- Sicher stellen, dass dem Problem der Substitution ausreichend Aufmerksamkeit geschenkt wird, sodass davon ausgegangen werden kann, dass sie unter 50% bleibt (Anteil der erfolgreichen Teilnehmer, die ihren Arbeitsplatz auf Kosten von jemand anderem finden).

Auch unter Anwendung solcher Grundsätze hängt die Rentabilität weiterhin von einigen Faktoren ab, die nicht kontrolliert werden können:

- Die Nutzen sollten mindestens 8 Jahre lang andauern, d.h. es sollte ausreichend Zeit verbleiben, bevor Altern, Verlust der Fähigkeiten, und Konjunkturzyklen die Arbeitsfähigkeit der erfolgreich Geschulten vermindern;
- Soziale Vorteile sollten mindestens 50% des individuellen Gewinns ausmachen (d.h. Bruttogehalt der gefundenen Arbeitsstelle).

Es fehlt ein erhebliches Wissen über die beiden oben genannten Umstände. Demzufolge wäre es schwierig, öffentliche Investitionen in Schulung von Arbeitslosen zu begründen, wenn nicht gleichzeitig Anstrengungen gemacht werden, diese Wissenslücken zu schliessen.
Aus den methodischen Tests lernen (siehe 5.2.1)

Basierend auf den wirkungsfeststellenden Evaluationen die in Zusammenarbeit mit vier Mitgliedstaaten gemacht wurden kommt die Studie zu dem Schluss, dass ausreichende Aufmerksamkeit, Fachfertigkeit und Ressourcen in Datenerhebungen und Impact-Analysen investiert werden sollte. Impact-Analysen sollten von Fall zu Fall neu gestaltet werden, da es keinen ultimativen Ansatz gibt.

Es dauert Jahre, aus den Impact-Analysen zu lernen. Dies bedeutet, dass die Impacts nicht in ein Leistungsüberwachungssystem einbezogen werden können.

Weiteres kann man aus Studien lernen, die sich auf spezifische Schulungsansätze, spezielle Zielgruppen und / oder bestimmte Regionen konzentrieren, und zwar vor allem, wenn solche Studien Regions- oder Länderübergreifend sind.

Aus mehreren Gründen lässt die Studie ernsthafte Zweifel daran aufkommen, ob es relevant ist, jährliche Zahlungen und Ergebnisse in Verbindung zu bringen, zumindest insofern die Leistungsindikatoren nicht mit gründlicheren Impact-Analysen ergänzt sind.


Die Studie wies erhebliche länderspezifische Unterschiede in Bezug auf Wirkungseinschätzungen auf. Solche Unterschiede erfordern eine sorgfältige Interpretation und könnten zum Beispiel dann Lernmöglichkeiten schaffen, wenn sie auf Veränderungen in der Qualität der Schulungen oder auf Unterschiede in den wirtschaftlichen Rahmenbedingungen zurückzuführen sind. Allerdings gibt es mindestens zwei Fälle, bei denen besonders hohe oder geringe Auswirkungen auf unzuverlässige Daten oder unangemessene Kontrollgruppen zurück geführt werden musste. Was daraus gelernt werden kann ist, dass die Datenerfassung und das Evaluationsdesign von Bedeutung sind, und dass daher genügend Aufmerksamkeit, Fachfertigkeit und Ressourcen in diese investiert werden sollte.

Je nach gewählter Analysemethode sollte die Evaluation bis zu einem Jahr vor der Schulung lanciert werden. Die Studie hat gezeigt, dass die Auswirkungen über einen Zeitraum von 1 bis 2 Jahren nach der Schulung analysiert werden sollten. Ein weiteres Jahr sollte für die Evaluation hinzugerechnet werden, um


Über die Lernprozesse lernen (siehe 7.1 und 7.3)

Die detaillierte Untersuchung vier erfolgreicher Impact-Analysen veranlasste die Studie dazu darauf zu schliessen, dass diese jeweils neue, zuverlässige und relevante Erkenntnisse lieferten, aber dass aus diesen Erkenntnissen nur begrenzt gelernt wurde.


Diese Vorschläge sind mit Verpflichtungen, Anreizen und Ratschlägen verbunden, die mit folgenden Aktivitäten zu tun haben:

- Systematisierung mehrjähriger Evaluationspläne und Weiterverfolgung der Evaluationen;
- Anwendung der EG-Verfahren der Impact-Einschätzung auf Programmebene;
- Das Lernen aus Impact-Analysen an sich zu einer finanziellen Bedingung machen;
- Earmarking finanzielle Mittel für Impact-Analysen um festgestellte Wissenslücken überbruecken;
- Förderung der Evaluationen der Impacts von Testschulungsansätzen;
- Verbreitung fachtechnischer Leitlinien für verschiedene Ansätze der Impact;


- Die Evaluationsergebnisse an politische Entscheidungsträger zu veröffentlichen trug nur sehr wenig dazu bei, solche Befunde tatsächlich aufzunehmen. Dies lag vor allem daran, dass den fachtechnischen Problemen auf Kosten von wesentlichen Aussagen zu viel Aufmerksamkeit gewidmet wurde.

- Die neu erworbenen Kenntnisse wurden politischen Entscheidungsträgern oft auf indirekte Weise nahegebracht, und zwar indem zahlreiche aufeinanderfolgende Evaluationen und Studien übermittelt wurden, bevor diese für die Politikgestaltung anwendungsreif waren.

- Politische Entscheidungsträger konnten Evaluationsergebnisse nur zu geringem Maße aufnehmen, ausser dann, wenn sie dauerhaft eng mit einer beständigen Gruppe von fachtechnischen- und Verwaltungsexperten zusammearbeitet haben.


Die Umsetzung dieser Empfehlungen würde weitere bedeutungsvolle Bemühungen Seiten der Verwaltungsbehörden ausdrücken. Unweigerlich würden solche Bemühungen als zusätzliche Belastung empfunden werden, sofern nicht gleichzeitig andere Verpflichtungen reduziert werden. Dies könnte durch die folgenden Maßnahmen erreicht werden:

- Verringerung der Management- und Audittätigkeiten mittels Ansätzen wie der Budgethilfe im Bereich der Entwicklungshilfe;

- Lockerung der Anforderungen für die Berichterstattung, vor allem im Bezug auf Kontroll-Indikatoren.
Politische Entscheidungsträger für das Lernen verantwortlich machen

Das Reglement für den nächsten Programmzyklus könnte einige Bestimmungen beinhalten, die die Kluft zwischen Evaluation, Beweisen und Wissen einerseits, und politischen Entscheidungen andererseits überbrücken sollten:

- Jedes Jahr zur gleichen Zeit sollten Verwaltungsbehörden einen jährlichen Folgebericht genehmigen, der die wichtigsten Erkenntnisse vorangegangener Evaluationen und die unternommenen Maßnahmen identifiziert.
- Eine Synthese von relevantem Wissen (Beweisgrundlage) sollte in (1) alle Evaluationsberichte und (2) alle Beschlüsse, die erhebliche finanzielle Auswirkungen haben, einbezogen werden. Dies würde die Verfahrensweise der Impact-Einschätzung der Europäischen Kommission widerspiegeln.
- Durch die Umsetzung der drei oben genannten Regeln sollten Verwaltungsbehörden Fortschritte darin zeigen, aus Impact-Analysen zu lernen und sich zur evidenzbasierten Politik hin zu bewegen. Ein Teil der Auszahlungen sollte von der Demonstration solcher Fortschritte abhängig gemacht werden.

Wissen durch Anreize handhaben

Spezielle finanzielle Mittel sollten für die Evaluation vorgesehen, und aus schliesslich den folgenden Aktivitäten zugeordnet werden:

- Planung und Evaluierung von Testschulungsansätzen;
- Durchführung von fundierten Impact-Analysen zur Überbrückung einer der identifizierten Wissenslücken;
- Auswertungen, die sich jeweils mit der gleichen Ursache-Wirkungsfrage in mehreren Regionen oder Ländern beschäftigen;
- Synthesen, Qualitätsbewertungen, Peer Reviews und Besprechung von Forschungs- und Evaluationsergebnissen innerhalb Wissensgemeinschaften (z.B. Netzwerken, Konferenzen) unter Miteinbeziehung von Fachleuten.

Bereitstellung fachtechnischer Beratung

Die Kommission oder ein entsprechendes von der Kommission unterstütztes Gremium sollte öffentlich zugängliche Leitlinien für das Folgende erstellen:

- Identifizierte Wissenslücken und fundierte Impact-Analysetechniken, die für Evaluationen zur Überbrückung solcher Lücken ausgerichtet sind;
- Zugriff auf Datenbanken und den Umgang mit Datenschutzfragen;
- Bewertung der Qualität der Impact-Analysen und die Auseinandersetzung mit Peer Reviews;
- Effektive Ansätze zur Umwandlung von Impact-Analysen in verwendbares Wissen.

Gegenseitiges Lernen über evidenzbasierte Politik
In den Versammlungen von Verwaltungsbehörden sollte genügend Zeit dafür aufgebracht werden, voneinander über evidenzbasierte Politikgestaltung zu lernen.
## Contents

1 Introduction .......................................................................................................................... 1

1.1 Purpose and scope of the study ......................................................................................... 1

1.1.1 Feeding into the design of the next ESF cycle ......................................................... 1

1.1.2 ESF-like investments in human capital ......................................................................... 1

1.2 Structure of this report ..................................................................................................... 4

2 Main lines of the study method .......................................................................................... 5

2.1 Questions asked and overall approach ........................................................................... 5

2.2 Tasks .................................................................................................................................. 5

2.3 Methodological limitations ............................................................................................... 7

3 Knowledge and knowledge gaps ......................................................................................... 8

3.1 Training of employees: knowledge and knowledge gaps ............................................ 8

3.1.1 An earlier review of company training literature .................................................... 8

3.1.2 Further results from studies of training of employees ............................................. 11

3.1.3 Concluding remarks ................................................................................................. 14

3.2 Training of the unemployed: knowledge and knowledge gaps ................................... 16

3.2.1 Kluve’s meta-study (2006) ...................................................................................... 16

3.2.2 Recent micro studies ................................................................................................. 18

3.2.3 Concluding remarks ................................................................................................. 26

3.3 Persons at risk of losing their jobs: another knowledge gap ....................................... 28

3.4 Macroeconomic impact estimates ................................................................................... 28

3.5 Views of the expert panel ................................................................................................. 30

4 Lessons from four country studies .................................................................................... 32

4.1 Impact of training: Hungarian country study ................................................................. 32

4.1.1 Dataset (HU) ........................................................................................................... 32

4.1.2 Surveyed participants (HU) .................................................................................... 32

4.1.3 Comparative analysis (HU) ..................................................................................... 36

4.1.4 Lessons (HU) ........................................................................................................ 41

4.2 Impact of training: Italian country study ......................................................................... 41

4.2.1 Dataset (IT) .............................................................................................................. 41

4.2.2 Surveyed participants (IT) ...................................................................................... 42

4.2.3 Comparative analysis (IT) ...................................................................................... 47

4.2.4 Lessons (IT) .......................................................................................................... 50

4.3 Impact of training: Polish country study .......................................................................... 51

4.3.1 Dataset (PL) .......................................................................................................... 51
4.3.2 Surveyed participants (PL) ................................................ 52
4.3.3 Comparative analysis (PL) ................................................ 56
4.3.4 Lessons (PL) ................................................................... 60
4.4 Impact of training: Belgian country study ......................... 60
4.4.1 Dataset (BE) ................................................................... 60
4.4.2 Methodological issues (BE) ................................................ 61
4.4.3 Descriptive analysis (BE) .................................................. 65
4.4.4 Comparative analysis (BE) ................................................ 67
4.4.5 Impact on employment status (Method III) ......................... 71
4.4.6 Lessons (BE) ................................................................... 73
4.5 Impact of training: lessons learned across four countries ...... 73
4.5.1 Synthesis of the country studies ........................................ 73
4.5.2 Concluding on the return on investment ......................... 74
5 Methodological issues ............................................................. 81
5.1 Return on investment ............................................................. 81
5.1.1 Financial and social return ................................................ 81
5.1.2 Costs and Benefits of human capital investments ............. 82
5.1.3 Rationales and determinants of human capital investments ... 84
5.2 Impact analysis ................................................................. 85
5.2.1 Methodological findings .................................................... 85
5.2.2 Managing methodological choices ................................... 87
5.3 Impact analysis techniques .................................................. 90
5.3.1 Duration model ............................................................... 90
5.3.2 Experimental design ........................................................ 91
5.3.3 Matching techniques ....................................................... 92
5.3.4 Instrumental variable ....................................................... 93
5.3.5 Regression discontinuity ............................................... 93
5.3.6 Fixed effects ................................................................. 93
5.4 Historical perspective .......................................................... 94
6 Learning from impact evaluations .......................................... 95
6.1 Analysing four impact evaluations ...................................... 95
6.2 Evaluation mandate (testing Assumption 1) ......................... 96
6.3 Evaluation method (testing Assumption 2) .......................... 98
6.4 New knowledge (testing Assumption 3) .............................. 101
6.5 Dissemination (testing Assumption 4) .................................. 102
6.6 Accumulation of knowledge (testing Assumption 5) ............. 104
6.7 Knowledge use (testing Assumption 6) ............................... 105
7 Recommendations .................................................................. 107
7.1 Findings from the four case studies ........................................... 107
7.2 Other problems identified in the study ........................................ 108
7.3 Suggestions for new rules, incentives and guidance ..................... 108
Appendix 1 – Study method .................................................................. 115
Appendix 2 – References of the literature review ............................... 118
Appendix 3 – Members of the expert panel ......................................... 123
Appendix 4 – Expert panel report ....................................................... 127
Appendix 5 – Learning from impact evaluation: Belgian case study ...... 144
Appendix 6 – Learning from impact evaluation: Hungarian case study ..... 150
Appendix 7 – Learning from impact evaluation: Italian case study ...... 155
Appendix 8 – Learning from impact evaluation: Polish case study ........ 167
Appendix 9 – Return on investments ................................................... 175
Appendix 10– Focus group ................................................................. 179
Appendix 11 – Vocabulary of the study ............................................... 186

**Tables**

Table 1 – Contribution of tasks to answering the questions .................. 5
Table 2 - Impact estimates in Card et al. (2009) ................................. 19
Table 3 - Estimates of the employment effects of training of a short and long duration ................................................................. 21
Table 4 - Profile of surveyed participants (HU) ................................... 33
Table 5 - Surveyed participants per type of activity (HU) ...................... 34
Table 6 - Surveyed participants and economic context (HU) ................. 35
Table 7 – Status of participants after 24 months (HU) ......................... 36
Table 8 – Impact on employment status after training (HU/Method 0) 37
Table 9 – Profile of participants and control group (HU) ..................... 38
Table 10 – Factors affecting participation (HU) .................................. 39
Table 11 – Impact on employment (HU/Method I) .............................. 40
Table 12 – Impact on employment per public (IT/Method I) ................. 40
Table 13 – Surveyed participants per type of training (IT) .................... 43
Table 14 – Participants and duration of training programme (IT) ........... 43
Table 15 – Participants and type of validation (IT) .............................. 43
Table 16 – Profile of surveyed participants (IT) .................................. 44
Table 17 – Surveyed participants per region (IT) ................................ 45
Table 18 – Duration of unemployment before training (IT) .................. 46
Table 19 – Employment status after training (IT) .............................. 46
Table 20– Impact on employment (IT/Method 0) .................................. 47
Table 21 – Factors affecting participation (IT) ........................................... 48
Table 22 – Impact on employment (IT/Method I) ....................................... 49
Table 23 – Impact on employment per target group (IT/Method I) .......... 49
Table 24 – Separate estimates per region (IT/Method I R) ....................... 50
Table 25 – Separate estimates per type of training (IT/Method I T) .......... 50
Table 26 – Participants per type of training activity (PL) ......................... 53
Table 27 – Participants and duration of training (PL) ............................. 53
Table 28 – Surveyed participants per public (PL) .................................... 54
Table 29 – Employment status after training (PL) ................................. 55
Table 30 – Improvements stated by participants (PL) ............................. 56
Table 31 – Impact on employment (PL/Method 0) ................................. 57
Table 32 – Job stability of employed interviewees (PL) .......................... 58
Table 33 – Job duration of employed interviewees (PL) .......................... 58
Table 34 – Impact on employment (PL/Method I) ................................. 59
Table 35 – Impact on employment per public (PL/Method I) ................. 60
Table 36 – Estimating impact with four methods (BE) ............................ 69
Table 37 – Factors affecting unemployment duration (BE) .................... 70
Table 38 – Impact on employment (BE/Method III) .............................. 71
Table 39 – Impact indicators vs impact estimates ................................. 90

Figures

Figure 1 – ESF investment in human capital (narrow definition) and other policy areas ................................................................. 3
Figure 2 – ESF investment per activity .................................................. 3
Figure 3 – Duration of prior unemployment (BE/Method I - unmatched) .... 63
Figure 4 – Duration of prior unemployment (BE/Method I - matched) ...... 64
Figure 5 – Duration of prior unemployment (BE/Method II - unmatched) ... 64
Figure 6 – Duration of prior unemployment (BE/Method II - matched) .... 65
Figure 7 – Duration of unemployment and gender (BE) ....................... 66
Figure 8 – Duration of unemployment and level of education (BE) ........ 66
Figure 9 – Duration of unemployment and age (BE) ............................. 67
Figure 10 – Probability of remaining unemployed (BE / unmatched) ........ 68
Figure 11 – Probability of remaining unemployed (BE – Method I, matched) 68
Figure 12 – Probability of remaining unemployed (BE / Method II, matched) 69
Figure 13 – Conditions for achieving good returns (first scenarios) ......... 76
Figure 14 – Conditions for achieving good returns (second scenarios) .... 77
Figure 15 – Conditions for achieving good returns (third scenarios) ....... 78
**Boxes**

Box 1  – Investing in the employed – the lessons learned ......................... 15  
Box 2  – Investing in the employed – knowledge gaps ............................. 15  
Box 3  – Investing in the unemployed – the lessons learned ........................ 28  
Box 4  – Investing in the unemployed – knowledge gaps ........................... 28  
Box 5  – Assessing reliability (IT) ........................................................... 51  
Box 6  – Probability of ending unemployment (BE) ................................. 62

**Acronyms**

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALMP</td>
<td>Active Labour Market Policy</td>
</tr>
<tr>
<td>CATI</td>
<td>Computer Assisted Telephone Interviewing</td>
</tr>
<tr>
<td>CGE</td>
<td>Computable general equilibrium (model)s or overlapping generations (OLG) models</td>
</tr>
<tr>
<td>EES</td>
<td>European Employment Strategy</td>
</tr>
<tr>
<td>ESF</td>
<td>European Social Fund</td>
</tr>
<tr>
<td>ISFOL</td>
<td>Istituto per lo Sviluppo della Formazione Professionale dei Lavoratori</td>
</tr>
<tr>
<td>ISTAT</td>
<td>Istituto Nazionale di Statistica</td>
</tr>
<tr>
<td>NPV</td>
<td>Net present value</td>
</tr>
<tr>
<td>OLG</td>
<td>Overlapping generations (models)</td>
</tr>
<tr>
<td>PSM</td>
<td>Propensity score matching</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised Controlled Trial</td>
</tr>
<tr>
<td>ROI</td>
<td>Return on Investment</td>
</tr>
</tbody>
</table>
Introduction

1.1 Purpose and scope of the study

1.1.1 Feeding into the design of the next ESF cycle

This study is intended:

- To inform future EU policy-making discussions on the types of human capital investment which are known to generate good returns and might therefore deserve to be prioritised within the next programming cycle of the European Social Fund (ESF), scheduled for 2014-2020;
- To suggest an EU-wide learning strategy in order to achieve more and faster feedback from good/poor return on ESF and ESF-like investments;
- To recommend evaluation and monitoring arrangements so that relevant, rapid and valid lessons can be added to existing knowledge on human capital investments and their impacts;
- To recommend analytical methods so that valid and transferable lessons can be learned from the evaluation of the impacts of human capital investments.

Inter alia, the conclusions of the study may feed into:

- The Fifth Cohesion Report to be issued in autumn 2010, and which will include an outline of the Commission’s policy proposals;
- The impact assessment of the next cycle of the EU social policy.

1.1.2 ESF-like investments in human capital

This section defines the scope of the study, and comments upon its relevance for the ESF.

Defining human capital investment

In this study, the following activities are identified as human capital investments (directly or indirectly):

- Education and training at any stage of a person’s life, targeted at employees (on- and off- the job), entrepreneurs, job-seekers and/or inactive people;
- Internships and temporary work placement associated with education or training;
- Building / strengthening the capacity of education and training systems and structures.

The following activities are not considered as human capital investments:

- Individual counselling and coaching of job-seekers and inactive people, in so far as they are not associated with education or training;
- Assistance for mobility and changing jobs;
- Subsidised jobs;
- Information, communications and awareness campaigns, e.g. in the areas of gender equality, social inclusion, or occupational safety and health;
- Building / strengthening the capacity of institutions in charge of managing the labour market.
Overall ESF expenditures

ESF-supported interventions potentially target almost all European people (employed, unemployed, and inactive) but there is an emphasis on the long-term unemployed, low-skilled workers and other employed people at risk of losing their jobs, young and elderly inactive people, and vulnerable social groups. Across all targeted publics the issue of gender is given constant attention.

The targeted publics are reached by:

- Co-funding national or sub-national programmes or schemes which share the same goals as that of the ESF (direct approach);
- Developing the capacity of national or sub-national systems and structures (education and training institutions, employment services) in order to attract them towards ESF targets and goals, and to enable them to develop adequate services on their own (indirect approach);
- Encouraging / enabling national or sub-national governments and social partners to develop / to reform policies and strategies in line with the objectives of the EU in the area of employment (double indirect approach or soft coordination).

At micro-level, the intended impacts can be summarised as follows: workers and enterprises are adaptable, new enterprises are created, unemployed people find jobs, people have longer working lives, and vulnerable people are socially included through their participation in the labour market.

At macro-level it is expected that unemployment remains limited throughout economic cycles and that a high proportion of people participate in the labour market, including vulnerable social groups.

ESF investments in human capital

The definition of human capital in the ESF documents is considerably narrower than the one above. ESF categories define ‘human capital investment’ as education essentially. In terms of this narrow definition, one third of ESF resources are invested in human capital.

This is visible in the figure below which is extracted from an on-going study1 covering all 117 ESF supported programmes in 27 Member States. The figure shows the planned allocation of EU funds (2007-2013) per policy area.

---

1 ESF implementation in the period 2007-2013 - Template for EC reporting and 2008 outline report. The figures are extracted from the interim report of this study, dated March 2010, and will need to be updated.
**Figure 1 – ESF investment in human capital (narrow definition) and other policy areas**

<table>
<thead>
<tr>
<th>Area</th>
<th>Allocation (€ m)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptability</td>
<td>11275</td>
<td>18%</td>
</tr>
<tr>
<td>Access to labour market</td>
<td>17729</td>
<td>29%</td>
</tr>
<tr>
<td>Social inclusion</td>
<td>7652</td>
<td>13%</td>
</tr>
<tr>
<td>Others</td>
<td>4556</td>
<td>7%</td>
</tr>
<tr>
<td>Human capital</td>
<td>20313</td>
<td>33%</td>
</tr>
</tbody>
</table>

The next figure displays the overall allocation of EU funds per activity, and suggests that the weight of human capital, as defined in this study, is considerably heavier.

**Figure 2 – ESF investment per activity**

<table>
<thead>
<tr>
<th>Area</th>
<th>Allocation (€ m)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placement &amp; subsidised jobs</td>
<td>11275</td>
<td>16%</td>
</tr>
<tr>
<td>Awareness raising</td>
<td>17729</td>
<td>26%</td>
</tr>
<tr>
<td>Counselling &amp; advice</td>
<td>7652</td>
<td>11%</td>
</tr>
<tr>
<td>Others (incl. training systems)</td>
<td>11594</td>
<td>17%</td>
</tr>
<tr>
<td>Education &amp; training</td>
<td>20313</td>
<td>30%</td>
</tr>
</tbody>
</table>

In the above figure, the largest share of funds (30%) is allocated to ‘education and training’. Other activities are to be considered at least partially as
human capital investments, such as counselling and support to training and education systems. Overall, it can be assumed that human capital investment account for about 50% of the funds if a wide definition is adopted.

**Scope of the study**

The study is intended to identify and validate lessons that are both relevant and transferable to future ESF support in the area of human capital, including lessons that have been learnt outside the EU or at national or sub-national level within the EU. In this respect, the scope of the study is delineated by the nature of human capital investments supported by the ESF.

As far as learning and evaluation are concerned, the study focuses on the ESF only. The recommendations should be applicable by and relevant to the needs of managing authorities across the 27 Member States, even if they build upon an in-depth investigation in a limited number of countries.

1.2 **Structure of this report**

The next chapters address the following points:

- Knowledge and knowledge gaps in the literature;
- Knowledge gained from four country studies;
- Evaluation methods;
- Learning from impact evaluations;
- Overall conclusions and recommendations.

The appendices are presented in an attached CDROM. They include:

- Study questions, process and method;
- List of reviewed papers;
- Report of the expert panel;
- Investigation into four cases of potential learning;
- Report of the focus group meeting;
- Slides to be used for presenting the study.
2 Main lines of the study method

2.1 Questions asked and overall approach

This study addresses the three questions below which are considered in more detail in Appendix 1:

- **Return on human capital investments**: Which types of human capital investments present the highest return for the individual participant or firm involved and for society at large?
- **Available evaluation methods**: What methods are available to estimate the impacts of human capital investments and to what extent are they applicable to ESF-supported interventions?
- **Learning from evaluation**: How could impact evaluation and dissemination of learning be incorporated in the design of the ESF-supported interventions from their outset?

The study was implemented through nine tasks (or work packages) as follows:

- Literature review and research synthesis;
- Expert panel focusing on research synthesis;
- Four country studies in Belgium (Flanders), Italy (Northern regions), Hungary, and Poland;
- Cross-country analysis focusing on return on human capital investments;
- Cross-country analysis focusing on learning from impact evaluations;
- Focus group on learning issues.

The following table shows how the various tasks contributed to addressing the three study questions.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Literature review and research synthesis</td>
<td>XX</td>
<td>XXX</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>2. Expert panel</td>
<td>XX</td>
<td>XX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 - 6. Country studies</td>
<td>XX</td>
<td>XX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Cross-country analysis - impacts</td>
<td>X</td>
<td>XXX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Cross-country analysis - learning</td>
<td></td>
<td>X</td>
<td>XXX</td>
<td></td>
</tr>
<tr>
<td>9. Focus group</td>
<td></td>
<td>X</td>
<td>XXX</td>
<td></td>
</tr>
</tbody>
</table>

2.2 Tasks

This section briefly introduces the nine tasks, which are further described and commented upon in Appendix 1.
**Literature review and research synthesis (Task 1)**

Through this task, the study team accessed, assessed, compiled, and synthesised 80 studies, research works and evaluations pertaining to the benefits of and economic return on ESF-like human capital investments. The purpose was to extract valid and relevant knowledge and to identify knowledge gaps.

The literature review was also an opportunity to draw a broad picture of the methodological approaches in use in the analysis of ESF-like activities.

A draft synthesis report was first produced and then submitted to an expert panel (see next task) in the form of a working document. The study team complemented the investigations and analyses on the basis of the recommendations of the expert panel.

The main output of this task is presented in Chapter 3 (Knowledge and Knowledge Gaps). The literature review also served as an input for Chapter 5 (Methodological Issues).

**Expert panel (Task 2)**

A panel of seven experts (see list of members in Appendix 3) was convened for a one-day meeting in Brussels and reflected on the basis of the draft research synthesis (see previous task).

The experts complemented their inputs through two rounds of comments on successive versions of the report. They made a number of useful recommendations for strengthening the literature review, delineating the transferable knowledge (“lessons learned”), and identifying knowledge gaps.

Appendix 4 presents the report of the expert panel. An excerpt is presented in 3.5.

**Country studies (Tasks 3 to 6)**

The study team estimated the impact of ESF-funded human capital investments through four country studies applying several analysis techniques.

The selected Member States had already carried out impact evaluations involving control groups during the previous years. Concerned Managing Authorities participated in the work on a voluntary basis and kindly provided the study team with the datasets that had been used in the previous evaluations. Some datasets were completed in the context of this study.

The four impact analyses are presented in Chapter 4. Two of them (Italy and Poland) are almost a replication of the previous national studies. In contrast, the analyses implemented in Belgium and Hungary provide new findings. On the basis of its impact estimates, the study team identified a series of lessons that might contribute to the accumulation of knowledge.

The country studies included a series of interviews and documentary reviews aimed at:

- Describing the human capital investments under study;
- Identifying the contextual factors which may affect the transferability of the lessons learned;
- Understanding the process through which the data had been collected and assessing the potential biases resulting from this process;
- Analysing the process through which the previous impact studies had been implemented and used.

The findings from these interviews and documentary reviews are presented in Appendices 5 to 8.
Cross-country analysis of impacts (Task 7)

Once finalised, the four impact analyses were subjected to a cross-section review which:

- Compared the lessons learned and reflected on their transferability;
- Confronted the lessons learned with the current state of knowledge as assessed in Chapter 3;
- Reflected on the pros and cons of the various methods applied.

In addition, the study team designed a series of scenarios in order to delineate the assumptions under which human capital investments are socially profitable.

The substantial findings of this task are presented at the end of Chapter 4 and in Appendix 9. Methodological findings are integrated in Chapter 5.

Cross-country analysis of learning (Task 8)

This task consisted of a cross-cutting analysis of the investigations carried out in the four countries, with a special focus on learning from impact estimates. Provisional lessons were drawn about the learning process, and draft recommendations were proposed in order to strengthen this process within the next years.

The recommendations were discussed and significantly improved after being presented in a focus group meeting (next task) and in the conference “Shaping the future of the ESF” held in Brussels in June 2010.

The output of this task is described in Chapter 6 of this report (Learning from Impact Evaluations).

Focus group (Task 9)

A focus group was convened in order to discuss the issue of learning from impact evaluations. Invited participants belonged to managing authorities and the Commission.

The report of the meeting is displayed in Appendix 10. A summary is presented at the end of Chapter 6.

2.3 Methodological limitations

There is no doubt that all the findings of this study have been affected by methodological limitations. However, these limitations have been the subject of several discussions involving various stakeholders and experts:

- The literature review and the research synthesis were discussed with the expert panel;
- The country studies were discussed with and checked by the concerned managing authorities in a bilateral way;
- The findings related to learning were also discussed with the concerned managing authorities, including a focus group deliberation.

Overall, this confirmation process has led to a distinction between what is known and what is not in a relatively consensual way, and the recommendations of the study team build upon this relatively sound basis.

The most risky part of this study is Section 4.5.2, in which the rationale for spending EU money in the ESF is assessed through a series of scenarios which were developed at a late stage in the study, with limited interaction. Should the corresponding recommendations be used for policy-making purposes,
there would be a need for a few rounds of discussions, comments and improvements.

3 Knowledge and knowledge gaps

This chapter results from the literature review and research synthesis (Task 1). In the first version it was discussed by the expert panel (Task 2). This version takes stock of the experts’ comments.

The focus on ESF-like human capital investments implies that attention is paid to human capital investments in terms of training people in (or potentially in) the labour force rather than other efforts to improve employment, such as general education and active labour market policy actions. Other types of human capital investments, like formal education, which obviously involves much larger investments, are therefore omitted from the study.

It is important to realize that ESF-like human capital investments account for only a relatively small proportion of the total of such investments, and cover only a marginal share of what EU Member States invest in this area.

3.1 Training of employees: knowledge and knowledge gaps

In general, human capital investments targeted at employed persons are understood as skills development through training.

Since the early work of Becker (1962) and Mincer (1958, 1958), a comprehensive set of microeconomic studies have been conducted with a view to assessing the returns to training and, in particular, to firm-sponsored training. The objective in these studies is primarily to clarify the employers' and the employees' gains from training, in terms of productivity and wage increases, respectively. With very few exceptions, none of them have actually estimated the returns on this type of human capital investment.

Many studies report substantial returns to training. However, the magnitude of the returns depends on a broad range of factors such as the type of training provided as well as the characteristics of the firm and the participants. The impact analysis method might also affect the findings. The fundamental problem lies in estimating the effect of training while controlling the selection bias (see 5.2.1). The following sections summarise the substantial findings of available studies. Methodological issues are addressed in Chapter 5.

Contrary to the microeconomic studies, the macroeconomic research is fairly limited. However, by incorporating human capital production in, for instance, computable general equilibrium (CGE) models or overlapping generations (OLG) models, some studies have analysed the effects of training on economic growth. The issue of macroeconomic impacts is discussed in 3.4.

3.1.1 An earlier review of company training literature

Rita Asplund (2005) reviewed the empirical research on economic benefits of company-provided training. The scope was relatively wide and the following main questions were addressed:

1. What kind of training do companies provide? (general and specific)
2. Who is trained, how intensively and for what reason?
3. Does training enhance employability, mobility, and promotion?
4. Do trained employees experience relative income advantages?
5. Do companies benefit from these investments?
6. Does society at large benefit from company-provided training?

Asplund (2005) concluded that “our current knowledge on the economic role of company-provided training demands great cautiousness in drawing policy implications concerning the crucial questions of inequality of access to company training and training market failures, among others”.

The main answers of Asplund (2005) to the mentioned questions are described in the following.

**General or specific training**

Contrary to the predictions of standard theories, the review concludes that companies do provide and pay for general training, and not only for company-specific training. Asplund refers to Loewenstein and Spletzer (1998, 1999) who found that, in the US, a large proportion of employer-paid training is essentially general. She also draws on data from the International Adult Literacy Survey (IALS) and from the European Community Household Panel (ECHP), which indicate that European companies pay for training that is generating general skills. Loewenstein and Spletzer (1999) contend that training can be considered general not because the acquired skills may be useful elsewhere, but because alternative employers recognise that the training raises the employees’ productivity elsewhere.

Labour market imperfections restricting workers’ mobility and reducing the risk of losing staff after providing training of a general nature are given as the main reasons for this.

**Who is trained, how and why?**

This question concerns differences between small and large companies and between employees with different levels of education. Asplund (2005) refers to the “widely established contention, supported by aggregate statistics”, that the probability of receiving employer-provided training increases substantially with the individual employee’s level of education. However, reference is made to OECD (2003), where the lower incidence and intensity of training is explained by demand rather than variations in supply. It is thus possible that the supply of training shows no variation with the educational level of employees, but that the less-educated do not demand training to the same extent as the better educated.

Asplund (2005) also concludes from the review that “larger companies are much more likely to provide their employees with formal training than are smaller companies” and the extent of training also tends to be positively related to the size of the employer.

A number of other parameters have been discussed on the basis of earlier research, but the picture was not considered sufficiently convincing, and no clear conclusions were drawn concerning occupation, age, gender, and union membership, for example.

**Mobility effects**

The review does not give any clear conclusions with regard to the correlation between training and mobility. Most evidence pointed to a negative correlation, but it is important to be aware that the effects may go both ways. Mobility may affect training, and training may affect the mobility of the trained person as well. Arguments are put forward for both negative and positive correlations between training and mobility. Some of the main arguments are:
• In particular, specific training tends to increase the commitment of employees and thus leads to less mobility and a negative correlation\(^2\).
• Employers are inclined to train the employees they wish to retain, which will also lead to a negative correlation\(^3\).
• As benefits from training are expected to be highest for employees who move to another firm, training may increase an employee’s incentive to and probability of leaving the employer and thus lead to a positive correlation\(^4\).

The latter argument is the background for the theories of companies’ underinvestment in training.

**Wage effects**

Rita Asplund (2005) refers to a large number of studies that report participation in employer-provided training as having a significantly positive impact on the wages of the participants. Methodological problems however generate biases in many previous studies. As a consequence, the resulting picture, according to the review, is a bit blurred, so that no clear conclusion can be drawn.

Findings from the UK, USA, and Switzerland\(^5\) are referred to, which indicate that company provided training results in negligible returns in terms of wages from current employers but in higher returns when the employees change jobs.

No clear and convincing picture is drawn by the review concerning the differences in the impacts of training of persons with different characteristics, but a number of sources are referred to concerning impacts of external versus internal training. The conclusion is that external training has a significant positive wage impact while the wage impact of internal training is insignificant. According to studies that are referred to by Asplund (2005) from Australia, the USA and the UK, an exception to this general picture is youth training where internal training offered larger returns than external training.

**Productivity effects**

Productivity is more difficult to measure, but it is the key parameter for understanding the value of training to the companies that pay for it and to society as a whole. Asplund (2005) refers to the focus that has been put in the literature on the wage effects of training, but points to a number of research works from the UK, Germany, Sweden and the US that are indicating significant positive productivity effects. Productivity is typically measured by means of standard Cobb-Douglas production functions.

Asplund (2005) mentions the study by Dearden, Reed and Van Reenen (2000) which focuses on the effects of training on productivity at industry level. The study uses data on British industries, where the estimated productivity effect of training is relatively large. This may be due to the inclusion of positive externalities that can be expected to arise from company-provided training.

According to the review, the size of the productivity effect varies, but existing evidence points to a significantly positive impact from external or general training.

\(^2\) Green et al. 2000
\(^3\) Dearden et al. 1997, and Frazis et al., 2000
\(^4\) Frazis et al, 2000
training, while the productivity effects of internal or specific training are typically negligible. It is suggested that a possible reason for this may be that formal company training has a more lasting, long-term productivity impact.

The review also refers to an older article by Black and Lynch (1996) that found certain aspects of training to be important for productivity impacts. They analysed the impact of training on sales in private US organisations with more than 20 employees. When using the number of staff involved in training as the criterion, they found no productivity impact. When however they accounted for other dimensions of training, such as the proportion of time spent in formal off-the-job training and specific contents of training, they found strong productivity effects.

Economic effects on society

The review by Asplund (2005) refers to a number of research works that are in line with the standard economic theory, according to which employers under-invest in training (and do not provide general training at all). The reason is the risk of competitors hiring away the trained employees and the resulting, implicit higher discount rates or the shorter timeline applied by companies as compared to society as a whole. As a result, many researchers have recommended public subsidies for training and especially for general training, in order to counterbalance this bias.

On the other hand, Moen and Rosen (2004), who are referred to in the review, concluded that under-investment is due to internal inefficiency. They suggested that internal efficiency may be achieved (as an alternative to subsidising) through long-term binding contracts or efficient bargaining between employers and employees. The problem is that employers do not set wages for experienced employees at a competitive level. Along the same lines, Leuven et al. (2005) argued that workers should be motivated by reciprocity to achieve an optimal investment in both general and specific company training.

Asplund (2005) concludes that a multitude of questions are left unanswered concerning the existence and magnitude of market failures in the area of training and their causes and consequences. It is further concluded that “there does not seem to exist a common solution for all countries and all situations. Instead, public interventions should be tailored, in a cost-efficient way, to fit the specific needs that exist in each country”.

3.1.2 Further results from studies of training of employees

Since Asplund's literature review in 2005, a number of studies have further contributed to the analysis of the effects of human capital investments in terms of employee training.

These more recent studies have generally been preoccupied with overcoming the limitations of earlier studies. In particular, the issues of unobserved heterogeneity and endogeneity of training are addressed. In this connection two aspects have received special attention.

- The importance of longitudinal data containing information on training and measures of corporate productivity (or other effect measures). The improved data availability has made it possible to control for various factors in the estimation;

---

• The use of an experimental approach when defining the control group. A correct definition of the non-participants is crucial for an accurate estimate of the returns to training.

Traditionally, the returns to human capital investments have been estimated in terms of the effects on wages and productivity, and there is still a long way to go in this field. However, improved data availability has allowed researchers to analyse a number of other impacts in more detail. For instance, Fouarge and Schils (2009) found evidence that training is a way to keep older workers in the labour market, and Conti (2005) analysed the role played by skills accumulated during a person’s working life.

The importance of imperfections in the labour market and the consequent impact on the incentives and returns to training is also an increasing area of research within human capital investments.

The conclusions from some of the recent studies are described below.

Wage effects

Based on data from interviews with a representative sample of the Dutch population aged 16-64, Leuven and Oosterbeek (2008, 2002) used an alternative approach to estimate the wage return to work-related training. Using different treatment and control groups, they found that the resulting impact of training on wages varied considerably, depending on which method was used. Three control groups were defined: those who wanted to receive the training but did not participate because of a random event; those who wanted to but did not participate for other reasons; and those who did not want to participate. The participant group was divided into two: those who participated at least once and those who participated once only. Leuven and Oosterbeek (2008) found that the estimated wage effect of a 40-hour training course varied from 17% to about zero, depending on the assumptions made concerning participation and the control group.

A recent study (Pavlopoulos, Muffels and Vermunt 2009) reported that training increases the likelihood of moving from low to higher pay and reduces the likelihood of transition from higher pay to low pay.

Productivity effects

Dearden et al. (2006, 2000) used a panel of British industries to analyse the returns to work-related training over an extended period of time and at an aggregated industry level. They employed a variety of estimation techniques and also found that training is associated with significantly higher productivity. For instance, a one-percentage point increase in the proportion of employees that receive training is associated with an increase in value added per hour of about 0.6 percent and an increase in hourly wages of about 0.3 percent. Dearden et al. (2006) criticise the focus on wage effects that, according to them, has led to an underestimation of the importance of training.

Conti (2005) uses an industry panel covering all sectors of the Italian economy to estimate the effects of training on productivity and wages. Based on several modelling specifications and a variety of panel data techniques, Conti argues that training significantly boosts productivity. For instance, an increase of one percentage point in the stock of trained employees in an industry results in a 0.4 percent increase in productivity and a 0.1 percent increase in wages. These effects are thus smaller than those estimated by Dearden et al. (2006). There could be several reasons behind this. For instance, it might be caused by the existence of unobserved industry-specific factors, like human resource management practices, that have not been controlled for. Nonethe-
less, both studies find that it is the firms that reap the largest part of the returns when their employees participate in work-related training.

Based on a panel of all large firms with over 100 employees, operating in Portugal, Almeida and Carneiro (2009) examine the returns to training for employers and employees as a whole, irrespective of how these returns are shared between the two parties. They find that the returns to firms providing the training amounts to 8.6 percent. This substantial return suggests that company training of employees is a sound investment, which apparently yields comparable returns to investments in physical capital and in education.

**Ageing and obsolescence of human capital**

In a recent study, Fouarge and Schils (2009) conclude that in most European countries training can keep older employees in the labour market for longer. Based on data from the European Community Household Panel, Fouarge and Schils examine the effects of early retirement incentives on the training participation of older employees. Their results confirm the standard prediction from human capital theory that older employees do indeed participate less in firm-sponsored training than do younger ones, due to the expected lower net returns of such investments for both the employee and the firm. They refer to descriptive evidence for the conclusion that investing in training promotes continued participation in paid employment. With a view to an ageing workforce, the study suggests that there are employability gains to training older employees. This seems to be a clear-cut case of under-investments. For the society as a whole, training older employees yields benefits in terms of enlarged workforce, especially if the demand for labour exceeds the supply.

In a review of recent literature in this field, De Grip (2004) shows that training participation in the Netherlands from 1998 to 2000 was 15-20% lower among 56-65 year-old employees than those in the 26-55 age-group. He also shows that in the recession two years later, the reduction in training participation was greater, the older the employee. For employees in the 56-65 age-group, participation in training fell from 45 to 32%. He points at human capital obsolescence as a major risk to the greying knowledge economies. Obsolescence of human capital may lower both the productivity of the respective employees and their participation in the labour force. At the same time, he notes that the literature on this topic is very limited, probably as a consequence of the limited longitudinal data on available skills, skills demanded, and learning activities. It is not the obsolescence of knowledge from company training that is relevant to look at here, but rather the impacts of the various types of human capital obsolescence on the value of company training.

De Grip (2004) concludes that there is a knowledge gap concerning the relations and interactions between human capital obsolescence, labour market institutions, including remuneration patterns, and labour market withdrawals. He suggests that the statistical basis be improved and initiatives to strengthen life-long learning be promoted with the aim of maintaining workers productive potential in our knowledge societies.

**Business cycle variations**

In an article on obsolescence, in which he analyses participation in training by sector of industry in 1998-2000 and in 2000-02, De Grip (2004) sheds light on the risk of human capital obsolescence in the Netherlands and shows how training has sharply declined due to the 2000-02 recession. This decline varies from one sector to the next, with chemicals, other manufacturing and commerce being the hardest hit. In contrast, the food and beverages industries, which often use ESF subsidies for financing training courses, have shown an increase in training participation. These statistical data show how company
training is affected by business cycles. The overall picture here is of a very clear pro-cyclical pattern, but the observations cover only a few years. It is not clear whether the returns to training have been affected accordingly, but there is apparently a circular cause and effect interaction, where the economic recession affects training that may in turn affect production and incomes.

There are however arguments for pro-cyclical, a-cyclical as well as contra-cyclical patterns of company training. The opportunity costs of training are lower in a downturn and this may encourage employers to increase training. Caponi et al. (2009) use the Canadian Workplace and Employee Survey dataset and find empirical as well as model-based evidence for the existence of two opposing drivers that affect company training in relation to business cycles. The conclusion is that firms train more in downturns (contra-cyclical) and that positive sectoral shocks have a positive impact on training in the sector (pro-cyclical). It is furthermore concluded that sectoral analysis is very important for the specification of the links between business cycles, sectoral shocks, innovation in firms, and the incidence and intensity of training.

The latter conclusion is in line with the analysis of De Grip (2004), but apart from that, the question of business cyclical patterns of company training can be considered to be unsettled.

Other effects

Jones et al. (2008) analyse the effects of training on job satisfaction and workplace performance, using the British 2004 Workplace Employee Relations Survey. They find clear evidence that training is positively associated with job satisfaction, and that job satisfaction in turn is positively associated with most measures of performance. The authors conclude that employers may be able to improve an organisation’s performance by increasing the volume of training and taking action to raise the job satisfaction of the workforce. However, success depends on the quantity and type of training offered. For instance, training courses shorter than two days do not appear to have beneficial effects on financial performance, productivity or product quality, although they do seem to induce lower drop-out and absence rates. The authors conclude that only when training covers a large proportion of the workforce, does it appear to have positive effects on financial performance and productivity.

3.1.3 Concluding remarks

The review shows a picture that seems to be consistent and logical. The main findings may be interpreted as follows.

There is no clear evidence of positive wage effect of training at employee or company level, but recent studies show a productivity effect at industry level, which is high enough to argue that human capital investment compares with other investments in terms of social profitability. The problem is that only part of these returns go to the investor, the company that pays for the training. Most benefits may go to other companies in the industry and to the employee who received the training and who moves to another employer. This partially explains why the wage effect and the productivity effect at company level are often questionable and insignificant.

If the larger part of the returns to human capital investments really does go to the industry as a whole rather than to the individual company, there is a case for involving industry organisations more in the promotion of training, rather than leaving the training decisions with the companies. The picture is not clear and there should be more focus on these aspects in future research and evaluations.
There are indications that large companies offer more training than do small ones, and that a large part of the training is of a general nature. No conclusions can be drawn regarding participation and effects by gender and age, except that older employees receive less training than do younger ones.

The questions of obsolescence and the effects of training elderly employees is another field that needs more research, especially with regard to the development of greying economies in Europe.

The lessons learned are summarised in Box 1 hereafter. Box 2 highlights the knowledge gaps that are particularly challenging in the context of the ESF.

**Box 1 – Investing in the employed – the lessons learned**

- Returns to firm-sponsored training are positive, especially with regard to productivity, and at the industry rather than the company level.
- Training increases job satisfaction, which can be seen as one of the reasons for increased productivity and may further reduce drop-out and absence rates.
- Returns on wages are lower. This means that the employers reap the largest part of the returns to training, at least in the short term.
- Wage effects are larger after changing to a new job than with the current employer, and training increases the likelihood of moving.
- Firms invest in general training, which accumulates human capital applicable in all or many firms, and general training leads to more significant and larger returns than specific training.
- External and more formalised training activities result in higher returns than internal and more informal types of training. An exception seems to be the training of young people.
- Larger firms provide more training to their employees than do smaller ones.
- Older employees are less likely to participate in training than are younger employees, in particular during recessions. This leads to a risk of under-investment in this field. Assuming that training postpones labour market exit, this goes against the need to lengthen working lives.
- There is probably some degree of under-investment in training by employers, which may be due to market imperfections but is also explained by lower returns at micro level (current employer) than at macro level (industry or society).

**Box 2 – Investing in the employed – knowledge gaps**

- Differences in the impact of training in relation to gender.
- Impact of training in terms of postponing the labour market exit of elderly people.
- Rates of return to society, including soft returns such as job satisfaction.
- Differences in the practice and impact of training in relation to education level.
- Effects of training on innovation and competitiveness, and further macroeconomic effects.
3.2 Training of the unemployed: knowledge and knowledge gaps

In the following section we will consider the academic literature on the training of unemployed persons. In this case, the investor is typically the public sector and the main expected effect is a higher probability of the participants of finding a job.

Training and other similar human capital investments in unemployed persons are very often part of broader programmes, the so-called Active Labour Market Policies (ALMPs) that also include other measures such as job search assistance, wage subsidies to the private sector, and direct job creation in the public sector. Research on human capital investments in unemployed persons is therefore often part of studies on the impacts of ALMPs.

In general, ALMPs aim to create contact between the labour market and the unemployed or to raise the unemployed worker's skills and abilities so that he or she becomes more attractive in the labour market. The latter can be characterised as human capital investments, as opposed to the former.

In the following, we will describe and compare the results of a meta-study of ALMP evaluations (Kluve, 2006), a later meta-study of ALMPs (Card et al., 2009) and a number of recent micro studies and evaluations of ALMP with a focus on human capital investment and particularly on training. The main focus will be on empirical evidence of the impacts of training unemployed persons.

3.2.1 Kluve's meta-study (2006)

Kluve (2006), concludes in his meta-study of ALMP-evaluations that "the evidence from existing evaluations remains inconclusive: there is little consensus on whether Active Labour Market Policies actually reduce unemployment or raise the number of employed workers, and which type of programme seems most promising". Kluve furthermore wonders whether it is at all possible for one country to learn from ALMP experiences in another country.

The objective of his paper is to overcome this deficit by studying the available cross-country evidence on ALMP effectiveness. He distinguishes between four types of ALMP elements, namely training programmes, (which constituted about 40 percent of all ALMP costs of EU15 in 2003), private-sector wage subsidy and start-up grant schemes, public sector direct employment programmes, and “Services and Sanctions”, a category comprising counselling and monitoring, job search assistance, and corresponding sanctions in case of non-compliance.

The meta-study was part of a major research project for the European Commission, DG Employment, Social Affairs and Equal Opportunities. In the study, a collection of 73 recent studies are presented. They are categorised as "third-generation European evaluation studies" from the period after the mid-1990s (1996 – 2005). The earlier systematic European reviews in the field are categorised as the reviews of "first generation evaluation studies" by Heckman et al. (1999) and the "second generation review" by Kluve and Schmidt (2002).

In the first generation evaluations, according to Kluve (2006), new policies were analysed “applying rather new econometric techniques on the basis of rudimentary data” and the second generation evaluations were “mostly characterised by more mature and a more extensive set of policies, by a deepened and rapidly developing methodological know-how, and frequently much improved data”. With “the methodological know-how”, Kluve refers to “the inherent unobservability of the counterfactual no-programme situation” which
was “discussed extensively in the literature (cf., for instance, Heckman et al. 1999)”.

Kluve (2006) analyses the correlation of the effectiveness of the programme with selected variables, including the type of programme, the study design, the institutional context, and the economic background in the country at the time of the particular programme. Eighteen (that is, one fifth) of the articles evaluate training programmes, and a total of 32 evaluate ALMP, which include training as a main component.

The conclusions of the analysis are clear, showing that there is little systematic relationship between programme effectiveness and all contextual factors other than the type of programme.

The general picture from the evaluations of training programmes may be summarised in the following main points (Kluve, 2006: 10-12):

- Training is the most widely used active labour market measure in Europe.
- The results of the analysis are mixed: there are several indications that training programmes do increase participants' post-treatment employment probability, but it is concluded that training programmes seem to have relatively small effects and often have a significant effect only in the long run.
- There are indications that the positive long-run treatment effects of training could outweigh possible negative lock-in effects, (the negative employment effect of the unemployed spending less time on searching for jobs as a result of spending time in a programme).
- The results of training are better for participants with better labour market prospects, and for women.
- No conclusion can be drawn on pro-cyclical patterns of effectiveness of training programmes or the opposite.

A general result of the qualitative review is that:

- All types of programmes have little impact when the target group is young people\(^7\), and
- Incentives for job creation, job search assistance, and sanctions for inefficient job search are more effective than training.

After this general, qualitative review of the latest evaluations, Kluve (2006) also conducted a quantitative cross-country analysis of available evidence over a longer period of time, ranging from the 1980s to 2006. The effectiveness of the programmes is categorised as positive, zero, or negative, and this is correlated with a number of variables describing the types of programme, study designs (randomised experiment or not, and size of sample), the institutional labour market context (dismissal protection, fixed terms and temporary employment indicators), and the economic background at the time of the programme (unemployment rate, annual growth, and level of expenditures on ALMP).

The conclusions of the quantitative meta-study are that:

- The type of programme is the only factor that has a clear impact on its effectiveness;
- Stand-alone training programmes are found to have a modest likelihood of recording a positive impact on post-programme employment rates;

\(^7\) Something which does not necessarily mean that such programmes should be abolished or redesigned.
• Private-sector incentive programmes, and "Services and Sanctions" show a significantly better response than training;
• Direct employment programmes in the public sector seem to be highly ineffective or even counter-productive as instruments to increase participants' employment prospects.

3.2.2 Recent micro studies

Another meta-analysis was conducted by David Card, Jochen Kluve and Andrea Weber in 2009, covering IZA- and NBER-studies⁸ from 1996 to 2008, with 46% of the studies covering (published in the period from 2006-08). This meta-study synthesises some of the main lessons in the recent micro-economic evaluation literature, using a new sample of studies. It is somewhat different from Kluve (2006), as there is less focus on training as such and more on the impacts of ALMP in general, and the conclusions that are drawn concern ALMP in general rather than training. About 55% of the ALMP's reviewed are either training programmes or combined programmes.

The conclusions of Card et al. (2009) confirm or modify some of the conclusions from Kluve (2006). These conclusions of other recent works are described below under the following headings:

• Employment effects of training
• Variations among target groups
• Pro- or contra-cyclical patterns
• Incentives and threat effect
• Economic rate of return.

Employment effects of training

The comparison of the effects of training among the various studies is hampered by the use of different measures of programme impact, resulting from the variation in methodological approaches in the ALMP literature. In Card et al. (2009), one third of the reviewed studies measured the effects on the hazard rate from registered unemployment and 45% measured the outcome in terms of the impact on the probability of employment at some date after completion of the programme. As these two measures of programme impact are not comparable, Card et al. classify the estimates in three categories: significantly positive, insignificantly different from zero, and significantly negative.

The meta-study from 2009 shows that 39% of the reviewed programmes have significantly positive impacts in the short term (less than 1 year), 50% in the medium term (1-2 years), and 54% in the long term. This is shown in the table below, where we see that only 10% and 6% respectively of the programmes have significantly negative impacts in the medium and long term.

---

Table 2 - Impact estimates in Card et al. (2009)

<table>
<thead>
<tr>
<th></th>
<th>Percentage of estimates that are...</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Significantly positive</td>
</tr>
<tr>
<td>Short-term impact estimates</td>
<td>39.3%</td>
</tr>
<tr>
<td>Medium-term impact estimates</td>
<td>50.0%</td>
</tr>
<tr>
<td>Long-term impact estimates</td>
<td>54.0%</td>
</tr>
</tbody>
</table>

Source: Card et al. (2009), Table 5.

It is further concluded in the meta-study that the effect of classroom and on-the-job training is relatively weak in the short run when compared to job search assistance. However, the impact of training programmes seems to rise between the short and the medium term. This result is in line with Kluve (2006) and with other studies such as Hotz, Imbens and Klerman (2006), who concluded that although job search assistance programmes dominate training in the short run, the training is relatively more beneficial over longer periods.

A large number of research studies have been conducted in this field over recent years. In general, all of them find negative employment effects of the training programmes in the very short term, mainly while the training programme is being conducted. The negative effect during the first period is due to a lack of time and motivation to wholeheartedly search for jobs before and during the training (lock-in effect). This may be a reason for the finding of Card et al. (2009), that a large number of training programmes have had significantly negative effects in the short term. There seems to be consensus on the existence of this effect, but there is less agreement whether or not the training programmes generate a sufficiently positive post-programme effect to outweigh the initial negative effect.

Zhang (2003) investigated the effects of active labour market programmes for Norwegian adults (25-50 years old) for the 1990 to 2000 period. Three types of active labour market programmes were evaluated (training programmes, wage subsidy programme and an employment programme). The training programme was mainly a qualification scheme, consisting of formal training courses including both general and occupational specific training offered by the public employment services. The duration of the programmes varied, but most of them lasted for 1-5 months and only a few lasted up to 10 months. A strong negative effect of about 45.6 percent reduction of the employment probability was found while the labour market programmes were ongoing, but after the training the participants achieved a 59.6 percent average increased probability of finding a job.

Four studies have evaluated the long-term effect of training of unemployed people in Germany in 1993 and 1994. Two of these studies have focused on West Germany and two on East Germany. The results are in general more positive in the two West German studies (Lechner et al. (2004) and Fitzenberger et al. (2006)) than in the two East German studies (Lechner et al. (2005) and Fitzenberger & Völter (2007)).

The training programmes have been classified differently, which makes a direct comparison difficult. All four studies divide training into three subgroups, and all studies have a ‘Retraining’ category of more formal and long duration training with a long lock-in period of around two years. The post-programme
long-term effect is however positive, with an employment effect of about 15% in West Germany, and around 5% to 10% in East Germany.

Lechner et al. (2004) & (2005) divide the remaining training with respect to duration: above or below six months. They find that shorter training programmes perform better than longer ones in terms both of lock-in period and post-programme effects. The post-programme effect of a short training outweighs lock-in effects by around 10%, whereas no significant added value is found in the case of longer training programmes.

Fitzenberger et al. (2006) and Fitzenberger & Völter (2007) distinguish general training and specific training. They find that specific training programmes perform better than general ones. The duration of the training programmes are more or less the same. The lock-in period is the same for the two types of training, but is longer in East Germany. The main difference is in the post-programme effect which is around 15% for the specific training and only 10% for the general training. Impact is lower in East Germany, and even insignificant in the case of general training.

These four studies show positive results of training, but they also highlight some aspects that can affect the outcome. The evaluation methods (data, technique, assumptions) in Lechner et al. (2004) & (2005) and in Fitzenberger et al. (2006) and Fitzenberger & Völter (2007) are very similar. They all find that training in West Germany performs better than in East Germany and that different types of training perform differently. They also find that the effects of longer training do not appear until after several years. The evaluation therefore requires a longer training programme to include the full effect.

Another three studies investigated the employment effect of training at the beginning of the 21st century: Rinne et al. (2007), Biewen et al. (2007) and Wunsch & Lechner (2007b). To some extent they all have the shortcoming of limited longitudinal data, so that long-run effects are not evaluated. Their results differ, although the estimation methods and data are quite similar.

- Rinne et al. (2007) find the most positive results with significant post-programme effects from 5 to 9 percent.
- Biewen et al. (2007) estimate positive but only partially significant post-programme effects of zero to ten percent higher employment probability.
- Wunsch & Lechner (2007b) do not estimate any significant positive post-programme effects at all.

One reason for the differences could be that the types of training differ slightly. Rinne et al. (2007) focus on three different kinds of training that all have a relatively short duration. Biewen et al. (2007) and Wunsch & Lechner (2007b) consider a more broad range of training. Both studies find most positive results in training programmes with a relatively short duration.

Part of the reason for the different results may be the different definitions of control groups, which is important for the resulting estimates (Stephan, 2008).

The employment effects of five of the training programmes studied in Germany in the 1990s and in the beginning of the present decennium are shown for comparison with groups of programmes with similar durations in the table below.

The training programmes have very different durations, and lock-in effects increase with the length of the programme. The same cannot be said about the post-programme effect, and as a result the shorter programmes tend to be the most effective. They have a smaller negative lock-in effect and apparently
about the same post-programme effect, and shorter programmes can be assumed to be less costly.

Table 3 - Estimates of the employment effects of training of a short and long duration

<table>
<thead>
<tr>
<th>Author and types of training</th>
<th>Average duration</th>
<th>Negative lock-in effect</th>
<th>Positive post-training effect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lechner et al. (2005): East Germany 1993-94</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short duration</td>
<td>Under 6 months</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Long duration</td>
<td>Over 6 months</td>
<td>30%</td>
<td>0 – 10%</td>
</tr>
<tr>
<td>Retraining</td>
<td>22 months</td>
<td>40%</td>
<td>5 – 10%</td>
</tr>
<tr>
<td><strong>Fitzenberger et al. (2006): West Germany 1993-94</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Practice firms</td>
<td>6 months</td>
<td>0 – 10%</td>
<td>0 – 10%</td>
</tr>
<tr>
<td>Special and professional skills and techniques</td>
<td>6 months</td>
<td>5 – 15%</td>
<td>10 – 20%</td>
</tr>
<tr>
<td>Retraining</td>
<td>15 months</td>
<td>10 – 25%</td>
<td>10 – 20%</td>
</tr>
<tr>
<td><strong>Fitzenberger and Völter (2007): East Germany 1993-94</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Practice firms</td>
<td>6 months</td>
<td>20%</td>
<td>0 – 10%</td>
</tr>
<tr>
<td>Special and professional skills and techniques</td>
<td>9 months</td>
<td>20 – 25%</td>
<td>5 – 15%</td>
</tr>
<tr>
<td>Retraining</td>
<td>18 months</td>
<td>20 – 40%</td>
<td>0 – 20%</td>
</tr>
<tr>
<td><strong>Biewen et al. (2007):</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-duration</td>
<td>2 months</td>
<td>5%</td>
<td>0 – 5%</td>
</tr>
<tr>
<td>Practical further training</td>
<td>7 months</td>
<td>10 – 20%</td>
<td>0 – 10%</td>
</tr>
<tr>
<td>Classroom further training</td>
<td>8 months</td>
<td>10 – 25%</td>
<td>0 – 10%</td>
</tr>
<tr>
<td><strong>Wunsch and Lechner (2007b):</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short duration</td>
<td>2 months</td>
<td>5%</td>
<td>-2 – -4%</td>
</tr>
<tr>
<td>General training</td>
<td>4 months</td>
<td>20 – 30%</td>
<td>-5 – -3%</td>
</tr>
<tr>
<td>Longer general training</td>
<td>10 months</td>
<td>30 – 45%</td>
<td>-6 – -3%</td>
</tr>
</tbody>
</table>

In a study by Jespersen et al. (2008), the effect of training was estimated by applying the same method as the seven studies in Germany but in another context. Jespersen et al. (2008) analysed the impact of training courses with an average duration of 28 weeks under the Danish active labour market programmes. They investigated data for an eleven-year period from 1995 to 2005. The estimated employment effects were negative, minus 5 percent to minus 10 percent in the short term, (1-3 years), and an insignificant 1 percent to 2 percent in the longer term (5-11 years). Jespersen et al. (2008) also estimates the income effect of training, which was negative in the short term and insignificant in the long run.

The reason for the difference between the results of Jespersen et al. and the German studies mentioned above may be found in contextual and institutional factors that are quite different from one country to another (labour market,
economic growth, and ALMPs). Alternatively, different definitions of counterfactuals may also be a decisive factor explaining the differences.

**Variations among target groups**

**Gender**

The general picture with regard to the importance of gender in the impact of training has been that women seem to benefit more from training than men. This was also concluded in the European study by Kluve (2006). Zhang (2003) estimated a 15.2 percent higher effect for women than for men. Biwien et al. (2007) find that training effects are generally greater for women than for men, and their analysis shows that women also gain from practical training while men do not. However, they do not find any positive effect for women in East Germany, where the estimated employment effects are generally insignificant.

Arellano (2005) and Lechner et al. (2005) find that females gain a bit more from long training programmes than do men, whereas Fitzenberger and Völter (2007) and Rinne et al. (2007) do not identify any gender effect. Lechner et al. (2005) finds that in particular long-duration retraining shows a high positive effect for women, in terms both of employment and incomes, whereas retraining is "completely ineffective" for men. The reason for this result is found in the fact that the majority of women were retrained towards occupations with average or below average unemployment rates. More than 70% of the men were trained towards construction-related occupations, which were severely hurt in 1995/1996, just when the retraining was completed.

Jespersen et al. (2008) concludes that the economic deficit from classroom training is smaller for women than for men, and Card et al. (2009) found that the comparisons by gender were not statistically significant in any of the cases considered.

There seems to be a tendency towards less gender differences, but the reasoning by Lechner et al. (2005) reminds us that there is often a contextual reason for such differences.

**Young people**

Cards et al. (2009) confirmed the findings in Kluve (2006) that all types of programmes had limited impacts when the target group is young people. However Card et al. did not analyse this for training as such, but for ALMP in general.

The estimations in Rinne et al. (2007) indicate that younger participants (between 17 and 34) face higher lock-in effects.

This was also found by Larsson (2000), who analysed Swedish data for two programmes for youth, a youth practise programme and a labour market training programme. The results in terms of both re-employment probability and subsequent earnings indicated either a zero or a slightly positive effect, but the training programme had the weakest performance.

It is obvious that firm conclusions cannot be taken on the basis of such studies, and it seems more reasonable to conclude that such efforts targeting young people need to be further developed. This should probably be designed for the particular group of persons again and again, for every new generation of young people.

**Other groups**

In addition, a number of other finding have been made for different groups of participants, but the specific national and labour market context may often prevail over specific employment effects for groups like immigrants, age
groups etc. Kluve (2006) concluded from his review of earlier articles that in addition to the gender effect, the results were generally better for participants with better labour market prospects. The estimations in Rinne et al. (2007) find no support for the hypothesis that participants with better labour market prospects gain more from training. It should be noted that Rinne et al. (2007) do not consider long-run training.

Hämäläinen (2002) compared the employment effect of three categories of participants, the disadvantaged, the average and the advantaged, and showed that the relative training effects were different between these groups, but that the effects varied over the business cycle.

The timing of the training in relation to the unemployment period may also affect the employment effect of the training. According to Wunsch and Lechner (2007b), individuals treated later than 5 month after the beginning of their unemployment period show smaller lock-in effects. The same pattern is pointed out by Gerfin and Lechner (2000) who find that participants starting their training more than 9 months after becoming unemployed gain more from training.

Fitzenberger et al. (2006) find that, especially for participants in long-run retraining, the results are less successful for participants with less than 6 months of unemployment before entering the programme. They observed both larger negative lock-in and less positive post-programme effects for this group.

Wunsch and Lechner (2007b) found that individuals without vocational education gain more from short-run training programmes than do the others.

**Pro- or contracyclical patterns**

The question here is mainly whether the impact of training varies over the business cycle.

Card et al. (2009) showed the shares of reviewed studies indicating significantly positive, statistically insignificant, and significantly negative short- and medium-term results for training programme implemented during four time periods: 1985-89, 1990-94, 1995-99, and 2000-07. The post training benefits vary from a high level in the beginning of the 1990s to lower impacts in the late 1990s, but as there is no distinction by countries, nothing can be said about the relationship with the business cycles. Further complexity is created by the fact that business cycles do not show uniform patterns across countries during such a period of time.

Others have paid more attention to the pattern of training impacts over the business cycle. Hämaläinen (2002), in an article on impacts of labour market training in Finland, focused on this aspect for the 1989 – 1994 period during which the unemployment rate increased from 3.4 percent to 18.4 percent. The estimated employment effect of training declined by 15-20 percentage points during the same period. It was therefore concluded that the effectiveness of labour market training was negatively correlated to overall unemployment. At the same time, the proportion of labour market trainees who benefited from participation in the training decreased by some 20 percentage points from 1992-94. This means that employment, training effectiveness and the intensity of training fell during the observed period.

Hämäläinen (2002) distinguishes between three types of participants, the disadvantaged, the average, and the advantaged, and shows very different training effects between these groups over the business cycle. The effects of training of persons in the disadvantaged groups were low during periods with low unemployment (1989), when training effects were greater for the advantaged. Five years later, when unemployment was high, the estimated effect of
training of disadvantaged and advantaged groups had increased by about 20 percentage points and decreased by 10 percentage points, respectively. This indicates a contra-cyclical effect of training for the disadvantaged and a pro-cyclical training effect for the advantaged groups.

Zhang (2003), in connection with an evaluation of the Norwegian active labour market programmes, 1989 – 2002, observed pro-cyclical patterns of both training and wage subsidy programmes. The effects of training programmes were stronger when the labour market conditions were good and job opportunities were more favourable. Zhang (2003) made use of monthly calendar dummies and was therefore able to capture the aggregate labour market conditions in terms of business cycle and seasonal changes on the labour market.

A number of reasons for the positive relation between the employment effect and overall employment conditions (the pro-cyclical pattern) have been suggested. A conventional explanation is that in situations with a fall in job opportunities and increased competition over existing jobs, the efforts put into training are weakened and consequently the benefits from the training decrease. Another argument is that a higher number of participants, during times of high unemployment, leads to decreasing returns to scale of training courses, which may imply a weaker outcome of the training (Hämäläinen 2002).

Hämäläinen (2002) suggested two other arguments, namely that the administrators may alter the selection of participants during times of high unemployment, or that employers may have different recruitment behaviours in times of low and high unemployment. When unemployment increases, administrators may select trainees with better employment prospects, or the employment targets may have created incentives to cream off the potentially successful participants. On the other hand, when unemployment is low, the employers may use labour market training as a signal of the ability and willingness to learn new skills. In a situation of mass unemployment, there are large numbers of skilled job-seekers, and this may reduce the importance of such signalling.

Lechner and Wunsch (2007a), on the other hand, found in Germany during the period 1986-95 a positive correlation between the unemployment rate at the start of a programme and its effectiveness, i.e. a negative relationship between labour market conditions at programme start and the effects of training, which is a contra-cyclical pattern. They found, in particular, that the negative lock-in effects are greater in times of low unemployment, but at the same time that the positive long-run effects are greater in times of high unemployment. On the basis of a 10-year administrative database compiled for evaluating German training programmes, they found a clear positive correlation between the effectiveness of the training and the unemployment rate.

Their result appears to conflict with that of Hämäläinen (2002) and Zhang (2003) but an important difference is in the timing of the labour market conditions. Hämäläinen and Zhang observed the unemployment rates at the same time as they observed effects, and related the two, whereas Lechner and Wunsch (2007a) related the employment effects of training to the labour market conditions at programme start. This obviously provides different and incomparable results.

According to Lechner and Wunsch (2007a), higher negative lock-in effects in times of low unemployment occur because the unemployed make less efforts to find work and receive fewer job offers from the Employment Service while in the programme. If unemployment is high, it takes longer to find a job, and the cost of reduced job searching because of attending a programme is therefore lower.
Finally, Fitzenberger et al. (2006) consider training in West Germany in 1986-87 and in 1993-94. The period from 1986 to 1993 was related to a boom in the West German economy due to unification, whereas the period after 1993 was one of recession. The estimated employment effects were essentially similar for the two periods, which indicates that the business cycles did not have an effect on the gains and losses from training in this case.

It is difficult to draw a clear-cut conclusion on this aspect of training effectiveness. The research that is referred to does not all point in the same direction, although there is a certain weight to studies and arguments pointing to a procyclical pattern in so far as training effectiveness is higher in times of high employment. But at this point a number of specific institutional and administrative behaviours also described by Hämäläinen (2002) may affect the observed patterns.

**Incentives and threat effect**

In his comparison of training programmes with alternative ALMPs, Kluve found that private-sector incentive programmes, job search assistance, and sanction schemes for inefficient job search show a significantly better performance than training. He recommended a well-balanced combination of basic services with appropriate sanctions for non-compliance and with training programmes and other measures.

A few researchers have analysed further the potential of incentives and sanctions, and have discussed the existence and importance of the "threat effect".

The existence of such an effect implies a positive employment effect of the training programmes even before the training is started. Rosholm and Svarer (2004) tested the assumption that “the threat of programme participation is so harsh that some individuals increase their search activity and/or lower their reservation wages in order to find a job before the programme starts”. Rosholm and Svarer (2004) found that the threat effect significantly reduced the duration of unemployment by approximately three weeks. They argue that the employment effect is underestimated if the threat effect is not considered because the counterfactual group is also affected by the threat effect. Their estimation is based on Danish data i.e. in a context with a high benefit replacement rate, whereas Black et al. (2003) estimated the threat effect at an average of 2.2 weeks of unemployment on the basis of US data.

The threat effect is seldom considered in evaluation studies, which is probably due to the difficulties in an estimation of this effect that occurs very early in the process and which may often affect both treatment and control groups. Rosholm and Svarer (2004) and Black et al. (2003) solved this problem in two different ways and estimated the threat effect.

Rosholm & Svarer (2004) estimate the probability of being enrolled in an ALMP and consider this probability as the threat. Black et al. (2003) use experimental data where only some individuals are in risk.

**Economic return to society**

Card et al. (2009) conclude that few studies provide enough information to perform even a crude cost-benefit analysis, and programme costs are often unknown or unreported. They express their hope that future studies will adopt a more substantive focus, enabling policy-makers to evaluate and compare the return on investments.

Fitzenberger et al. (2006) argue that an overall cost-benefit assessment of the microeconomic effects is not possible since certain information required for a comprehensive cost-benefit-analysis are lacking.
Winter-Ebmer (2001) notes that a formal cost-benefit analysis is only rarely done in evaluation studies, mainly because in most cases the measured benefits are small or marginal in the short term.

However, in a few cases attempts were made to estimate the economic rate of return through a traditional cost-benefit analysis.

Jespersen et al. (2008) estimated the costs and benefits of several Danish active labour market programmes. First they made impact estimates, including lock-in and post-training employment effects, applying a propensity score-matching method to a very detailed set of administrative data (Jespersen et al., 2008). They then estimated the income effects and actual costs of the training.

The cost-benefit analysis took into account earnings, income transfers, direct operation costs, wage subsidies associated with job training programmes, and loss of taxation. For a stand-alone training course, the direct costs (EUR 12,000) are far from being offset by the benefits. The net present value (NPV) is clearly negative.

The displacement effect of job training and the value of leisure time were not taken into account in the calculations of Jespersen et al. (2008). The value of leisure time might be estimated by using the concept of happiness. Winkelmann and Winkelmann (1998) estimated and compared the effect on happiness of being unemployed, with the effect of household incomes and found that the negative effect of unemployment was much higher than the positive value of increased household incomes. To offset the negative effect on happiness of being unemployed, a person's income has to rise by a factor of between 4 and 7!

It may be argued that the intangible social effects of training programmes leading to employment are quite large, but on the other hand, the net value should also include the loss of income and happiness of those persons that might be replaced as a consequence of the employment effect of labour market training.

3.2.3 Concluding remarks

Studies at micro level have shown very mixed results and most recent research indicates that labour market training has a positive employment effect, i.e. a positive medium- and long-term employment effect is probably sufficient to counterbalance a negative short-term lock-in effect. It may be concluded though that labour market training has a relatively limited short term effectiveness in comparison with other labour market instruments. The effectiveness of training is improved if there is a mix with other instruments such as job search assistance and financial incentives / sanctions.

Very little research exists on the economic returns on human capital investments in terms of labour market training, and no examples have yet been identified with documentation of benefits from labour market training programmes covering the costs over a lifetime. This does not mean that this is not the case, but the lack of data concerning programme costs and benefits has been the main barrier. This is particularly the case for benefits at macro level, where the replacement effect seems to eliminate the positive effects, but where there may be other benefits from the human capital investments, in particular in the years to come, where problems with ageing and a declining labour force are foreseen.

9 These issues are addressed in Section 4.5 hereafter.
Very limited evidence also seems to exist for concluding on specific general patterns of employment effects in relation to target groups. There is no reason to believe in a special gender effect except under specific circumstances. There seems to be consensus that it is difficult to achieve success with labour market training for young people, but this may also be caused by insufficient efforts or ability to adapt the efforts to the demand of this particular group. In addition, this group is quite heterogeneous and a number of other socio-economic factors should also be taken into account.

No clear conclusion is drawn concerning the possible pattern of training effectiveness over the business cycle. There are some indications that the training effect is pro-cyclical, in the sense that it is higher in periods with good job opportunities and low unemployment, but the picture is not clear.

We conclude that the existing general knowledge on the employment and income effect of labour market training compared to its costs is very limited.

The aim of further knowledge production in this field might focus on:

- Developing instruments for cost-benefit analysis and concrete calculations of costs and benefits of specific programmes. In particular there is a knowledge gap concerning the costs of the programmes and of the possible positive externalities of labour market training of different types;
- Improving the employment effects of labour market training, for example with the use of combinations of labour market training and instruments that help to provide appropriate incentives to the participants;
- Too little is known regarding the types of training that will be effective and help young people to find jobs. Just concluding that the previous efforts have not yielded positive results is not a satisfactory answer to such an important question.

The lessons learned are summarised in the Box 3 below. Box 4 then highlights the knowledge gaps that are particularly challenging in the context of the ESF.
Box 3 – Investing in the unemployed – the lessons learned

- Training is the most widely used active labour market measure in Europe.
- Labour market training has a negative lock-in effect in the short run, during and immediately after the training, when the unemployed spend their time on the programme instead of searching for jobs.
- Training programmes have relatively limited effects on employment in the short run, but there is a positive employment impact in the long run.
- In many instances positive long-run impact outweighs the negative lock-in effect.
- Impact of training on young job seekers is limited.
- There may be an additional positive impact originating from the threat effect, i.e. the fact that potential participants make an extra effort to find a job before the training starts.
- Impacts are better with short training programmes than with longer ones, partly as a result of lower lock-in effects and partly because the marginal effectiveness of training may decrease with duration.
- Training is more effective if it is associated with other instruments such as job-search assistance, and if it takes place in a context of incentives and sanctions for taking / not taking job opportunities.

Box 4 – Investing in the unemployed – knowledge gaps

- Effects of various combinations of return-to-work training, coaching and counselling services, and the overall incentive framework for finding jobs.
- Specific effects of various mixes of instruments on young job-seekers.
- Effects of training older unemployed people in terms of postponing their exit from the labour market.
- Rates of return to society, including soft benefits and externalities associated with the fact that someone exits unemployment.
- Variations of the impact of training the unemployed over the business cycle.

3.3 Persons at risk of losing their jobs: another knowledge gap

People at risk of losing their jobs fall more or less into a category between employees and the unemployed, i.e. the two sections considered above. Such people are a very important target of the ESF, and very little is known on this issue.

Major knowledge gaps relate to:
- Impact of training in terms of preventing skills obsolescence.
- Impact of training employees in at-risk firms and sectors.
- Specific pattern of under investments in at-risk firms and sectors.

3.4 Macroeconomic impact estimates

Micro-level impact studies cannot stand alone if there is a need for a global picture of the effects of training. The inclusion of potentially important secon-
dary effects requires a macroeconomic analysis, whether the evaluation concerns company training or training of the unemployed.

**Company training**

The value of company training may be underestimated by focusing only on the employer and on the employee’s wage effect. Some studies have taken into account the effects on wages with future employees, but the full picture must also include the wider impacts through subsequent job changes and replacements for the trainee as well as for others in the labour market.

A few studies assess the macroeconomic impact of training. The primary focus in these studies is the effect of training on economic growth. Depending on the objective of the study, the starting point is a computable general equilibrium (CGE) model or a non-overlapping or overlapping generation model. The parameters and elasticities in these models are based on the microeconomic analyses. Thus, the underlying foundation of the models reflects the results from the microeconomic field described above. For instance, when setting the elasticity of productivity with regard to firm-sponsored training, the experts behind the development of the European Commission’s Labour Market Model (Berger et al. 2009) refer to Dearden et al. (2006).

Most macroeconomic studies conclude that, all things being equal, training has a positive impact on economic growth. For instance, Scicchitano (2005) demonstrates that the composition of human capital, i.e. formal education as well as on-the-job training, is important in determining the probability of innovation occurring and the growth rate of the economy. Hence, training programmes are essential to bridge gaps in human capital generated by schools and colleges and the society’s needs. In line with the findings of microeconomic research (see 3.1.3), Scicchitano also concludes that general training targeting all employees is more effective than firm-specific training.

In this connection, Scicchitano emphasises the importance of an improved access to training for low-skilled employees. The findings also show that general training prevents low development traps, when R&D is absent.

Some macroeconomic studies examine the effects of training on economic growth indirectly, when analysing the impact of population ageing on economic growth. In an ageing society labour will be relatively scarce and capital relatively abundant. Hence, interest rates will fall. Fougère and Mérette (1999) conclude that population ageing could create more opportunities for future generations to invest in human capital formation, which would stimulate economic growth. Thus, human capital investments may be a means to reduce significantly the negative impact of ageing on output per capita.

A similar conclusion is reached by Ludwig et al. (2010), who find that endogenous human capital formation is an important adjustment mechanism, which substantially mitigates the macroeconomic impact of population ageing. In their simulations for the American economy they estimate that the rate of return will fall by only a 0.3 percentage point until 2025 with endogenous education, compared to a 0.7 percentage point in a model with a fixed human capital profile.

A number of recent articles analyse the productivity effect of training at industry level. This is an alternative way of including the positive externalities that can be expected to arise from company-provided training.

**Training the unemployed**

The wider impact of the training of the unemployed may be even more important for the full picture of the effect of training. Micro studies show that training of the unemployed and other ALMP efforts may or may not give the
participants a higher chance of being employed. In the wider context, this may have secondary effects that are positive or negative. A negative effect is substitution, i.e. the fact that a participant finds a job at the expense of someone else who becomes or remains unemployed something which means that the macro-level effect is nil.

There may nevertheless be a positive productivity effect from the training if the improvement of the trainees’ skills make themmore productive than the persons they replace. Substitution may also lead to an improved matching in the labour market and may thus increase the productivity of other employees. This type of positive externality has also been more or less neglected in the research on labour market training effects. There are thus many potential outcomes of the training, and even with the best, most reliable results at micro level, little is known about the total macro-level effects.

There are very few macro studies on the effects of human capital investments, especially on the background of the large and increasing number of micro studies. They often consider ALMPs as a whole rather than training alone. Boone and van Ours (2004), however, have investigated the effectiveness of active labour market policies on an aggregate level from both an empirical and a theoretical point of view.

In the empirical analysis, the effects of ALMP on both unemployment rates and employment-population rates were analysed using aggregate time series – cross-section data from 20 OECD countries. The effects of specific categories of ALMP focusing on training, public employment services and subsidised jobs were analysed for the time period 1985-1999. They concluded that labour market training was the most effective programme to bring down unemployment.

Macro studies consider how expenditures on ALMPs/training affect the labour market performance. This gives rise to major estimation problems, especially concerning endogeneity. There is no standard solution to overcome these problems and Boone and van Ours (2004) pursue two different approaches. They can both be questioned, the first with respect to country-specific effects and the second with respect to endogeneity problems. However, the estimation results are the same; expenditures on training affect the labour market performance positively, with respect to both lowering the unemployment rate and raising the number of employed individuals.

### 3.5 Views of the expert panel

In an earlier version, the content of this chapter was reviewed and commented upon by a panel of seven experts (see Appendices 3 and 4). In addition, the experts were asked to reach a consensus, as far as possible, on a series of issues. Unsurprisingly, their views are close to what arises from the research synthesis.

The panellists reached a quasi-unanimous agreement on the following points:

- In the short run, training is likely to produce productivity gains for employers, and unlikely to pay off in earnings for workers;
- Training for unemployed individuals may be ineffective, unless it is used in conjunction with other instruments within a consistent activation policy;
- It is likely that there are some specific areas of under-investment in training by the private sector, something which creates social justification for public investment;
- General equilibrium effects matter and need to be considered.

There was a relatively high level of consensus on the following point:
In order to avoid the 'creaming effect', profiling of the unemployed on an individual basis is a feasible alternative to targeting specific social groups.

The following issues were considered as knowledge gaps:

- Comparing the impact of training on unemployed women and men respectively;
- Specific areas of under-investment in training by the private sector, especially as regards at-risk employees;
- Understanding the relationship between the impact of training and the business cycle;
- Nature and magnitude of the main general equilibrium effects (e.g. substitution effect).
4 Lessons from four country studies

This chapter displays the impact analyses which were carried out successively in Hungary, Italy, Poland and Belgium (Tasks 3 to 6). In the following sections, the study team endeavours to present the main findings and technicalities in a way which is intended to be accessible to those readers who are unfamiliar with statistical techniques.

A cross-section analysis then identifies the lessons learned across the four countries, and discusses their transferability (Task 7).

Finally, the study team reflects on the assumptions under which human capital investments could be profitable enough from the standpoint of the society.

4.1 Impact of training: Hungarian country study

4.1.1 Dataset (HU)

The dataset originates from two distinct sources:

- A sample of 1,071 beneficiaries of training and employment subsidies financed by the ESF through the Human Resource Development Operational Programme (HEFOP).
- A control group of 5,654 unemployed jobseekers drawn from the records of the Hungarian Employment Service, and who have not been affected by any active labour market measure.

The training and subsidised employment programmes took place in 2006. The employment status of participants and non-participants was measured in May 2008. The surveyed participants consists of a 20 percent random sample from the total number of programme beneficiaries. The control group represents about five percent of unemployed jobseekers. The data originate from seven of the 20 Hungarian regions.

4.1.2 Surveyed participants (HU)

The descriptive analysis displays the characteristics of the sample of participants, the type of attended activities, and the labour market status 24 months after completion of the training.

Profile of surveyed participants

Table 4 below displays the profile of the sample of participants in terms of a number of individual characteristics.

---

10 See more details in Appendix 6.
Table 4 - Profile of surveyed participants (HU)

<table>
<thead>
<tr>
<th>Type of public</th>
<th>Share of sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>48.6</td>
</tr>
<tr>
<td>Female</td>
<td>51.4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td>30 years old or younger</td>
<td>41.4</td>
</tr>
<tr>
<td>between 31 and 50</td>
<td>47.5</td>
</tr>
<tr>
<td>50 or older</td>
<td>11.1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td>Primary education</td>
<td>30.5</td>
</tr>
<tr>
<td>Secondary education</td>
<td>61.9</td>
</tr>
<tr>
<td>Tertiary education</td>
<td>7.6</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td>Not a long-term job seeker</td>
<td>84.7</td>
</tr>
<tr>
<td>Long-term job seeker</td>
<td>15.3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>
| **Total (number)**   | **1,071**       

**Type of human capital investment**

The sample of surveyed participants had benefited from either subsidised jobs or training, but never both, as shown in Table 5 below.
Table 5 - Surveyed participants per type of activity (HU)

<table>
<thead>
<tr>
<th>Type of public</th>
<th>Subsidised employment</th>
<th>Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>53.1</td>
<td>47.0</td>
</tr>
<tr>
<td>Female</td>
<td>46.9</td>
<td>53.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td>30 years old or younger</td>
<td>36.1</td>
<td>43.4</td>
</tr>
<tr>
<td>between 31 and 50</td>
<td>50.7</td>
<td>46.3</td>
</tr>
<tr>
<td>50 or older</td>
<td>13.3</td>
<td>10.3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td>Primary education</td>
<td>20.4</td>
<td>34.7</td>
</tr>
<tr>
<td>Secondary education</td>
<td>70.1</td>
<td>58.5</td>
</tr>
<tr>
<td>Tertiary education</td>
<td>9.5</td>
<td>6.8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td>Not a long-term job seeker</td>
<td>85.4</td>
<td>84.9</td>
</tr>
<tr>
<td>Long-term job seeker</td>
<td>14.6</td>
<td>15.1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

The above table shows that the share of women who participate in training is larger than the share of women who take part in subsidised employment. Persons who have participated in training are on average somewhat younger than those who have participated in subsidised employment. The share of persons below 30 is higher in the sample of training participants. The share of persons with primary education is also larger in the sample of training participants. These differences are statistically significant. The share of long-term job seekers is nearly the same within both groups.

Geographical profile

Seven counties were selected for the study. Table 6 below shows the profile of surveyed participants per county, in association with economic information about these counties.
The above table shows that the participants are not spread out evenly across the seven counties. A comparison with the size of the counties in terms of the economically active population shows that there is a negative relationship. The share of participants is not larger for counties with a large population, but smaller. A relatively larger number of participants were selected in counties with high unemployment rates.

**Employment status after training**

Table 7 below shows the labour market status of participants 24 months after participation.

The likelihood of finding employment is higher for both the job seekers who have received training (56.5 percent) and those who have participated in subsidised employment (47.5 percent). The table furthermore shows that the differences between men and women among participants are small. Of those who have received training, participants from the youngest age group are most successful in finding employment. Among the participants who took part in subsidised employment, the middle age group shows the highest percentage of job finders. The youngest age group shows the highest percentage of job finders and the oldest age category shows the lowest percentage. More education means higher odds of finding employment.

---

**Table 6 - Surveyed participants and economic context (HU)**

<table>
<thead>
<tr>
<th>County</th>
<th>Subsidised employment</th>
<th>Training</th>
<th>Total</th>
<th>Economically active population (thousands)</th>
<th>Unemployment rate 2006 (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Budapest</td>
<td>5.4</td>
<td>11.7</td>
<td>9.2</td>
<td>788</td>
<td>4.8</td>
</tr>
<tr>
<td>Baranya County</td>
<td>4.1</td>
<td>13</td>
<td>9.7</td>
<td>159</td>
<td>7.6</td>
</tr>
<tr>
<td>Békés County</td>
<td>25.2</td>
<td>20.2</td>
<td>20.3</td>
<td>142</td>
<td>7.8</td>
</tr>
<tr>
<td>Fejér County</td>
<td>24.1</td>
<td>14.7</td>
<td>16.5</td>
<td>190</td>
<td>4.9</td>
</tr>
<tr>
<td>Győr-Moson-Sopron County</td>
<td>0.3</td>
<td>11.1</td>
<td>9.5</td>
<td>195</td>
<td>4.3</td>
</tr>
<tr>
<td>Heves County</td>
<td>11.9</td>
<td>7.6</td>
<td>8.3</td>
<td>124</td>
<td>9.1</td>
</tr>
<tr>
<td>Szabolcs-Szatmár-Bereg County</td>
<td>28.9</td>
<td>21.7</td>
<td>26.4</td>
<td>216</td>
<td>13.6</td>
</tr>
<tr>
<td><strong>Total (percent)</strong></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>8.3</td>
<td>13</td>
</tr>
<tr>
<td><strong>Total (number)</strong></td>
<td>294</td>
<td>709</td>
<td>1,003</td>
<td>1,813</td>
<td></td>
</tr>
</tbody>
</table>

---


12 Ibid.

13 average
### Table 7 – Status of participants after 24 months (HU)

<table>
<thead>
<tr>
<th>Type of public</th>
<th>Subsidised employment</th>
<th>Training</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not employed</td>
<td>Employed</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>43.6</td>
<td>56.4</td>
</tr>
<tr>
<td>Female</td>
<td>43.5</td>
<td>56.5</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 years old or younger</td>
<td>38.7</td>
<td>61.3</td>
</tr>
<tr>
<td>Between 31 and 50</td>
<td>46.3</td>
<td>53.7</td>
</tr>
<tr>
<td>50 or older</td>
<td>46.2</td>
<td>53.8</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary education</td>
<td>48.3</td>
<td>51.7</td>
</tr>
<tr>
<td>Secondary education</td>
<td>43.2</td>
<td>56.8</td>
</tr>
<tr>
<td>Tertiary education</td>
<td>35.7</td>
<td>64.3</td>
</tr>
<tr>
<td>Duration of unemployment (before)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not a long-term job seeker</td>
<td>44.6</td>
<td>55.4</td>
</tr>
<tr>
<td>Long-term job seeker</td>
<td>37.2</td>
<td>62.8</td>
</tr>
<tr>
<td>Total</td>
<td>43.5</td>
<td>56.5</td>
</tr>
<tr>
<td>Total (number)</td>
<td>128</td>
<td>166</td>
</tr>
</tbody>
</table>

**4.1.3 Comparative analysis (HU)**

This section compares the sample of participants and the control group. Two methods were applied successively:

- Method 0 - Selection bias not corrected
- Method I - Selection bias corrected by exact matching

Technical terms such as selection bias, propensity score matching and pooled analysis are explained in Chapter 4.4.

**Method 0 - Selection bias not corrected**

Method 0 starts with a simple comparison of the overall samples of participants, as displayed in Table 8. It shows that the proportion of employed people after two years is 19.1% higher among those who benefited from subsidised jobs, and 10.1% higher among trainees, in comparison with persons
from the control group at the same time interval after they were first sampled from the Hungarian Public Employment Service job seekers’ database.

Table 8 – Impact on employment status after training (HU/Method 0)

<table>
<thead>
<tr>
<th>Individual characteristics</th>
<th>Subsidised employment</th>
<th>Training</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not employed</td>
<td>Employed</td>
<td>Not employed</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>43.6</td>
<td>56.4</td>
<td>53.5</td>
</tr>
<tr>
<td>Female</td>
<td>43.5</td>
<td>56.5</td>
<td>51.6</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 years old or younger</td>
<td>38.7</td>
<td>61.3</td>
<td>54.2</td>
</tr>
<tr>
<td>Between 31 and 50</td>
<td>46.3</td>
<td>53.7</td>
<td>49.7</td>
</tr>
<tr>
<td>50 or older</td>
<td>46.2</td>
<td>53.8</td>
<td>57.5</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary education</td>
<td>48.3</td>
<td>51.7</td>
<td>58.5</td>
</tr>
<tr>
<td>Secondary education</td>
<td>43.2</td>
<td>56.8</td>
<td>49.4</td>
</tr>
<tr>
<td>Tertiary education</td>
<td>35.7</td>
<td>64.3</td>
<td>47.9</td>
</tr>
<tr>
<td>Duration of unemployment (before)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not a long-term job seeker</td>
<td>44.6</td>
<td>55.4</td>
<td>50.8</td>
</tr>
<tr>
<td>Long-term job seeker</td>
<td>37.2</td>
<td>62.8</td>
<td>61.7</td>
</tr>
<tr>
<td>Total</td>
<td>43.5</td>
<td>56.5</td>
<td>52.5</td>
</tr>
<tr>
<td>Total (number)</td>
<td>128</td>
<td>166</td>
<td>372</td>
</tr>
</tbody>
</table>

row percentages

The above table furthermore shows that the differences between men and women in both treatment groups are small. The effect of age varies across the three groups. Amongst those who have received training, participants from the youngest age-group are most successful in finding employment. Amongst the participants who took part in subsidised employment, the middle age-group shows the highest percentage of job finders. In the control group, the likelihood of finding employment declines with age. The youngest age-group shows the highest percentage of job finders and the oldest age-group shows the lowest percentage. More education means higher odds of finding employment. This phenomenon is strongest in the control group. The fact that participants with only primary education have a higher likelihood of finding employment in comparison to non-participants with primary education suggests that the effect of the ESF investment is particularly strong for persons with a low level of education.
**Method 0 – Potential selection bias**

Table 9 below displays the differences and similarities in the composition of the surveyed participants and control group, given that such differences may affect the success of participants in the labour market, and therefore create a "selection bias". Since this study focuses on the employment impact of investment in training or education, the participants who entered subsidised employment are excluded from the analysis. The remaining group of surveyed participants now consists of 563 persons who took part in training.

**Table 9 – Profile of participants and control group (HU)**

<table>
<thead>
<tr>
<th>Types of public</th>
<th>Participants</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>48.0</td>
<td>49.2</td>
</tr>
<tr>
<td>Female</td>
<td>52.0</td>
<td>50.8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td>30 years old or younger</td>
<td>49.0</td>
<td>31.2</td>
</tr>
<tr>
<td>between 31 and 50</td>
<td>43.5</td>
<td>48.3</td>
</tr>
<tr>
<td>50 or older</td>
<td>7.5</td>
<td>20.5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td>Primary education</td>
<td>35.7</td>
<td>40.0</td>
</tr>
<tr>
<td>Secondary education</td>
<td>58.4</td>
<td>55.1</td>
</tr>
<tr>
<td>Tertiary education</td>
<td>5.9</td>
<td>4.8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td>Not a long-term job seeker</td>
<td>85.6</td>
<td>73.9</td>
</tr>
<tr>
<td>Long-term job seeker</td>
<td>14.4</td>
<td>26.1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td><strong>Total (number)</strong></td>
<td><strong>563</strong></td>
<td><strong>5,654</strong></td>
</tr>
</tbody>
</table>

The composition of the two groups does differ in terms of age. There are more young job seekers in the sample of surveyed participants and the share of job seekers that are 50 or older is higher in the control group. Participants and the control group also differ with respect to the level of education. The share of persons who have completed primary education is 35.7 percent among surveyed participants and 40.0 percent in the control group. This difference is statistically significant. The percentage of long-term job seekers is higher in the control group (26.1 percent) than it is in the sample of participants (14.4 percent). This difference is also statistically significant.

**Method I – Propensity to participate**

As a first step in this analysis, a model was estimated in order to predict the probability of participating, depending on various individual characteristics. The model was adjusted on the whole dataset of participants and non-participants.

A positive and statistically significant parameter implies that a given characteristic contributes positively to the probability of participation. In the situation that a dummy variable (0-1) is used, one category has to be left out. For instance, one of the education categories (primary education) is excluded.
from the specification. The effects of the other two education categories are assessed relatively to the category ‘primary education’.

Table 10 – Factors affecting participation (HU)

<table>
<thead>
<tr>
<th>Explanatory factors</th>
<th>Estimated parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>/</td>
</tr>
<tr>
<td>Female</td>
<td>.099</td>
</tr>
<tr>
<td>Age 30 or under</td>
<td>/</td>
</tr>
<tr>
<td>Age between 31 and 50</td>
<td>-.499**</td>
</tr>
<tr>
<td>Age 50 or older</td>
<td>-1.326**</td>
</tr>
<tr>
<td>Primary education</td>
<td>/</td>
</tr>
<tr>
<td>Secondary education</td>
<td>.118</td>
</tr>
<tr>
<td>Tertiary education</td>
<td>.340*</td>
</tr>
<tr>
<td>Not a long-term job seeker</td>
<td>/</td>
</tr>
<tr>
<td>Long-term job seeker</td>
<td>-.638**</td>
</tr>
<tr>
<td>Unemployment level in the county</td>
<td>.033</td>
</tr>
<tr>
<td>Intercept</td>
<td>-2.290**</td>
</tr>
<tr>
<td>Number of observations</td>
<td>6 217</td>
</tr>
</tbody>
</table>

** 95% confidence level, * 90% confidence level

Table 10 above shows that women and men do not have a different likelihood of participating in the programme. Persons between 31 and 50 are less likely to participate than people below 30. For persons aged 50 or above the likelihood of participation is even lower. There is a weak positive relationship between the level of education and the likelihood of participating. Long-term job seekers are less likely to enrol in a programme than persons who have been looking for a job for less than a year.

Method I - Impact analysis through exact matching

The exact matching technique is used to produce an impact estimate without selection bias. Each person in the sample of participants is matched with a single person from the control group for whom all characteristics are the same, i.e. gender, age category, education level, and unemployment duration. Gender and unemployment duration both consist of two categories. For age and education there are three categories. In total, there are 36 possible combinations of characteristics (2x3x3x2). As the control group is about five times larger than the treatment group and the number of possible combinations is limited, it was possible to find a pair from the control group for every person in the sample of participants.

After matching, 536 pairs of participants and non-participants could be analysed. A consequence of this one-to-one exact matching is that about 4,000 individual data from the control group were not included in the analysis.
Table 11 – Impact on employment (HU/Method I)

<table>
<thead>
<tr>
<th></th>
<th>Surveyed participants A</th>
<th>Control Group B</th>
<th>Impact estimate (I) A-B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated probability of being employed after 24 months (percent)</td>
<td>45.7</td>
<td>40.0</td>
<td>+ 5.6**</td>
</tr>
</tbody>
</table>

** - significant at 95% confidence level

After correcting for selection bias by means of exact matching, the effect of programme participation is still positive (+5.6%) but smaller than what was estimated through Method 0 (+8.2 for the unemployed people who participated in training only. This shows that it was worth correcting the selection bias.

Impact estimates per public (Method I)

Table 12 below applies the same model to estimating the impact of training that might be observed if all participants and non-participants were successively men, women, youngsters, and so on.

Table 12 – Impact on employment per public (IT/Method I)

<table>
<thead>
<tr>
<th>Additional probability of being employed after 24 months</th>
<th>Gender</th>
<th></th>
<th>Age</th>
<th></th>
<th>Education</th>
<th>Duration of unemployment (before)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>3.8</td>
<td>30 years old or younger</td>
<td>-0.7</td>
<td>Primary education</td>
<td>7.9</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>7.9</td>
<td>Between 31 and 50</td>
<td>7.5</td>
<td>Secondary education</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50 or older</td>
<td>11.7</td>
<td>Tertiary education</td>
<td>-1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Not a long-term job seeker</td>
<td>2.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Long-term job seeker</td>
<td>9.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The programme is shown to have a strong positive impact on some difficult target groups, such as long-term job seekers (+9.6%), elderly job seekers (+ 11.7%) and less educated people (+ 7.9%).

On the other hand, the impact on young people is close to null.
4.1.4 Lessons (HU)

The impact of analysed interventions, assessed in terms of the probability of being employed after 24 months, is globally assessed as about 5%.

However, the activities under study include training and subsidised jobs, and the dataset suggests that the impact of subsidised jobs is greater.

As regards the specific publics, the impact of training is:

- Higher for women, lower-educated people, elderly people, and long-term job seekers;
- Particularly low for young participants.

4.2 Impact of training: Italian country study

4.2.1 Dataset (IT)

The dataset was provided by ISFOL. It covers 10 regions of Northern Italy (Operational Programme Obj. 3) and originates from two distinct sources:

- A sample of 47,494 participants were interviewed 12 months after having finalised their training. The data are derived from a series of regional placement surveys conducted during the period 2003-2005 and applying to ESF co-funded training courses completed in the years 2001, 2002 and 2003. After a long process of harmonising regional surveys and local archives, the surveys could be deemed to be comparable enough for aggregating in a single dataset. Trained participants were 18 to 45 year old, and unemployed at the beginning of the training. Interviews were complemented by administrative archives of the Managing Authorities and administrative archives of the beneficiary organisations implementing the training activities.

- Another sample of 8,049 non-trained people was extracted from the micro-data of the Quarterly Labour Force Survey of the National Institute of Statistics (ISTAT). In order to compare it with the participants having completed their training in 2001 and 2002, ISFOL used 12-month panel data, so that the before/after employment status could be shown. This approach became impossible in 2003 since ISTAT had changed its survey method. A specific retrospective section was therefore added in the ISTAT questionnaire where respondents were asked to provide information on the previous year. The control group was then consistently screened in order to ensure comparability with the sample of surveyed participants (regions, age, labour status and so on). Moreover, only people who had not attended any training course in the recent years were included.

---

14 See more details in Appendix 6.
15 Some regions covered the entire population of beneficiaries while others surveyed a sample. Some regions had their surveys carried out by an independent firm, while others relied on the beneficiary organisations providing the training.
16 Derived from birth date. Missing data dropped.
17 Students and inactive people were excluded.
18 This could be done since the ISTAT questionnaire included such specific demands.
4.2.2 Surveyed participants (IT)

This section describes the surveyed participants in terms of: (1) attended training activity, (2) socio-economic profile, (3) regions, (4) duration of unemployment, and (5) labour market status 12 months after completing the training.19

Participation in training activities

The surveyed participants attended various categories of training activities:

- **A**: Investments intended to enhance vocational and other key skills for the unemployed and to remove barriers to finding a job.
- **B**: Investments focused on equal opportunities for all and promoting social inclusion, especially by providing access to basic skills to disadvantaged groups excluded from or under-represented in the workplace.
- **C**: Investments aimed at widening the participation in lifelong learning so that more people continue throughout their life to develop their knowledge, skills and understanding, and to improve their employability in a changing labour market.
- **D**: Investments aimed at improving the skills base and the adaptability of the labour force, and at increasing the level of entrepreneurship. Only a very small percentage of the surveyed people benefited from such training.
- **E**: Investments aimed at reducing the level of disadvantage faced by women in the labour market by improving access to learning and removing barriers to employment.

In fact, these five categories were the five ‘axes’ (policy areas) of the 2000-2006 Italian Community Support Framework. They were used as building blocks in the design of all Operational Programmes at regional level.

Table 13 below displays a breakdown of surveyed participants in four categories. Due to the small number of participants, investment targeted at business start-ups (D) was not included in the analysis.

The table shows that category A (training unemployed job seekers) is by far the most numerous in the sample of participants, something which reflects the overall priorities of the ESF in Northern Italian programmes.

---

19 The figures shown in the tables reflect the distribution observed in the sample and not the population. Figures per year are not displayed in this version of the report. There are two reasons why an analysis per year could lead to different findings: (1) training activities may have changed over the years, and (2) changes may have occurred in the business cycle and economic context. For instance, the regional unemployment rate is used in the estimated models, and it shows annual differences.
Table 13 – Surveyed participants per type of training (IT)

<table>
<thead>
<tr>
<th>Category</th>
<th>Type of training programme</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
</tr>
<tr>
<td>A</td>
<td>Training unemployed job seekers</td>
<td>25,295</td>
</tr>
<tr>
<td>B</td>
<td>Providing access to basic skills to disadvantaged groups</td>
<td>2,429</td>
</tr>
<tr>
<td>C</td>
<td>Widening the participation in lifelong learning</td>
<td>12,535</td>
</tr>
<tr>
<td>E</td>
<td>Improving women’s access to learning and removing barriers to employment</td>
<td>7,235</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>47,494</td>
<td>100</td>
</tr>
</tbody>
</table>

As shown in Table 14 below, the typical duration of training activities was in the range of 7 to 9 months.

Table 14 – Participants and duration of training programme (IT)

<table>
<thead>
<tr>
<th>Category</th>
<th>Type of training programme</th>
<th>Duration in months</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-3</td>
<td>4-6</td>
</tr>
<tr>
<td>A</td>
<td>Training unemployed job seekers</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>Providing access to basic skills to disadvantaged groups</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>Widening the participation in lifelong learning</td>
<td>6</td>
</tr>
<tr>
<td>E</td>
<td>Improving women’s access to learning and removing barriers to employment</td>
<td>9</td>
</tr>
</tbody>
</table>

Row percentages

Finally, Table 15 below displays the type of validation which participants acquire through their training. Category A (Training of unemployed job seekers) provides for the most formal training.

Table 15 – Participants and type of validation (IT)

<table>
<thead>
<tr>
<th>Category</th>
<th>Type of training programme</th>
<th>Diploma</th>
<th>Qualification</th>
<th>Attendance certificate</th>
<th>Nothing</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Training unemployed job seekers</td>
<td>20</td>
<td>36</td>
<td>17</td>
<td>27</td>
<td>100</td>
</tr>
<tr>
<td>B</td>
<td>Providing access to basic skills to disadvantaged groups</td>
<td>21</td>
<td>19</td>
<td>50</td>
<td>9</td>
<td>100</td>
</tr>
<tr>
<td>C</td>
<td>Widening the participation in lifelong learning</td>
<td>18</td>
<td>6</td>
<td>9</td>
<td>67</td>
<td>100</td>
</tr>
<tr>
<td>E</td>
<td>Improving women’s access to learning and removing barriers to employment</td>
<td>10</td>
<td>36</td>
<td>8</td>
<td>46</td>
<td>100</td>
</tr>
</tbody>
</table>

Row percentages
Socio-economic profile

A breakdown of participants per gender, age, and education level is shown in the next table, where columns A, B, C, D and E are types of training programmes as defined above.

There are slightly more women participants than men, although not in the case of the larger types of programme (training unemployed job seekers).

Young people are particularly numerous in Type A programmes (training unemployed job seekers) while elderly people represent a larger part of Type E programmes (women’s access to learning and employment).

Finally, Type B programmes (basic skills for disadvantaged groups) attract the largest proportion of low-skilled people.

Table 16– Profile of surveyed participants (IT)

<table>
<thead>
<tr>
<th>Type of training programme</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>E</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>48</td>
<td>57</td>
<td>48</td>
<td>6</td>
<td>47</td>
</tr>
<tr>
<td>Female</td>
<td>52</td>
<td>43</td>
<td>52</td>
<td>94</td>
<td>53</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>15-24 years old</td>
<td>37</td>
<td>22</td>
<td>20</td>
<td>12</td>
<td>26</td>
</tr>
<tr>
<td>25-34 years old</td>
<td>46</td>
<td>41</td>
<td>67</td>
<td>49</td>
<td>50</td>
</tr>
<tr>
<td>35-44 years old</td>
<td>14</td>
<td>32</td>
<td>12</td>
<td>32</td>
<td>19</td>
</tr>
<tr>
<td>Over 45 years</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>No or elementary education</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Lower secondary education</td>
<td>19</td>
<td>60</td>
<td>1</td>
<td>25</td>
<td>21</td>
</tr>
<tr>
<td>Upper secondary education</td>
<td>12</td>
<td>6</td>
<td>3</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Tertiary education</td>
<td>57</td>
<td>20</td>
<td>61</td>
<td>46</td>
<td>51</td>
</tr>
<tr>
<td>Higher education</td>
<td>12</td>
<td>7</td>
<td>35</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Geographic profile

Table 17 below shows a breakdown of surveyed participants per region and per type of investment. More than three quarters of them originate from three regions: Piemonte, Lombardia and Emilia Romagna.
Table 17– Surveyed participants per region (IT)

<table>
<thead>
<tr>
<th>Region</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>E</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of training programmes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piemonte</td>
<td>47</td>
<td>59</td>
<td>5</td>
<td>27</td>
<td>32</td>
</tr>
<tr>
<td>Valle d'Aosta</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Lombardia</td>
<td>17</td>
<td>-</td>
<td>20</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>Veneto</td>
<td>15</td>
<td>24</td>
<td>36</td>
<td>23</td>
<td>21</td>
</tr>
<tr>
<td>Friuli Venezia Giulia</td>
<td>5</td>
<td>3</td>
<td>12</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Liguria</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Emilia Romagna</td>
<td>6</td>
<td>9</td>
<td>13</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>Toscana</td>
<td>6</td>
<td>2</td>
<td>5</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Umbria</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Marche</td>
<td>2</td>
<td>-</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Bolzano</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Trento</td>
<td>0</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Duration of unemployment before training

Table 18 below shows the duration of unemployment when the training started. Over one third of the training starts during the first quarter after the participants have become unemployed. Over one third starts after the participants have been unemployed for more than a year. Small differences are seen in this parameter. Training for disadvantaged groups (line B) starts relatively late after the participant has become unemployed, whereas employability enhancing training starts relatively soon after the participant has become unemployed.

---

20 Bolzano and Trento are classified not as regions but as administrative zones.

21 Piemonte stands as the most well-represented region in the dataset because the placement survey covered all participants in this region.
Table 18 – Duration of unemployment before training (IT)

<table>
<thead>
<tr>
<th>Category</th>
<th>Type of training programme</th>
<th>Duration of unemployment before training (months)</th>
<th>0-3</th>
<th>4-6</th>
<th>7-9</th>
<th>&gt;9</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Training unemployed job seekers</td>
<td></td>
<td>47</td>
<td>24</td>
<td>15</td>
<td>14</td>
<td>100</td>
</tr>
<tr>
<td>B</td>
<td>Providing access to basic skills to disadvantaged groups</td>
<td></td>
<td>30</td>
<td>29</td>
<td>16</td>
<td>25</td>
<td>100</td>
</tr>
<tr>
<td>C</td>
<td>Widening the participation in lifelong learning</td>
<td></td>
<td>66</td>
<td>19</td>
<td>8</td>
<td>7</td>
<td>100</td>
</tr>
<tr>
<td>E</td>
<td>Improving women’s access to learning and removing barriers to employment</td>
<td></td>
<td>47</td>
<td>24</td>
<td>15</td>
<td>14</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>49</td>
<td>22</td>
<td>13</td>
<td>16</td>
<td>100</td>
</tr>
</tbody>
</table>

Row percentages

Employment status after training

Table 19 below shows the labour market status of the surveyed participants, 12 months after completion of the training.

Table 19  – Employment status after training (IT)

<table>
<thead>
<tr>
<th>Employed after one year (percent)</th>
<th>per type of training programme</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td>69.6</td>
<td>61.7</td>
<td>75.8</td>
<td>67.8</td>
<td>66.8</td>
<td>69.6</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td>69.2</td>
<td>56.6</td>
<td>74.9</td>
<td>70.8</td>
<td>65.0</td>
<td>69.2</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-24 years old</td>
<td></td>
<td>68.3</td>
<td>54.2</td>
<td>74.4</td>
<td>68.5</td>
<td>66.4</td>
<td>68.3</td>
</tr>
<tr>
<td>25-34 years old</td>
<td></td>
<td>72.0</td>
<td>61.1</td>
<td>76.8</td>
<td>73.1</td>
<td>69.1</td>
<td>72.0</td>
</tr>
<tr>
<td>35-44 years old</td>
<td></td>
<td>65.1</td>
<td>59.8</td>
<td>70.6</td>
<td>68.6</td>
<td>58.6</td>
<td>65.1</td>
</tr>
<tr>
<td>Over 45 years old</td>
<td></td>
<td>62.9</td>
<td>66.0</td>
<td>63.5</td>
<td>66.0</td>
<td>54.5</td>
<td>62.9</td>
</tr>
<tr>
<td>Level of education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No title or elementary</td>
<td></td>
<td>56.2</td>
<td>61.6</td>
<td>64.3</td>
<td>57.1</td>
<td>51.9</td>
<td>56.2</td>
</tr>
<tr>
<td>Middle school</td>
<td></td>
<td>60.9</td>
<td>56.4</td>
<td>67.5</td>
<td>70.4</td>
<td>53.4</td>
<td>60.9</td>
</tr>
<tr>
<td>Diploma qualification (2-3 years)</td>
<td></td>
<td>71.4</td>
<td>60.4</td>
<td>71.9</td>
<td>66.8</td>
<td>64.6</td>
<td>71.4</td>
</tr>
<tr>
<td>School diploma and high school</td>
<td></td>
<td>70.8</td>
<td>65.4</td>
<td>73.1</td>
<td>69.4</td>
<td>68.3</td>
<td>70.8</td>
</tr>
<tr>
<td>University degree</td>
<td></td>
<td>75.2</td>
<td>66.3</td>
<td>79.8</td>
<td>75.7</td>
<td>74.1</td>
<td>75.2</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>69.4</td>
<td>59.5</td>
<td>75.4</td>
<td>70.6</td>
<td>65.8</td>
<td>69.4</td>
</tr>
</tbody>
</table>

There are some differences among the various types of training. The proportion of interviewed participants who found a job appears to be relatively low in programmes targeting the disadvantaged groups (Measure B). On the other hand, this proportion is relatively high for the unemployed participating in
programmes targeting women (categories C and E). The educational level of the participants in these two measures is relatively high compared with Measure B, for instance.

Unsurprisingly, the lowest proportion of employed participants is found among people with a low level of education.

4.2.3 Comparative analysis (IT)

This section compares the sample of participants and the control group. Four distinct methods were applied as follows:

- Method 0 - Pooled analysis, selection bias not corrected
- Method I - Pooled analysis, selection bias corrected by propensity score matching
- Method I.R – Same as Method I, but analysis per region
- Method I.T – Same as Method I, but analysis per type of training

Method 0 - Pooled analysis, selection bias not corrected

Method 0 starts with a simple comparison of the overall samples of participants (all regions, all types of training). This comparison is displayed in Table 20 below and shows that the proportion of employed people after one year is 33.8% higher among participants.

<table>
<thead>
<tr>
<th>Surveyed participants</th>
<th>Control group</th>
<th>Impact estimate (0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>A-B</td>
</tr>
<tr>
<td>Estimated probability of being employed after one year (percent)</td>
<td>70.7</td>
<td>36.9</td>
</tr>
</tbody>
</table>

Method I - Pooled analysis, selection bias corrected

This method uses the propensity score matching (PSM) technique to correct the selection bias (see Chapter 5).

Table 21 below gives insight into the issue of selection bias. The study team has estimated the ‘propensity to participate’ through a model including age, gender, education level, region, and regional level of unemployment. The model is adjusted to the dataset of participants and non-participants. It shows that some people have a specifically high probability of participating (e.g. people older than 29, people with a low level of education). Such an uneven likeliness to participate may entail a bias because these characteristic also play a role in the eventual success or failure of the trainees on the labour market. Therefore, a simple comparison between the surveyed participants and the control group (Method 0) shows the impact not only of the training but also of the factors explaining participation in the training.
Table 21 – Factors affecting participation (IT)

<table>
<thead>
<tr>
<th></th>
<th>Model coefficient</th>
<th>Significant</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.224</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.016</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under 25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 to 29</td>
<td>-0.037</td>
<td></td>
<td>People over 29 are more likely to participate</td>
</tr>
<tr>
<td>30 to 34</td>
<td>0.322</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>35 to 44</td>
<td>0.760</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>over 45</td>
<td>0.847</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary or no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower secondary</td>
<td>-0.171</td>
<td>**</td>
<td>People with a higher level of education are less likely to participate</td>
</tr>
<tr>
<td>Upper secondary</td>
<td>-0.264</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Tertiary</td>
<td>-0.465</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Higher</td>
<td>-0.579</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td><strong>Region</strong>¹</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other regions</td>
<td></td>
<td></td>
<td>People living in the four main regions are more likely to participate than those living in the other regions</td>
</tr>
<tr>
<td>Piemonte</td>
<td>0.856</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Lombardia</td>
<td>0.410</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Veneto</td>
<td>0.721</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Emilia Romagna</td>
<td>0.900</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td><strong>Economic context</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regional unemployment rate</td>
<td>0.007</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* ** - significant at a 95% confidence level,

On the basis of the model displayed in the above table, the study team estimated the 'propensity to participate' of all people in the dataset. Each participant in the surveyed sample was then matched with the person of the control group having the closest propensity score. The findings are shown in Table 22 below.

---

¹ Since there are more people among the surveyed participants than in the control group, one person from the control group could be matched with several participants.
Table 22 – Impact on employment (IT/Method I)

<table>
<thead>
<tr>
<th>Estimated probability of being employed after one year (percent)</th>
<th>Surveyed participants A</th>
<th>Control Group B</th>
<th>Impact estimate (I) A-B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>70.7</td>
<td>41.2</td>
<td>+29.5**</td>
</tr>
</tbody>
</table>

** - significant at 95% confidence level

After correcting for selection bias the impact of the ESF-programmes is still assessed as positive (+29.5%), although smaller than what was estimated with Method 0 (+33.8% as seen in Table 20).

Impact estimates per public (Method I)

The next table applies the same model to estimating the impact that might be observed if all participants and non-participants were successively men, women, youngsters, and so on.

Table 23– Impact on employment per target group (IT/Method I)

<table>
<thead>
<tr>
<th>Employed after one year (percent)</th>
<th>Impact estimate (I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>per gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>32.7</td>
</tr>
<tr>
<td>Female</td>
<td>26.2</td>
</tr>
<tr>
<td>per age</td>
<td></td>
</tr>
<tr>
<td>20-24 years</td>
<td>22.9</td>
</tr>
<tr>
<td>25-29 years</td>
<td>30.8</td>
</tr>
<tr>
<td>30-34 years</td>
<td>31.6</td>
</tr>
<tr>
<td>35-44 years</td>
<td>29.7</td>
</tr>
<tr>
<td>More than 45 years</td>
<td>22.9</td>
</tr>
<tr>
<td>per level of education</td>
<td></td>
</tr>
<tr>
<td>No or elementary education</td>
<td>22.4</td>
</tr>
<tr>
<td>Lower secondary education</td>
<td>27.3</td>
</tr>
<tr>
<td>Upper secondary education</td>
<td>31.6</td>
</tr>
<tr>
<td>Tertiary education</td>
<td>27.8</td>
</tr>
<tr>
<td>Higher education</td>
<td>28.6</td>
</tr>
<tr>
<td>Average</td>
<td>29.5</td>
</tr>
</tbody>
</table>

As shown by the above table, the analysis finds a higher impact for men than women, for people aged 25 or more rather than younger ones, and for those with a medium level of education.
**Method I.R – Analysis per region, selection bias corrected**

In this section the study team applies the same approach as that of Method I, but once per region. The samples of surveyed participants are described in Table 24. Regional control groups are extracted from the overall control group. The analysis finds an impact which is higher than average in Piemonte, and lower in Lombardia.

The last column on the right shows the impact estimate with no correction of the selection bias, and highlights the importance of correcting this bias since the ranking of regions is not the same in the two columns.

<table>
<thead>
<tr>
<th>Region</th>
<th>Impact (Method I.R)</th>
<th>Impact (Method 0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piemonte</td>
<td>+36.9</td>
<td>+37.5</td>
</tr>
<tr>
<td>Lombardia</td>
<td>+13.1</td>
<td>+23.6</td>
</tr>
<tr>
<td>Veneto</td>
<td>+31.8</td>
<td>+40.0</td>
</tr>
<tr>
<td>Emilia Romagna</td>
<td>+27.6</td>
<td>+33.4</td>
</tr>
<tr>
<td><strong>All regions</strong></td>
<td><strong>+29.5</strong></td>
<td><strong>+33.8</strong></td>
</tr>
</tbody>
</table>

**Method I.T – Analysis per type of training, selection bias corrected**

As in the above case of regions, a series of separate analyses were undertaken per type of training and the findings are displayed in Table 25 below. The highest impact is that of vocational education.

Once again, the table shows how important it is to correct the selection bias.

<table>
<thead>
<tr>
<th>Type of training</th>
<th>Impact Method I.T</th>
<th>Impact Method 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internships</td>
<td>+27.1</td>
<td>+36.4</td>
</tr>
<tr>
<td>Re-integration</td>
<td>+25.2</td>
<td>+24.7</td>
</tr>
<tr>
<td>Part of compulsory education</td>
<td>+27.7</td>
<td>+22.3</td>
</tr>
<tr>
<td>Vocational education</td>
<td>+32.2</td>
<td>+35.8</td>
</tr>
<tr>
<td><strong>All types of training</strong></td>
<td><strong>+29.5</strong></td>
<td><strong>+33.8</strong></td>
</tr>
</tbody>
</table>

**4.2.4 Lessons (IT)**

Overall, the analysis of the Italian dataset reveals that 30% more unemployed people return to work when they benefit from ESF-supported training. Separate analyses show considerable variation across regions (in a range of 13-37%). The impact estimate is very high and the differences among regions are considerable. These findings cast a doubt on the reliability of the study, a point which is discussed in the following box.

In addition, impact estimates on particular target groups (see Table 23) show some differences with the Italian evaluation building on the same dataset. For
instance, this study finds a higher impact for men than women, contrary to the findings of Italian evaluations. This problem could not be resolved in the framework of this study.

**Box 5 – Assessing reliability (IT)**

The study team observes a series of facts that may support an explanation of the surprisingly positive findings of the analysis:

- Two of the regions which surveyed the largest number of participants (Piemonte, Veneto) did so by entrusting the training organisation with the task of collecting data. These regions show the highest success rates.
- Emilia-Romagna also has a very high success rate, although the data were collected by an independent research team. During the study period this region enjoyed an exceptionally low unemployment rate (3%).
- In Lombardia as in Emilia Romagna, the data were collected by an independent research team. The unemployment rate was higher. The impact estimate (+13% - see Table 24) is in line with that of other countries.

An interpretation that would fit in with these facts is that the dataset is partly biased by the data collection approach used in Piemonte and Veneto, and that the consequences of the exceptionally positive situation of the labour market could not be totally controlled.

Variations across types of training programmes (25 to 32%) seem limited in light of the generally very high level of impact estimates.

### 4.3 Impact of training: Polish country study

#### 4.3.1 Dataset (PL)

The dataset was provided by the Managing Authority in charge of the “Sectoral Operational Programme Human Resources Development 2004-2006”. It is drawn from the “Study on ultimate beneficiaries” of the programme, which involved several waves of surveys during the period 2006-2008. The dataset includes:

- A sample of 5,541 participants in ESF-supported training activities, who were interviewed about their employment status six month after the end of their participation. They had participated in several types of activities such as off-the-job training, on-the-job training, and/or counselling and advice. Activities were part of two “actions” of the programme, respectively called “Perspectives for youth” and “Counte-racting and combating long-term unemployment”. The main targeted groups were respectively: (1) young unemployed jobseekers and (2) long-term unemployed older than 24. Surveyed participants were selected through a layered randomised process (layers = action, gender, age, region) from the lists of recipients whose details were submitted by implementing bodies.
DG EMPL - Study on the Return on ESF Investment in Human Capital

- A control group was created in two steps. First, 81 local offices of the employment service were selected randomly and asked to provide lists of people who were unemployed at the time when participants were entering the projects, and who did not participate in any kind of training. Then a sample of 2,414 non-participants was selected through a layered randomised process (layers = gender, age, level of education, region, and local rate of unemployment). They were surveyed roughly at the same time as the sample of participants.

4.3.2 Surveyed participants (PL)

The descriptive analysis focuses on the participation in ESF-supported activities and the employment status after completing the training.

**Participation in ESF-supported activities**

Participants took part in three types of activity, defined as follows:

- On-the-job training, including internship, vocational education, and subsidised employment;
- Off-the-job training and education, including language courses, computer courses, advice or training on how to run a company, training or vocational workshops, training allowance / scholarship;
- Counselling and advice, e.g. vocational counselling, training on looking for a job and CV writing, psychological support.

Table 25 below displays a breakdown of surveyed participants in these three categories. It shows that participants have usually benefited from one or more types of training plus some advice and counselling. It also shows that the average participant is involved in at least one type of training, even if some counselling of advice is provided in addition. Therefore, it is justified to say that they participated in "training activities".

---

23 However, they have typically received standard job search assistance.
Table 26 – Participants per type of training activity (PL)

<table>
<thead>
<tr>
<th>Type of training activity</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
</tr>
<tr>
<td><strong>Activities for young unemployed jobseekers</strong></td>
<td></td>
</tr>
<tr>
<td>A1 - Off-the-job training and education</td>
<td>48</td>
</tr>
<tr>
<td>A2 - On-the-job training</td>
<td>68</td>
</tr>
<tr>
<td>A3 - Counselling and advice</td>
<td>58</td>
</tr>
<tr>
<td>A Total</td>
<td>2,773</td>
</tr>
<tr>
<td><strong>Activities for long-term unemployed older than 24</strong></td>
<td></td>
</tr>
<tr>
<td>B1 - Off-the-job training and education</td>
<td>72</td>
</tr>
<tr>
<td>B2 - On-the-job training</td>
<td>28</td>
</tr>
<tr>
<td>B3 - Counselling and advice</td>
<td>53</td>
</tr>
<tr>
<td>B Total</td>
<td>2,768</td>
</tr>
</tbody>
</table>

As shown by Table 27 below, young unemployed people participate in longer training activities than do older ones.

Table 27 – Participants and duration of training (PL)

<table>
<thead>
<tr>
<th>Type of training activity</th>
<th>Average duration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Days</td>
</tr>
<tr>
<td>Activities for young unemployed jobseekers</td>
<td>145</td>
</tr>
<tr>
<td>Activities for long-term unemployed older than 24</td>
<td>87</td>
</tr>
</tbody>
</table>

The available dataset does not inform on whether a qualification certificate is issued at the end of the training.

Socio-economic profile

A breakdown of surveyed participants per gender and education level is shown in Table 28 below, which also displays the types of training in lines.

Women participants are more involved in ‘on-the-job’ training than are the men, but not in the case of the larger type of programme (training unemployed job seekers).

Young people with a low level of education are more often enrolled in a mix of off-the-job training and counselling. Older people with a low level of education tend to be enrolled in on-the-job training.

---

24 Total is more than 100% because a given participant typically benefits from several types of support.
Table 28 – Surveyed participants per public (PL)

<table>
<thead>
<tr>
<th>Activity</th>
<th>Total (number)</th>
<th>Gender</th>
<th>Education level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>Activities for young unemployed jobseekers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Off-the-job training and education</td>
<td>2,773</td>
<td>1,591</td>
<td>1,182</td>
</tr>
<tr>
<td>On-the-job training</td>
<td>1,319</td>
<td>73</td>
<td>57</td>
</tr>
<tr>
<td>Counselling and advice</td>
<td>1,449</td>
<td>62</td>
<td>70</td>
</tr>
<tr>
<td>Total (number)</td>
<td>2,768</td>
<td>1,319</td>
<td>1,449</td>
</tr>
<tr>
<td>Activities for long-term unemployed older than 24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Off-the-job training and education</td>
<td>2,768</td>
<td>1,319</td>
<td>1,449</td>
</tr>
<tr>
<td>On-the-job training</td>
<td>1,319</td>
<td>73</td>
<td>57</td>
</tr>
<tr>
<td>Counselling and advice</td>
<td>1,449</td>
<td>62</td>
<td>70</td>
</tr>
<tr>
<td>Total (number)</td>
<td>2,768</td>
<td>1,319</td>
<td>1,449</td>
</tr>
</tbody>
</table>

Figures in percentages represent the share of target groups participating in various types of activity. The totals of some lines exceed 100% because surveyed people typically participated in more than one activity.

The available dataset does not inform on the geographical profile of surveyed participants. However, it must be remembered that the surveyed samples were representative of the whole country in geographical terms.

**Employment status**

Table 29 below shows the employment status of participants six months after they have completed the training. The overall picture is that 58 to 67% of participants are employed or self-employed, whereas all of them were unemployed before being enrolled in training activities.

The table also shows that women are systematically less successful than men, and the same applies to participants with the lowest level of education. This picture is very much in line with the objectives of the ESF.
Table 29– Employment status after training (PL)

<table>
<thead>
<tr>
<th>Target Status</th>
<th>Total</th>
<th>Female</th>
<th>Male</th>
<th>Basic</th>
<th>Vocational</th>
<th>Intermediary</th>
<th>Higher</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activities for young unemployed jobseekers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>42</td>
<td>49</td>
<td>33</td>
<td>52</td>
<td>38</td>
<td>48</td>
<td>40</td>
</tr>
<tr>
<td>Employed, not in own business</td>
<td>47</td>
<td>44</td>
<td>51</td>
<td>40</td>
<td>50</td>
<td>43</td>
<td>49</td>
</tr>
<tr>
<td>Runs own business</td>
<td>11</td>
<td>7</td>
<td>16</td>
<td>8</td>
<td>12</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Total (percent)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total (number)</td>
<td>2,773</td>
<td>1,591</td>
<td>1,182</td>
<td>993</td>
<td>785</td>
<td>922</td>
<td></td>
</tr>
<tr>
<td>Activities for long-term unemployed older than 24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>33</td>
<td>42</td>
<td>25</td>
<td>40</td>
<td>32</td>
<td>42</td>
<td>26</td>
</tr>
<tr>
<td>Employed, not in own business</td>
<td>35</td>
<td>33</td>
<td>37</td>
<td>44</td>
<td>36</td>
<td>30</td>
<td>33</td>
</tr>
<tr>
<td>Runs own business</td>
<td>32</td>
<td>25</td>
<td>38</td>
<td>16</td>
<td>33</td>
<td>29</td>
<td>41</td>
</tr>
<tr>
<td>Total (percent)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Total (number)</td>
<td>2,768</td>
<td>1,319</td>
<td>1,449</td>
<td>230</td>
<td>1,573</td>
<td>460</td>
<td>497</td>
</tr>
</tbody>
</table>

Six months after the end of the training

Self-assessment of capability improvements

Table 30 below displays the answers of surveyed participants to a series of questions about their capabilities:

- Job-search: Do you now know better how to look for a job than before?
- Acquired skills: Do you currently know better how to do a certain type of work than before participation?
- Self-confidence: Are you now more confident than before participation?

The table shows a very high level of positive self-assessment, something which is quite frequent with this type of questionnaire. Opinions are more positive among participants in the activities targeted at young unemployed jobseekers, in comparison with older ones.
Table 30– Improvements stated by participants (PL)

<table>
<thead>
<tr>
<th>Improved capabilities</th>
<th>Total</th>
<th>Gender</th>
<th>Education level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>Activities for young unemployed jobseekers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job search</td>
<td>73</td>
<td>73</td>
<td>72</td>
</tr>
<tr>
<td>Work competencies</td>
<td>85</td>
<td>86</td>
<td>85</td>
</tr>
<tr>
<td>Self confidence</td>
<td>83</td>
<td>84</td>
<td>82</td>
</tr>
<tr>
<td>Activities for long-term unemployed older than 24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job search</td>
<td>55</td>
<td>56</td>
<td>55</td>
</tr>
<tr>
<td>Work competencies</td>
<td>67</td>
<td>71</td>
<td>65</td>
</tr>
<tr>
<td>Self confidence</td>
<td>76</td>
<td>79</td>
<td>73</td>
</tr>
</tbody>
</table>

4.3.3 Comparative analysis (PL)

Two distinct methods were applied as follows:

- Method 0 - Analysis per type of activity, selection bias not corrected
- Method 1 - Analysis per type of activity, selection bias corrected by propensity score matching.

Method 0 – Analysis per type of activity, selection bias not corrected

Method 0 consists of comparing the surveyed participants and the control group for both types of training. This comparison applies to several variables, i.e.

- Employment status after training
- Stability of jobs
- Duration of jobs.

Table 31 below displays a comparison per employment status after six months. It shows that the share of employed people is higher (+10%; +19%) among participants than in the control group. However, the participants are clearly less successful at finding permanent jobs.
Table 31 – Impact on employment (PL/Method 0)

<table>
<thead>
<tr>
<th>Employment status after training (percent)</th>
<th>Surveyed participants A</th>
<th>Control Group B</th>
<th>Impact estimate (0) A-B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activities for young unemployed jobseekers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>42</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>58</td>
<td>48</td>
<td>+10</td>
</tr>
<tr>
<td>- full time, permanent</td>
<td>8</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>- full time, temporary</td>
<td>38</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>- part time, permanent</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>- part time, temporary</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>- specific or work order</td>
<td>6</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td><strong>Activities for long-term unemployed older than 24</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>33</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>67</td>
<td>48</td>
<td>+19</td>
</tr>
<tr>
<td>- full time, permanent</td>
<td>12</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>- full time, temporary</td>
<td>44</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>- part time, permanent</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>- part time, temporary</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>- specific or work order</td>
<td>5</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Table 31 and Table 33 describe the situation of participants after six months, while Table 33 gives an insight on what might happen to successful participants in a longer term. As regards job stability, the surveyed participants are more optimistic than the control group members in the case of long-term unemployed job seekers, but not in the case of younger participants.
Table 32– Job stability of employed interviewees (PL)

<table>
<thead>
<tr>
<th>“I will remain employed next year” (percent)</th>
<th>Surveyed participants A</th>
<th>Control Group B</th>
<th>Impact estimate (0) A-B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activities for young unemployed jobseekers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highly likely</td>
<td>52</td>
<td>57</td>
<td>-2</td>
</tr>
<tr>
<td>Likely</td>
<td>35</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td>10</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Unlikely</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Highly unlikely</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td><strong>Activities for long-term unemployed older than 24</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highly likely</td>
<td>66</td>
<td>57</td>
<td>+8</td>
</tr>
<tr>
<td>Likely</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td>18</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>Unlikely</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Highly unlikely</td>
<td>7</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

Table 33 shows how long the employed interviewees had been at work at the time of the survey, i.e. six months after the training. We see that the duration of employment is clearly longer in the control group than in the group of participants.

Table 33– Job duration of employed interviewees (PL)

<table>
<thead>
<tr>
<th>In the current job since... (percent)</th>
<th>Surveyed participants A</th>
<th>Control Group B</th>
<th>Impact estimate (0) A-B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activities for young unemployed jobseekers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 months or less</td>
<td>72</td>
<td>41</td>
<td>+31</td>
</tr>
<tr>
<td>7 months or more</td>
<td>29</td>
<td>59</td>
<td>-30</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td><strong>Activities for long-term unemployed older than 24</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 months or less</td>
<td>66</td>
<td>44</td>
<td>+22</td>
</tr>
<tr>
<td>7 months or more</td>
<td>35</td>
<td>57</td>
<td>-22</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

The interpretation of this finding is not that participants would have less stable jobs than non-participants. First, such an interpretation would contradict Table 32. Second, and even more importantly, the findings below can be explained in terms of a lock-in effect.
In fact, at the time of the survey, both groups had quite a different story:

- Participants and non-participants were equally unemployed before the training started;
- During the 3 to 5 following months, participants were in training and people in the control group were searching for jobs;
- At the time of the survey, participants had a 6-month period of job-hunting, but that period extended to 9-11 months for those in the control group.

This interpretation suggests that the analysis is strongly affected by the fact that the survey took place six months only after the end of the training. This may limit the use of the findings.

**Method I– Analysis per type of activity, selection bias corrected**

This method uses the Propensity Score Matching (PSM) technique in order to correct the selection bias (see Chapter 5). The study team has estimated the probability to participate, depending on gender, education level, and intensity of job search before training (propensity score). Then each participant in the surveyed sample was matched with the person of the control group having the closest propensity score.

The findings show a positive impact for both target groups under study, although less for young unemployed job seekers.

**Table 34 – Impact on employment (PL/Method I)**

<table>
<thead>
<tr>
<th>Employed or self-employed after training (percent)</th>
<th>Surveyed participants A</th>
<th>Control Group B</th>
<th>Impact estimate (I) A-B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activities for young unemployed job seekers</td>
<td>58.0</td>
<td>51.7</td>
<td>+6.3**</td>
</tr>
<tr>
<td>Activities for long-term unemployed older than 24</td>
<td>67.0</td>
<td>47.7</td>
<td>+19.3**</td>
</tr>
</tbody>
</table>

** - significant at 95% confidence level

Correcting the selection bias has made a major difference as regards the impact on young unemployed job seekers (6% instead of 10% - see Table 31). This is not the case for the other target group.

Table 35 below applies the same model to estimating the impact that might be observed if all participants and non-participants were successively men, women, youngsters and so on.

---

25 One person from the control group could be matched with several participants.
Table 35 – Impact on employment per public (PL/Method I)

<table>
<thead>
<tr>
<th>Additional probability of being employed after six months</th>
<th>Young unemployed job seekers</th>
<th>Long-term unemployed older than 24</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>+8.1</td>
<td>+23.1</td>
</tr>
<tr>
<td>Female</td>
<td>+3.9</td>
<td>+16.4</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic</td>
<td>+8.1</td>
<td>+19.7</td>
</tr>
<tr>
<td>Vocational education</td>
<td>+14.7</td>
<td>+19.9</td>
</tr>
<tr>
<td>Intermediary education</td>
<td>+2.7</td>
<td>+11.4</td>
</tr>
<tr>
<td>Higher education</td>
<td>-7.3</td>
<td>+20.6</td>
</tr>
</tbody>
</table>

The table shows that women systematically perform less well than do men and that the same applies to younger job seekers in comparison with older ones. The pattern of impacts per education level is unclear.

4.3.4 Lessons (PL)

Overall the analysis of the Polish dataset shows an impressive success in the case of long-term unemployed people. In this case, training increases the probability of finding a job after six months by nearly 20%. Note that this public participated in a 3-month training course, in average, most often combined with counselling for job searching and advice.

Impact is by far smaller in the case of young unemployed job seekers, although they participate in longer training activities (5 months in average).

This analysis shows that on-the-job training (i.e. internships, vocational education, and subsidised employment) is more effective for finding a job, and off-the-job training is more effective for creating one's own business.

The study also shows that women are systematically less successful than men, and that the same applies to younger job seekers in comparison with older ones.

4.4 Impact of training: Belgian country study

4.4.1 Dataset (BE)

All data were supplied by the Flemish Public Employment Services and are briefly described below.

The unemployment database consists of 821,820 persons who were unemployed at some period between January 2005 and January 2010. Information about unemployment, unemployment duration and individual background characteristics are included in the dataset. The key variables of interest are 61 monthly binary variables indicating whether a person is searching for work or not.

However, the dataset lacks some crucial information:

- The dataset does not provide any information about social benefits. A person who is not categorised as “searching for a job” may however still receive a social benefit.
It is not clear why a person is not searching for work in a certain month. The person could have found a job but various other reasons could also apply. The person could, for example, be ill or be in training and therefore not actively seeking work.

In the framework of this analysis, the “participants” consist of unemployed people who received training between January 2005 and January 2010. There are 226,181 people who belong to this category. The Employment Service supplied data on – among other things – the incidence, the frequency and the duration of the training that Flemish unemployed people received in this period. Also, the data provide information on whether the training was ESF-subsidised.

The lack of information on social benefits in the unemployment database implies that the results of the analysis cannot directly be translated in terms of a reduced or increased period in which a person receives a social benefit. It can merely indicate whether training will increase or decrease the duration of unemployment, assuming that a person is no longer unemployed if he or she is not actively looking for work anymore. This immediately relates to the second data issue: the lack of information on the reasons why a person is not actively searching for work in a certain month. This lack of information implies that certain assumptions have to be made about what is regarded as a period of unemployment. These assumptions will be explained hereafter.

The availability of such interesting time series calls for testing the Duration Model technique (see explanation in Chapter 5).

### 4.4.2 Methodological issues (BE)

**Lock-in effect**

The ‘lock-in’ effect refers to the extra unemployment duration that is directly caused by the time spent on training. While unemployed individuals who do not receive training can search for a job immediately, the unemployed who do receive training cannot start looking for a job before the end of their training.

The lock-in effect has two implications with respect to this study. The first implication is that in terms of return on investment, the lock-in effect means a longer period of unemployment benefits and other costs related to unemployment.

However, it is reasonable to assume that training increases the probability of finding work once the training is finished. In this study, the lock-in effect is neglected and the analysis concentrates on the probability of finding a job.

**Survival analysis**

The tested approach is based on Survival Analysis. This technique analyses the time at which participants and non-participants exit unemployment (see Box 6 below).

In this analysis, the duration of unemployment is determined by subtracting the first month in which a person starts searching for work from the last month of the period in which he/she is actively looking for work.\(^{26}\)

The fact that the dataset covers the years 2005-2009 raises the problem of “censoring”. Left censoring applies to the persons who were unemployed before January 2005 and still are in January 2005. Right censoring applies to

---

\(^{26}\) Assuming that interrupting job searching is synonymous with exiting unemployment, something which is not entirely true.
persons who became unemployed before January 2010 and still are after January 2010. In fact, these persons have a longer duration of unemployment than what is measured by the study team. The influence of censored observations on estimates is tested hereafter.

**Box 6 – Probability of ending unemployment (BE)**

The following formula reflects the study team’s assumptions about the probability of being out of unemployment:

\[
h_j(t | x_j) = pt^{p-1} \exp(\beta_0 + x_j \beta_x)
\]

The equation explains the probability \( h_j \) that a person \( j \) exits the unemployment database during a given month \( t \), depending on his/her participation in a training activity \( x_j \).

Irrespective of the training, the probability of being out of unemployment is assumed to increase over time, but at a diminishing pace. This is modelled by the term \( pt^{p-1} \) where \( 1 < p < 2 \). It is reasonable to assume that the probability of ending unemployment increases with time. The longer the time in unemployment the longer the period the individual is looking for work. It is moreover reasonable to assume that the chance of a successful application decreases with time because a long period of unemployment signals low productivity. In other words: the marginal probability of ending employment diminishes with time.

The end of the formula reflects the assumption that participants and non-participants have different probabilities of ending unemployment, which are reflected in the parameters \( \beta_0 \) and \( \beta_x \).

(Cleves and al., 2008\textsuperscript{27})

**Correcting the selection bias**

It is reasonable to assume that people who received training had specific difficulties in finding a job. It is thus likely that people who receive training have a higher average unemployment duration than people who do not receive training. This ‘selection bias’ may lead to an underestimation of the impact of training\textsuperscript{28}.

One of the ways to control for selection bias is the use of an instrumental variable. Such a variable should be sufficiently correlated with participation in the training, but not with the employment status after training. However, the dataset does not provide such a variable.


\textsuperscript{28} The other three country studies suggest that the reverse is more frequent, i.e. training service providers tend to select participants who are likely to succeed in finding jobs, in order to show good performances (‘creaming effect’).
Therefore, another technique had to be applied, i.e. stratified matching. Participants and non-participants were matched on the basis of the duration of unemployment at the time when participants started or ended the training. The two methods are described as follows:

- Matching before training (Method I) - Participants who start training at time $t_b$ are matched with individuals in the control group who were unemployed until at least time $t_b$. In this case, the impact estimates include the lock-in effect and the subsequent benefits in terms of finding jobs.
- Matching after training (Method II) - Participants who finish training at time $t_a$ are matched with individuals in the control group who were unemployed until at least time $t_a$. In this case, the lock-in effect is neglected and the probability of finding work after the end of the training is the main focus.

**Matching at the start of the training (Method I)**

Figure 3 below displays the distribution of participants and the control group in terms of the duration of prior unemployment (in months). For the participants, unemployment duration is measured until the start of training, and for the control group it is measured until exit from unemployment. Clearly visible is the large group of individuals in the control group who find a job in the first months of unemployment.

![Figure 3 - Duration of prior unemployment (BE/Method I - unmatched)](image)

After matching participants who start the training and who exit unemployment at the same time, the distribution of both groups is considerably changed (see Figure 4 below). The distribution of the control group is flattened out and there is no peak at the beginning of the distribution. The comparison of both figures suggests that matching has resulted in some correction for the selection bias.
Matching at the end of the training (Method II)

Figure 5 also displays the distribution of participants and the control group over time in terms of the duration of prior unemployment (in months). The unemployment duration is measured until the end of training for participants, and until unemployment exit for the control group. Again, the large group of individuals in the control group who find a job in the first month of unemployment is clearly visible.

Figure 5 – Duration of prior unemployment (BE/Method II - unmatched)
Figure 6 shows how the previous one has changed as a result of matching. Again, the distribution of the control group is flattened out and there is no peak at the beginning of the distribution.

**Figure 6 – Duration of prior unemployment (BE/Method II - matched)**

![Duration of prior unemployment](image)

4.4.3 **Descriptive analysis (BE)**

This section describes the duration of unemployment and respectively gender, age and educational level. The relation between training and unemployment duration is also described. The impact of training is not considered.

**Unemployment duration per type of public**

Figure 7 displays the distribution of men and women with respect to unemployment duration. Censored observations are filtered out. There is no large difference in unemployment duration between men and women.
Educational level is correlated to the unemployment duration. As Figure 8 shows, there is a larger share of highly-educated people that find a job in the first few months of unemployment. This share is smaller for medium-educated and even smaller for low-educated people. Moreover, the share of highly-educated people who remain unemployed for a relatively long time is lower than the share of medium- or low-educated people.

Finally, there is also a clear relationship between age and unemployment duration. Figure 9 below displays the distribution of different age categories over
unemployment duration. The younger the person, the shorter the unemployment duration will be.

Figure 9 – Duration of unemployment and age (BE)

4.4.4 Comparative analysis (BE)

Overall estimate

Figure 10 below shows the estimated survival function for participants and the control group. No matching was applied at this stage.

---

29 In this section, the survival function estimates the probability of being unemployed after t months (Kaplan Meier survival estimator). The survival function thus shows at a certain point in time t the proportion of unemployed that will be unemployed for at least that time or longer. In formula:

\[ \hat{S}(t) = \prod_{t_j \leq t} \left( \frac{n_j - d_j}{n_j} \right) \]

where \( n_j \) is the number of individuals in unemployment at time \( t_j \) and \( d_j \) is the number of individuals who are no longer unemployed at time \( t_j \). Over all observed times the product is less than or equal to t.
The probability of remaining unemployed decreases over time for both groups, but the participants have systematically less chances of leaving unemployment than does the control group.

This is no longer the case when matching methods are applied. When individuals are matched at the beginning of the training (Method I) those who have received training are more successful than the control group (lower probability of remaining in unemployment). However, this only holds for a sub-group of participants who leave unemployment between 0 and 15 months after the end of the training.
If individuals are matched at the end of the training (Method II) the two curves have the same pattern. However, the time at which the curves intersect is longer. Participants remain more successful than the control group for 22 months, as shown in Figure 12 below.

**Figure 12 – Probability of remaining unemployed (BE / Method II, matched)**

Overall, the analysis shows a positive impact of the training on the probability that an individual will become employed. However, this only holds for a subgroup of participants who benefit from the training in the first one or two years. There is also a sub-group of people who have a high probability of remaining unemployed in the long run, even after receiving training.

**Marginal impact per type of public**

This section applies the previously estimated models in order to provide impact estimates per type of public. Table 36 below lists the four methods that have been applied.

**Table 36 – Estimating impact with four methods (BE)**

<table>
<thead>
<tr>
<th>Method</th>
<th>Matching</th>
<th>Censoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No matching</td>
<td>No subjects are excluded from the analysis</td>
</tr>
<tr>
<td>0/b30</td>
<td>No matching</td>
<td>Subjects who are unemployed at t=1 or t=61 are excluded from the analysis</td>
</tr>
<tr>
<td>I</td>
<td>Matching at the beginning of the training</td>
<td>No subjects are excluded from the analysis</td>
</tr>
<tr>
<td>II</td>
<td>Matching at the end of the training</td>
<td>No subjects are excluded from the analysis</td>
</tr>
</tbody>
</table>

30 The findings of this approach are not displayed in the previous section.
All methods lead to impact estimates that are negative as far as all types of training are considered together (probably due to lock-in effects being larger than subsequent benefits). However, the impact of labour-market oriented courses is positive. The study shows that participants have a significantly reduced duration of unemployment.

Interesting gender effects have been identified. With Methods 0, I, and II, being a male involves a longer duration of unemployment. According to the estimate of Method 0, being male results in a 0.46 month longer period of unemployment, *ceteris paribus.*

Age is also related to the duration of unemployment. Unemployment duration increases with age, but the effect also becomes smaller as age increases.

Educational level is related to the unemployment duration. With Method 0, having a high educational level leads to 3.31 fewer months of unemployment, compared to an individual with a low educational level.

Having a subsistence level of income or a (labour-related) handicap does not have a large effect on unemployment duration.

With Method 0/b, the effects of training are similar but their magnitude is generally smaller than with Method 0. The fact that censored observations are filtered out of the data may result in an under-estimation because the individuals with long unemployment durations are more likely to be right-censored.

With Methods I and II (matching at the beginning/end of the training) the pattern of the findings is similar.

Table 37 – Factors affecting unemployment duration (BE)

<table>
<thead>
<tr>
<th>Individual characteristics</th>
<th>Marginal effect (months) estimated through four methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Training</td>
<td></td>
</tr>
<tr>
<td>One or several courses in period 2005-2010 (1=yes)</td>
<td>4.69*</td>
</tr>
<tr>
<td>Number of courses</td>
<td>7.84*</td>
</tr>
<tr>
<td>Number of ESF-subsidised courses</td>
<td>1.72*</td>
</tr>
<tr>
<td>Number of labour-market oriented courses</td>
<td>-4.09*</td>
</tr>
<tr>
<td>Gender (male =1)</td>
<td>0.46*</td>
</tr>
<tr>
<td>Age (years)</td>
<td>1.41*</td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>Low educational level (reference category)</td>
<td></td>
</tr>
<tr>
<td>Medium educational level</td>
<td>-1.46*</td>
</tr>
<tr>
<td>High educational level</td>
<td>-3.31*</td>
</tr>
<tr>
<td>Vulnerability</td>
<td></td>
</tr>
<tr>
<td>Subsistence level of income (1=yes)</td>
<td>0.76*</td>
</tr>
<tr>
<td>Labour-related handicap (1=yes)</td>
<td>-0.10</td>
</tr>
<tr>
<td>Heavy labour-related handicap (1=yes)</td>
<td>0.31*</td>
</tr>
</tbody>
</table>

* - significant at a 90% confidence level
4.4.5 Impact on employment status (Method III)

In addition to the analysis in terms of unemployment duration, a second analysis was performed on a cross-sectional selection of the dataset. The purpose was to apply an approach which is less demanding in terms of technical capabilities, and more comparable to what was done in the other three countries.

For this analysis, two samples were extracted from the overall database:

- People who were not registered as job seekers throughout 2005, who were looking for a job in January 2006, and who have subsequently received training (7,964 persons, of whom 2,479 underwent training that was co-funded by the ESF).
- People with the same employment history, but who did not receive training (33,832 persons).

The study team estimated four models explaining the probability that job seekers were still looking for a job at the start of 2007, 2008, 2009 and 2010.

The group of participants consisted of individuals who had received training between 2001 and the year of the analysis. Someone who first received and completed training during 2008 was therefore in the control group in the analysis for 2007, but in the sample of participants in the analysis for 2009. This meant that the treatment group was larger in the analysis for 2010 than in the 2009 analysis.

The factors analysed in this method (Method III) are not the same as in the previous ones. For instance, the number of courses, the number of ESF-subsidised courses and the number of labour-market oriented courses are excluded from this analysis for technical reasons.

Table 38 below shows the estimated parameters of the model.

Table 38 – Impact on employment (BE/Method III)

<table>
<thead>
<tr>
<th>Individual characteristics</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation in training</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (control group)</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Yes</td>
<td>-.397***</td>
<td>-.500***</td>
<td>-.676***</td>
<td>-.615***</td>
</tr>
<tr>
<td>Yes (ESF-subsidised training)</td>
<td>-.083</td>
<td>.058</td>
<td>-.085</td>
<td>-.102*</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Female</td>
<td>-.047*</td>
<td>-.034</td>
<td>.083***</td>
<td>.183***</td>
</tr>
<tr>
<td>Age</td>
<td>-.017***</td>
<td>-.013***</td>
<td>-.012***</td>
<td>-.008***</td>
</tr>
<tr>
<td>Level of education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Medium</td>
<td>.407***</td>
<td>.470***</td>
<td>.495***</td>
<td>.494***</td>
</tr>
<tr>
<td>High</td>
<td>.748***</td>
<td>.757***</td>
<td>.795***</td>
<td>.880***</td>
</tr>
<tr>
<td>Vulnerability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subsistence level of income (1=at or below subsis-</td>
<td>-.239</td>
<td>-.473**</td>
<td>.192</td>
<td>.017</td>
</tr>
</tbody>
</table>

31 Multivariate logit models.
32 Including these variables along with a dummy variable for treatment versus non-treatment leads to collinearity.
The table shows that the likelihood for participants to exit unemployment is lower than for the control group, especially in 2009 and 2010. Whether or not training was subsidised by the ESF does not make much of a difference according to the analysis.

Women interestingly have an estimated lower probability of success in the analysis for 2007, but a higher likelihood of finding employment in 2009 and 2010.

Age has a negative effect on the probability of exiting unemployment.

Job seekers with a medium level of education have a higher probability of success than those with a low level of education. For job seekers with a high level of education, this difference is even greater.

There is a positive impact for people with a labour-related handicap in the 2007 and 2008 estimates. Note that the number of hours of training received is higher for job seekers with a labour-related handicap, which might explain this positive coefficient.

Some regions differ significantly from the reference category (the region of Tongeren is the base category). Job seekers from Antwerp have a lower probability of exiting unemployment. Although this could be related to the local labour market, this is probably the result of selection. On average, job seekers in Antwerp probably have a larger distance to the labour market than job seekers in Tongeren and many other parts of Flanders. Job seekers in the region of Kortrijk – Roesselare have a higher probability of success, according to the 2007 and 2008 estimates. This effect disappears in the estimates from the 2009 and 2010 analysis.
4.4.6 Lessons (BE)

The lessons learned from the analyses carried out in Belgium are:

- In general, training has a negative effect in terms of the desired achievements on the labour market, something which is probably due to lock-in effects being greater than subsequent benefits.
- However, training reduces the unemployment duration for a sub-group of participants who find a job within the first or second year after the end of their participation.
- Labour-market oriented training does not show the above quoted negative impact but it does not reduce the unemployment duration either (see Table 37).
- There are some less negative impacts for women and people with a labour-related handicap, but these findings are not stable over the years.

4.5 Impact of training: lessons learned across four countries

4.5.1 Synthesis of the country studies

Impact of training in general

All four country studies analyse the impact of training for unemployed people in terms of change in their employment status within a few months (or years) following the training.

Impact differs considerably between the four countries:

- In Belgium, the assessed impact is the number of months before exiting the files of the Public Employment Service (a proxy of exiting unemployment). It is generally negative, i.e. participants stay longer in unemployment than the control group. However, such a negative impact is not observed in the case of labour market oriented training programmes.
- In Hungary, the assessed impact is the probability of being employed after 24 months. The impact estimate is about +5%.
- In Poland, the assessed impact is the probability of being employed after 6 months. The impact estimates range from 5 to 20%, depending on the targeted public.
- In Italy, the assessed impact is the probability of being employed after 12 months. The impact estimate is near to 30%.

The studies carried out in Belgium and Italy end in surprisingly small and large impact estimates, and deserve to be discussed in terms of reliability.

In the case of Belgium, the study finds that the impact of training is negative or null, and this is obviously due to the lock-in effect. Beyond this first explanation, it can also be said that the study may not have controlled an important factor. In fact, the impact estimate derives from a database containing all the unemployed in Flanders, among which a relatively large number of people are unemployed only for a short period, for instance because they are in between two jobs. This means that the participants in ESF supported training were compared with a control group having a smaller distance to the labour market, and this factor could not be controlled in the comparison because the information processed in the study was not rich enough.

On the other hand, the Italian study finds that the participants are almost two times more likely to be employed than the control group. This would be an exceptionally successful impact if the analysis was fully credible, but the study
team has some doubts about the reliability of the estimate, especially because a large part of the dataset originates from performance measures which might be biased (see Box 5).

Considering that Belgian and Italian estimates are probably not reliable, and building on the findings of the Hungarian and Polish country studies the authors of this report consider that the impact of training on the probability of being employed is in a range of +5 to +20%. These figures are slightly more optimistic than that of the literature review which suggests a bracket of [0-20%] (see Table 3).

**Types of publics**

Impact estimates are more positive for women than for men in Hungary, and less so in Poland. The situation is mixed in Belgium. These findings are less conclusive than that of the literature review, which suggest that women are more successful than men in general.

In both Poland and Hungary, specifically positive impacts have been found for long-term job seekers, and younger trainees show a clearly lower probability of finding a job, although still positive, but this is not visible in Belgium. These findings are consistent with that of the literature review.

**Types of training**

In Belgium and Poland, there are converging indications that training is more effective in terms of return to work if it is (1) labour-market oriented, and (2) associated with counselling. These findings are fully consistent with that of the literature review.

### 4.5.2 Concluding on the return on investment

**Investing in employed people: an open question**

This study could not conclude on the social profitability of investing public money in training employed people. It is clear that there is a rationale for taking public action at the level of industries rather than firms. How should such actions be structured remains an open question, and especially the following sub-questions:

- Should public policies be based on regulations (obligation to train) or public investment through financial incentives?
- Under which assumptions would public investments be socially profitable?
- Which is the relevant level for designing such a policy (regional, national, European)?
- If the relevant level for policy-making is not that of the EU as a whole, then is there a residual need for taking action at EU level? ... and if yes, what should be the instruments (soft coordination, financial incentives, or both)?

---

33 In Italy, there is a discrepancy between this study and previous Italian evaluations building upon the same dataset
Investing in the unemployed: a reverse calculation

The research synthesis did not reveal any clear picture concerning the return on investment (see pages 11 and 25), an issue which is assessed as a knowledge gap. Therefore, the study team carried out a series of reverse calculations showing the conditions that need to be fulfilled for a training programme to be profitable enough from society’s point of view (see Appendix 9).

Considering the coverage of ESF-supported activities, three types of investment would deserve to be studied in this section:

- Training of long-term unemployed with a return-to-work purpose;
- Training of low-skilled employees in at-risk enterprises in order to prevent unemployment;
- Support to training systems at sector level in order to compensate for underinvestment at the level of individual firms.

In fact, only the first type of training is considered in this study because:

- It is relatively well known in terms of benefits, and is covered by the four country studies of this report
- Both other types of training correspond to knowledge gaps.

Basic assumptions

Training programmes targeted at the unemployed may be combined with other types of activities such as counselling or job search assistance. The literature review indicates that such a mix of instruments is the most effective approach. This is why the calculation applies to a 4-month training course for a group of 17 long-term unemployed individuals who also benefit from 9 months of individual coaching.

The costs of such training are assumed to be €47,000, i.e. €2,700 per participant, including the cost of running the training system.

The success rate, i.e. the proportion of participants who find a job, is assumed to be in a range of 5 to 15%.

The individual benefit for successful participants is assumed to be €1,200 per month.

Another key assumption is the duration of the benefits, which may be related to ageing, skills obsolescence, and business cycles. This duration is assumed to be in a range of 3 to 10 years.

Scenarios with unknown substitution

In a first series of scenarios, substitution (i.e. the fact that someone else becomes or remains unemployed for the very reason that a participant found a job) is assumed to be in a range of 0 to 100%.

In these scenarios, social benefits are assumed to amount to 50% of individual benefits. These benefits include satisfaction with having a job (see page 25), plus the social benefits arising from the fact that someone exits unemployment, e.g. improvement of family and social relationships, reduction of informal economy and anti-social behaviour.

Despite the fact that some pieces of research are available such as Almeida and Carneiro (2009) who find a 8.6 percent rate of return on training investments in large Portuguese firms (see 3.1.2), and Jesperen (2008) who finds a negative profitability in the case of a Danish return-to-work training programme.
Figure 13 below displays the findings of the reverse calculation in the case of the above assumptions. The graph should be read as follows:

- The two coloured curves correspond to the success rates: 10% and 15%. Social returns cannot reach the desired threshold with a success rate of only 5%.
- The horizontal axis corresponds to the duration of benefits: from 3 to 10 years.
- The vertical axis corresponds to the proportion of substitution among successful participants.
- In the area below a given curve the return on investment is higher than 4%.

Depending on the above assumptions, a satisfactory return on investment could be achieved in the following cases:

- Point A on Figure 13 below: optimistic assumptions in terms of duration of benefits (8 years), reasonable assumption in terms of success rate (10%), optimistic assumption about substitution (20% only).
- Point B on Figure 13 below: same duration of benefits (8 years), more optimistic assumption about success rate (15%), pessimistic assumption about substitution (50%).

Both sets of assumptions above are consistent with the literature review and the country studies. This means that there are several sets of reasonable assumptions under which public investment in training the long-term unemployed is socially profitable.

**Figure 13 – Conditions for achieving good returns (first scenarios)**

---

**Scenarios with unknown social benefits**

All assumptions are the same as in the previous scenarios, except the following ones:
Social benefits are no longer fixed at 50% of individual benefits. On the contrary, they may vary in a range of 10 to 120%.

Substitution is assumed to be 50%.

Figure 14 below displays the findings of the reverse calculation in the case of this second series of scenarios. The graph should be read as follows:

- The two coloured curves correspond to the success rates: 10% and 15%. Social return cannot reach the desired threshold with a success rate of only 5%.
- The horizontal axis corresponds to the duration of benefits: from 3 to 10 years.
- The vertical axis corresponds to the proportion of social benefits in relation to individual ones.
- In the area above a given curve the return on investment is higher than 4%.

Depending on the above assumptions, good returns on investment could be achieved in the following cases:

- Point A on Figure 13 below: optimistic assumption in terms of duration of benefits (8 years), optimistic assumption about success rate (15%), reasonable assumption about social benefits (50%)
- Point B on Figure 13 below: same duration of benefits (8 years), reasonable assumption in terms of success rate (10%), very optimistic assumption about social benefits (120%).

Again, these sets of assumptions are assessed by the study team as reasonable.

**Figure 14 – Conditions for achieving good returns (second scenarios)**

**Scenarios with unknown cost-effectiveness**

In these scenarios, the duration of the training is in a range of 1 to 7 months and the cost varies accordingly, while the success rates remain the same as
previously assumed, i.e. 10% or 15%. Consequently, these scenarios explore the issue of cost-effectiveness.

The main other assumptions are the following ones:

- Social benefits amount to 50% of individual benefits.
- Substitution is assumed to be 50%.
- The duration of benefits varies from 3 to 8 years

Figure 15 below displays the findings of the reverse calculation in the case of this third series of scenarios. The graph should be read as follows:

- The two coloured curves correspond to the success rates: 10% and 15%. Social return cannot reach the desired threshold with a success rate of only 5%.
- The horizontal axis corresponds to the duration of benefits: from 3 to 10 years.
- The vertical axis corresponds to the cost of the training, which is approximated by the number of months.
- In the area below a given curve the return on investment is higher than 4%.

Depending on the above assumptions, good returns on investment could be achieved in the following cases:

- Point A on Figure 13 below: optimistic assumption in terms of duration of benefits (8 years), reasonable assumption in terms of success rate (10%), optimistic assumption in terms of cost-effectiveness (same benefits achieved at half the cost of previous scenarios).
- Point B on Figure 13 below: same duration of benefits (8 years), more optimistic assumption about success rate (15%), reasonable assumption about cost-effectiveness (same benefits achieved at the same cost as that of previous scenarios)
The above scenarios suggest that there is a cluster of reasonable assumptions under which training the unemployed is socially profitable, and which can be summarised as follows:

- The success rate (additional probability of finding a job) should be in a range of 10-20%. The smallest figure is a profitability threshold and the largest figure is the most optimistic assumption that can be reasonably made (see 4.5.1).
- The training programme should be combined with job-search assistance and its duration should be in a range of 2 to 4 months. Other types of training might also be profitable, provided that their cost does not exceed that of a 2-4 month programme (see Figure 15).
- The duration of benefits should be 8 years or more, an assumption which seems to be consistent with recent and credible research works\textsuperscript{35}.
- No more than 20 to 50% of the participants having found a job should have succeeded at the expense of someone else (substitution). The largest figure is a profitability threshold and the smallest figure is a very optimistic assumption. However the study team cannot justify these assumptions which are an absolute knowledge gap.
- Social benefits should amount to at least 50% of individual benefits (i.e. gross salary of the job found), and this profitability threshold may have to be increased to 100% or more if other assumptions are pessimistic. In this area also, the study team cannot justify these assumptions in the absence of accumulated knowledge.

Looking at the country studies in the light of the above assumptions, it seems that ESF supported training was:

- Socially profitable in the case of the long term unemployed in Poland;
- Not profitable in Hungary and Belgium in general;
- Not profitable for young people in general.

At the end of this chapter, and despite the shaky nature of the available evidence, the study team concludes that training the unemployed may be profitable from a social point of view, but this should be considered as an exception rather than a rule.

In order to justify public investments in training, policy-makers and programme managers should demonstrate that they control the factors that make training programmes socially profitable. In particular, they should:

- Avoid targeting young unemployed people;
- Combine training and job-search assistance;
- Keep the cost of training reasonable, i.e. equivalent to that of a 2 to 4 month programme;
- Ensure that subsidised programmes have a high enough quality so that they can be assumed to reach a success threshold of 10%;
- Ensure that the issue of substitution is paid sufficient attention so that it can be assumed to remain under 50% (proportion of successful participants who find their job at the expense of someone else).

\textsuperscript{35} In a recent and credible study carried out in the USA the annual benefits were shown to grow constantly from year 1 to year 4, suggesting that they remain significantly positive over a longer time period (reported in C.Heinrich’s contribution to the Conference ‘Shaping the Future of the ESF’ in June 2010).
Even if such principles are applied, the profitability of public investments in training remains conditional to the following factors which cannot be controlled:

- The duration of benefits should be 8 years or more, which means that there should be a long enough time period before ageing, skill obsolescence, and business cycles reduce the employability of successful trainees.
- Social benefits should amount to at least 50% of individual benefits (i.e. gross salary of the job found).

In the case of the two above conditions, there is a complete lack of knowledge. As a consequence, it would be difficult to justify public investments in training the unemployed if no effort is made in parallel towards bridging these knowledge gaps.
5 Methodological issues

This chapter introduces the main concepts and techniques which pertain to the evaluation of the costs and benefits of public investments in human capital. It draws on the literature review (Task 1), the discussions in the expert panel (Task 2), and the experience gained by the study team in carrying out the four country studies (Tasks 3 to 6).

Section 5.2.2 presents the study team’s view on how analysts could make their methodological decisions in the context of evaluating the impact of training.

Section 5.4 offers a historical perspective on methodological issues, drawn from the literature review.

5.1 Return on investment

As shown in this section, the very purpose of this study (assessing the return on ESF investments in human capital) was certainly too ambitious with respect to the current state of knowledge. This is why the study team adopted an alternative approach based on scenarios in Section 4.5.2.

This section shows the distance which is still ahead on the way towards justifying public investment in training by a review of costs and benefits.

5.1.1 Financial and social return

The “return on human capital investments” is sometimes understood as just the impact of the training without any consideration of the money invested. This approach is followed by most of the literature on the subject, which focuses on selected key effects such as wage increases after one or two years. In practice, the benefits are difficult to assess over several years and many studies have therefore focused on the short-term benefits. This may of course yield misleading results if benefits are continuing over time, which seems to be the case (Card et al. 2009).

In this study, a wider perspective is chosen and the approach is that of the rate of return on investments, i.e. the discount rate at which the cost and the sum of discounted benefits are equal.

The rate of return may be determined in either financial or social terms.

- Financial return indicates the profitability of an investment for the private investor only, taking into account only the benefits accruing to the investor, e.g. individual persons investing in their own education or companies investing in the training of their staff.

- Social return takes into account all costs, either public or private, and the benefits accruing to society as a whole. In the case of company training, the costs to the company and the combined net benefits accruing to the employer and the employee, as well as to other affected stakeholders, shall be taken into account.

There is a rationale for public-sector investment if there is a positive social return, and if private actors assess their financial return as insufficient. In such cases, there is a tendency for companies and individuals to invest less than

36 The social rate of return is also called the economic rate of return.
what would be justified from the standpoint of society. There are several assumptions supporting the idea of underinvestment in the private sector:

- Healthy firms do not invest enough because a part of the benefits will be lost if the employee moves to another company;
- Firms facing difficulties of competitiveness cannot invest in training, something which accelerates skills obsolescence in a vicious circle;
- Disadvantaged or older people do not participate in training because nobody trusts in the return on the investment, possibly including themselves.

However, other assumptions may suggest that under-investment in training may be less severe than previously believed. For instance, recent literature in experimental economics shows that individuals are motivated not only by monetary gains but also by reciprocity and fairness. Typically, this is ignored in standard human capital models. Furthermore, market imperfections may give employers more incentives than previously thought to invest in general training of their employees. The rationale for public investment in human capital should therefore be checked carefully.

5.1.2 Costs and Benefits of human capital investments

Costs

Costs of training programmes consist of direct expenses for the trainer, premises and various training tools and inputs, as well as indirect costs of participation in terms of income forgone or lost production and earnings. Most research focuses on the benefit side and in fact very few have estimated the programme costs. This is surprising, as such information is often easily accessible, but information on training duration, which is often indicated, may be seen as a reasonable proxy of the cost (Card et al. 2009).

Cost of the programme

An assessment of the returns to human capital investment is only feasible when the cost of the programme is available. The country studies show that information on costs is available only at the aggregate level, or else missing. None of the data contain information on costs at the level of the participant. This to some extent limits the evaluation of the ESF programme. For instance, questions such as what measures work for whom (in what period) can only be answered partially.

It is hard to make an explicit link between budgets and expenditures. In most countries the active labour market programme is financed using different sources. In the micro data it is often not clear whether a participant is on a programme that is financed as part of an ESF-programme or from another source. For instance in Flanders, budgets are allocated on an ex-post basis. Therefore, it is hard to measure the added value of ESF investment on the return to human capital investment.

Benefits

The benefits from training activities may be assessed in terms of increased productivity and earnings that is achieved as a result of the training.

For the individual training participants, the main benefits are the resulting knowledge and capabilities in relation to their job functions which may lead to and be measured as increased wages, better job opportunities, and increased satisfaction and self-esteem.
Benefits for the unemployed

For an unemployed person the direct impact may be measured as employment in a suitable job or the resulting increase in income over and above the unemployment benefits that the person might alternatively be entitled to. The salary obtained may furthermore be higher than if the same unemployed person had been offered a job without having received the training, which is important for the estimation of the value of the benefit in a further time perspective. In addition to this, a full inventory of benefits must also include increased self-confidence and the value to the person of having a job, and reduced leisure time. The effects such as the possibly more active job searching are intermediary ones and should not be counted as benefits per se.

Benefits of company training

For the employers who pay for the training of their staff, the direct benefit is the increase in productivity, plus indirect benefits that may be reaped in terms of team-building, job satisfaction, improved social relationships, and so on. The productivity benefits to the employer may be reduced to the extent that they are shared with the employee who may be offered higher wages after the training. Of course, direct benefits to the employer may completely disappear if trained employees move to other firms or retire.

The benefits enjoyed by staff receiving training from their employer, on the other hand, is the increased salary they receive beyond what they would otherwise have obtained, whether they get it from the current employer or from any other future employer. In addition to this, increased self-confidence, which is more difficult to quantify, may also be a benefit resulting from the training.

Benefits for society as a whole

The picture is more complex where the focus is on the value of human capital investments to the economy as a whole. In such a context, all positive and negative impacts on all actors need to be taken into account without double counting. For instance, if a participant exits unemployment, there may be indirect benefits such as a reduction in health care and anti-social behaviour, assuming that workers are usually healthier and socially integrated than unemployed.

In case of training programmes for unemployed persons that are leading to the placement of participants, this positive impact may be partly or fully due to a substitution effect. The persons who have received the training may have done so at the expenses of others who have not participated, and when seen from the macro level the net result may be nil or at least lower than an estimate based on the increased earnings of the persons who received the training and found a paid job.

In the case of company training, the full return to the economy should include all productivity increases, whether these are gained by the current or future employer or by the employee in terms of wage increases.

Benefit profile over time

The threat effect is the positive effect of increasing employment among potential participants who make extra efforts to avoid being forced to take part in an ALMP. This effect is often difficult to identify and quantify.

37 The administrative costs of the Public Employment Service and the Social Security Board may also decrease.
The lock-in effect is the (necessary) negative effect on the probability of finding employment during participation in the programme. The resulting net benefits may vary over the short, medium and long term, and may at a later stage be reduced by the phenomenon of skills obsolescence, whether it is enjoyed by trainees, employers or society as a whole. A full picture of the returns on human capital investments must of course take all these potential, relevant impacts into account.

**5.1.3 Rationales and determinants of human capital investments**

Theoretically, investments shall be made to the extent that the expected marginal returns to the investment equals the expected remuneration from other alternative investments, at least as seen from the investors and the decision makers’ point of view. This assumes, however, that all investors are fully informed on their options and the consequences of these in terms of both costs and benefits over time.

For human capital investments, there are good reasons to believe that the returns to investments and in particular the economic returns to society as a whole will vary over time and across countries and sectors, even if the duration and quality of the training would be assumed to be the same. There are many reasons for this.

**Different training investment behaviour**

Some human capital investments are made by public authorities, namely in the case of training of the unemployed, while company training of staff is offered by private companies. These two groups of decision-makers may pursue different investment behaviour and their implicit demand for returns on investment may be different. Additionally, as mentioned above, the economic returns to society of training paid by an employer may be different from the returns to the investor.

**Specific factors related to company training**

For company training of own staff, the impacts and the returns on investment must be expected to depend on a number of factors, of which the main ones are explained below:

- The HR policy of the company as such and the reasons for providing the training may differ considerably among companies. The company may need targeted competencies, or it may offer training as a fringe benefit to its employees. In general, the company context in which the training is given may affect its impact. In particular, the way in which the cost of training is shared between employer and employee does affect the rate of return.
- Technological development, which will be different in different sectors of the economy, may of course affect the payoff period and the value of the training.
- The staff turnover rate, the age of the participant and the age at which he or she left the labour force may have some influence on the expected pay-back period and hence the expected returns on the investment.
- General and specific training may or may not have different impacts, but there could be reasons for the employer to believe that the risk of losing the employee to another employer as a result of the training is higher for general than for specific training.
DG EMPL - Study on the Return on ESF Investment in Human Capital

- General employment, the labour market situation and the structural development of the economy may affect the short- and long-term value of the training to the employer and to other stakeholders very differently.
- Existing competencies and the adaptability of the employees receiving the training may vary and give reasons to expect different returns on the human capital investment from different groups of participants.

Specific factors related to training of the unemployed

For training of the unemployed, the impacts and hence the returns must be expected to depend on:

- The institutional context and environment creating incentives and disincentives to look and apply for a job. The incentives may therefore vary among labour market groups with different wage levels and cultures, and among countries with completely different institutional set-ups.
- The labour market demand for specific competencies, which is affected by:
  - the changing economy from recession to a booming economy
  - the structural development that creates new "dying sectors" and new "growth sectors"
  - technological development.
  All of this may change the overall demand for training and the demand for specific competencies.
- The existing competencies and the general employability of the persons receiving the training, as well as their general labour market experience and behaviour (before and after the training), may vary and give reasons to believe that the returns on the human capital investment will differ among groups of participants.
- The way the training is given in a package or not may also influence the impacts, and the resulting employment may be different for training given purely as training and training given in a package of support and advice to the unemployed.

These aspects are all important to take into account not only in an ex ante assessment of returns on human capital investments but also for an evaluation of different ex post assessments of different training programmes. They are therefore indications of parameters that should be focused on in a study of the returns to human capital investments.

5.2 Impact analysis

5.2.1 Methodological findings

Findings from the literature review

The literature review suggests a number of issues which are related to data collection or data availability and which deserve to be mentioned:

- Access to monitoring and/or statistical databases has been a major constraint for many years but, increasingly, data have been made available with a high level of detail and covering large samples and time periods. This is the case for both employer-sponsored training (Asplund, 2005) and public training programmes for the unemployed (Kluve, 2006). The increased data availability has made it possible to control for various factors in the analyses and thereby further improve the resulting estimates.
The lack of data on costs of training is an important barrier to estimation of returns on human capital investments, but this problem seems to be caused by the lack of researchers’ interest rather than by a technical limitation.

**Opinion of the expert panel**

The experts were asked to reach a consensus, as far as possible, on a series of issues. As far as methodology is concerned, the panellists agreed on the following points:

- Data quality is of the utmost importance;
- Privacy issues should be manageable and should not restrict the development of sound impact evaluations;
- RCT is likely to produce strong evaluation evidence and there is no justification for neglecting this method, as has almost been the case up to now in Europe;
- The impact of training should be assessed in the long run, given that there are a range of short-term negative effects (e.g. foregone productivity, lock-in);
- The transferability of lessons learned in evaluation is likely to be possible, although it must be done carefully with caveats.

Some methodological challenges are still ahead:

- Nature and magnitude of the main general equilibrium effects (e.g. substitution effect)
- Specific areas of under-investment in training by the private sector, especially as regards at-risk employees
- Connecting the costs and impacts of training into some cost-effectiveness or cost benefit analysis.

**Findings from the four country studies**

**Constructing the counter-factual**

The content and quality of the counterfactual dataset increases the possibilities and the quality of the analysis. In the countries that use survey data, i.e. Italy and Hungary, the counterfactual dataset was constructed after the programme was finished, by taking the unemployed from another survey or an administrative database (external control group). A disadvantage of external control groups is that the definition or the classification of variables can differ between treatment and control group. This makes it hard to compare the two datasets.

In Poland the counterfactual was constructed for the sake of the evaluation (internal control group). Hence, the treatment and control group where given similar questionnaires. This increased the possibilities to apply different kinds of methods and improved the quality of the analysis.

Constructing a counterfactual has a cost, and this means that impact analyses need to focus on a few carefully selected areas. For instance, in Poland five actions were monitored under the ESF programme, but only two of them could be subjected to an impact analysis. Unfortunately no control group was available for the training programme targeted at the employed at risk of losing their jobs.

In the case of Belgium, both participants and non-participants were drawn from an existing database, without any additional survey. This approach is potentially very cheap and flexible, but in practice entailed a serious methodological limitation since the comparability of the two samples was put into question (see 4.5.1).
Information on actual human capital investment

Only limited information is available regarding the quality of training. In most countries, apart from Hungary, information on the duration of training is available. The Italian dataset also contains information on whether the training was completed successfully. None of the datasets includes detailed information on the nature of training actually supported.

Impact indicators

In general, the datasets include information on the employment status after training, and this information reflects the objectives of the ESF. Since all programmes under study targeted unemployed people, their success was assessed in terms of return to work. However, other objectives could have been considered, such as enhancing social inclusion or equal opportunities.

The survey data from Poland, Italy and Hungary contain information on the employment status at a certain moment after completing the training, respectively 6, 12 and 24 months. The Polish dataset also contains interesting qualitative measures. For instance, whether the participants have improved their job finding skills and whether they are more confident owing to the programme. It also contains the respondents’ judgement on the quality of their new job.

In the case of Flanders, the fact that unemployed people had found a job was not known, and a proxy had to be used: the fact that they exited the files of the Public Employment Service. This limitation was largely balanced by the fact that impacts could be analysed over a long period, something which is quite relevant in the context of the ESF.

An open question is how long the observation period should be to assess the return on investment in human capital. In Poland, for instance, impact could be assessed only six months after the training and this was considered as too short. A longer observation period also allows for the quality of job placement to be taken into account in an evaluation. A promising option would be to combine a survey after 12 months and the use of an administrative database over several subsequent years.

5.2.2 Managing methodological choices

This section displays the various options to be taken in order to design a sound impact analysis in the context of the ESF. Acknowledging that there is no absolute hierarchy of methods, the study team followed the logic of a decision tree with successive couples of branches.

Evaluating impacts with / without a counterfactual

A counterfactual is a quantitative estimate of what would have occurred in the absence of the evaluated intervention. Through a comparison with what did actually occur, the counterfactual provides a quantitative impact estimate. It can be developed in several ways:

- Randomised control group created by the evaluators in the case of an experimental design;
- Internal control group created by the evaluators and matched with a sample of participants (e.g. case of Poland in Section 4.3);
- External control group extracted from a statistical database and matched with a sample of participants (e.g. case of Italy in section 4.2);
- Simulation on the basis of a model (e.g. case of Flanders).
Is it always possible to develop a counterfactual? If not, where are the boundaries delineating the area of relevance of this approach to impact evaluation? And is there an alternative approach so that a professional impact evaluation can be carried out without a counterfactual? Here are some answers to these difficult questions:

- A counterfactual is feasible if, and only if, the intervention (or “treatment”) is implemented in a uniform way and if it creates a substantial change in a “system” which is stable enough to be predictable at least in probabilistic terms. For instance, a counterfactual would not make sense in the case of a programme providing small subsidies (marginal change) to encourage emerging local partnerships (unstable systems) to take innovative training initiatives (no uniform implementation).
- Another approach is also available for doing a professional impact evaluation, i.e. a “theory driven” evaluation. It relies upon a systematic test of cause-and-effect assumptions, most often in the framework of a case study design. An example of this approach is provided in the next chapter of the present study, which presents a challenging impact analysis, i.e. assessing the extent to which Managing Authorities have learned from previous impact evaluations.
- Both approaches (counterfactual and theory-based) have the capacity to draw transferable lessons, although through very different types of reasoning. The area of relevance of theory-based evaluation is much wider by far, but only a counterfactual can provide a quantitative impact estimate.

The following sections describe a range of techniques relying on counterfactuals.

**Modelling vs control group**

In order to compare a sample of participants and a control group, it is necessary to know the employment status of both groups before and after. This is done through various combinations of surveys and extracts from monitoring or statistical databases, always at specific dates, e.g. beginning of training and 12 months after finalising the training.

However, most of the returns on human capital investment materialise through long and changing individual stories which cannot be meaningfully summarised in just two pictures. Typical ESF recipients experience successive jobs of various types, unemployment periods, and participation in various human capital investments. The impact of a given training course tends to be small or even negative in the first months (lock-in effect, foregone productivity), and then to unfold over the years (placement, better earnings, increased productivity), before decreasing again (skills obsolescence).

The study team considers that a very relevant way of analysing such impacts is the duration model (see 5.3.1 and 4.4.2) which is tested in the Belgian country study. However, this approach is quite demanding in terms of data availability, and may be feasible in only a small number of circumstances, thus leaving a wide space for all the other approaches relying on control groups.

---

38 This may be a human body in the case of testing a medicine, an active person in the case of assessing a training approach, a class in the case of assessing a primary education method, an enterprise in the case of assessing a given incentive.

39 A few studies reviewed in the research synthesis also rely on duration models, e.g. Zhang (2003) and Cockx (2003).
**Experiment or quasi-experiment**

If a control group is considered, then the study team recommends that the feasibility of an experimental design (see 5.3.2) be assessed. The first condition is that the analysis be part of a well thought-out multi-annual evaluation programme (because experiments tend to be long and difficult to connect to the policy-making agenda), and the second condition is that ethical problems be solved.

**Internal or external control group**

If a quasi-experiment is considered, then the study team suggests that an internal control group be developed as in the case of the Polish country study (see 4.3) because this approach enables the evaluator to collect a rich set of fully comparable data on both participants and non-participants. This is a precondition for applying the matching techniques in the most satisfactory way (see 5.3.3).

If an internal control group cannot be developed, then the analysis needs to rely on an external control group drawn from one or several databases.

**Controlling the selection bias**

Several techniques are available for controlling the selection bias, e.g. matching (see 5.3.3), instrumental variable (see 5.3.4), regression discontinuity (see 5.3.5), fixed effects (see 5.3.6).

In practice, the selection bias leads to overestimation of impacts by 0% to +10% in the country studies, the latter figure being of the same order of magnitude as the impact itself. The country studies do not show any pattern which could help in predicting the importance of the selection bias. Moreover, the selection bias might even result from deliberate implementation decisions in the case of the so-called "creaming effect".

The research synthesis confirms the study team’s views. Basic approaches have been criticised on the grounds that they suggest excessively positive effects of training (e.g. on employment, wages, productivity). Asplund (2004) mentions a number of examples illustrating this point. For instance, a Norwegian study showed a 5% wage effect of a very short training programme, which is on a par with one year of education (Schöne 2004). The analyses suggested that most of the high returns were due to an uncontrolled selection bias. Trainees seemed to have some favourable unobserved characteristics correlated with wages. After controlling for this, the returns to training fall to approximately 1 per cent (which is still relatively satisfactory for short training).

As a consequence, there is no alternative to controlling the selection bias and the study team recommends that evaluators do not rely on basic approaches which fall short of that, such as simple comparisons of participants and non-participants. Ordinary multiple regression analysis\(^{40}\) should be avoided as well.

**Impact analysis or impact indicator**

It goes without saying that relying on impact indicators, the ultimate branch of the tree, is no longer acceptable.

---

\(^{40}\) In this approach, the employment status after training is explained by a model including participation (in the form of a yes/no variable) and a number of socio-demographic characteristics.
In the country studies, the picture given by impact indicators overestimates the actual impact by a factor ranging from 2 to 9 (see Table 39). Moreover, the study suggests that there is nothing like a constant relationship between impact indicators and actual impacts. This means that an impact indicator cannot even be considered as a good predictor of the actual impact.

Table 39 – Impact indicators vs impact estimates

<table>
<thead>
<tr>
<th>Country</th>
<th>Impact Indicator</th>
<th>Impact Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hungary (Table 11)</td>
<td>being employed after 24 months</td>
<td>45.7 + 5.6</td>
</tr>
<tr>
<td>Italy (Table 22)</td>
<td>being employed after 12 months</td>
<td>70.7 + 29.5</td>
</tr>
<tr>
<td>Poland – Young people (Table 34)</td>
<td>being employed after 6 months</td>
<td>58.0 +6.3</td>
</tr>
<tr>
<td>Poland – Long term unemployed (Table 34)</td>
<td>being employed after 6 months</td>
<td>67.0 +19.3</td>
</tr>
<tr>
<td>Poland – Young people (Table 32)</td>
<td>having a stable job after 6 months</td>
<td>87.0 -2.0</td>
</tr>
<tr>
<td>Poland – Long term unemployed (Table 32)</td>
<td>having a stable job after 6 months</td>
<td>73.0 +8.0</td>
</tr>
</tbody>
</table>

Knowing that sound impact analyses require time and resources, and cannot apply to all impact indicators at any time, the Commission should be extremely cautious in using impact indicators on an annual and extensive basis, either for reporting purposes or for managing performance.

5.3 Impact analysis techniques

The next sections introduce some of the above-mentioned techniques.

5.3.1 Duration model

In a duration model (Abbring and van den Berg, 2003), not only information on whether an individual participates in a programme or not is considered, but also the timing of the treatment within the unemployment spell. Duration or survival analysis encompasses a wide variety of methods for analyzing the timing of events. In essence we are interested in the conditional probability of an event occurring in the next time period, for instance outflow from unemployment, given that the event has not occurred so far. Some duration models are estimated on the basis of specific assumptions about the

---

41 Except indicator of job stability in Poland
probability distribution\textsuperscript{42}. We may also use a non-parametric distribution and let the data show the exact distribution.

Duration models are extremely demanding in terms of data since the employment status and the training activities of all participants and non-participants should be known on a monthly basis over several years. In the Belgian country study, this was made possible because such data were recorded by the Public Employment Service and made accessible to the study team. It is also worth noting that the approach is relatively demanding in terms of technicality and data-processing capacity.

\textbf{5.3.2 Experimental design}

The experimental design, also called Randomised Control Trial (RCT), relies upon a random assignment to either a treatment group or a control group. The random assignment makes sure that members of the treatment group do not systematically differ from members of the control group except for the training and that access to the treatment group is not selective. Under the additional condition that leakages do not create a bias (attrition bias), and assuming statistical validity, the difference between average before-after change in the employment status of the treatment group and control group can be fully attributed to the training. The technique both identifies the causal relationship and produces a quantitative estimate.

Unquestionably, the strength of experiments lies in the high reliability of causal attribution with regard to the treatment under consideration. Secondly, with experiments, it is possible to achieve statistically significant results with relatively small samples. Moreover, the technique has an excellent persuasive power since its logic is easy to communicate and the same holds for the conclusions drawn from the findings.

However, there are good and bad reasons why experiments are difficult to implement in practice:

- An experimental design needs to be planned long enough in advance (in order to prepare the random assignment) and to be continued until the final survey (typically one year after the end of the training). The overall process is longer than that of other approaches, and therefore more difficult to adjust with the policy-making agenda.
- Field-level operators may be reluctant to implement a random assignment which, in addition, may raise difficult ethical debates\textsuperscript{43}.
- Random selection may affect (demotivate) those who are not selected and (motivate) those who are, and thus create an artificial effect.
- Finally, a random selection may be perceived as ethically questionable and politically unfeasible.

\textsuperscript{42} Parametric assumptions such as Gompertz or Weibull distribution, both of which can have declining, increasing, or constant hazards, depending upon the parameters of the distribution.

\textsuperscript{43} This technique is used extensively in testing new medicines. In such cases, random assignment does not raise too many ethical questions, even if the treatment tested may be harmful, because there is a very wide consensus that medical knowledge is a common good. This is far from being the case in the area of labour market policies, even if the treatment is always assumed to be a plus, because such policies are often subjected to heated political debates, and because they target vulnerable people. Moreover, public authorities may fear being blamed for preventing people from accessing training opportunities, but this problem may be addressed by delaying treatment to certain groups. This has been done, for instance, in the PROGRESA anti-poverty programme in Mexico.
Another way of addressing the difficulties of RCTs consists in relying on “natural experiments”. For instance, Leuven and Oosterbeek (2008) took into consideration the employees who wanted to participate in training but did not do so because of some random event. This natural experiment reduced the estimated wage returns to training considerably.

Black et al. (2003) selected their treatment group by assigning profiling scores that take on integer values from 1 to 20, with higher scores indicating claimants with longer expected periods of unemployment. The reemployment and training services were then allocated a recruitment ceiling for each profile. Where the number of claimants exceeded the capacity, a random selection was done, and the non-selected were assigned to a control group. This allowed a random selection without disrupting the programme and without denying services to those most in need. The main reason for using the described approach, however, was a capacity problem and not any plans to make a random selection for the evaluation. The resulting difference between the experimental group and the control group was a 2.2-week reduction in benefit receipt, a reduction in mean benefit payments of US$143, and increased earnings, but only in the short term.

All other approaches involving counterfactuals are qualified as quasi-experimental. They are reviewed in the following sections.

5.3.3 Matching techniques

Matching is based on the assumption that the employment status of non-participants has the same distribution as participants would have experienced had they not participated in the programme and vice versa (Heckman et al., 1997). Similar to randomisation in a classical experiment, matching balances the distributions of all relevant, pre-treatment characteristics in the treatment and comparison group. By relevant we mean all factors that influence the assignment to treatment as well as the potential impact.

Exact matching

The exact matching technique proceeds by finding for each participant a counterpart in the non-treated population and vice versa. There are limitations if: (1) some influential factors are not observable, and/or (2) the number of influential factors to be considered is too large.

Propensity Score Matching

The Propensity Score Matching technique (PSM) consists in computing a score which reflects the propensity to be enrolled in the training for all individuals belonging to the treatment and control groups. Then each member of the treatment group is matched with a member of the control group through one of the available algorithms, e.g. the Nearest Neighbour matching with replacement.

44 This technique is used in several country studies in the context of that contract. The analyses are performed using Stata version 11.0 with the module psmatch2 installed. See: E. Leuven and B. Sianesi. (2003). ‘PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing’. http://ideas.repec.org/c/boc/bocode/s432001.html. Version 3.1.5.

45 Other algorithms belong to the family of Kernel matching (Epanechnikov, Gaussian, biweight, Uniform and Tricube)
Deheija and Wahba (2002) contend that most of the matching algorithms yield similar results if there is substantial overlap in the distribution of the propensity score between the comparison and treatment groups. This means that there is a limitation in the use of this technique when the propensity scores of the two groups are too different from each other.

Another limitation is that of endogeneity, i.e. the fact that the same set of socio-demographic variables explains the propensity to participate and the probability of finding a job. Endogeneity is a serious threat to the validity of the impact estimate arising from a PSM analysis. Internal control groups, which enable the analyst to use a wide set of variables, offer a better chance of identifying distinct enough factors explaining participation and impact.

5.3.4 Instrumental variable

This technique proceeds through adjusting two equations to the dataset through a regression analysis. The first equation simulates participation in the training, which is itself introduced in the second equation simulating the employment status after training.

This technique requires that the analyst find variables which affect training participation, but at the same time have no direct effect on the employment status after training, in other words, variables that can be introduced in the first equation and not in the second one. This problem tends to hamper the application of the technique.

5.3.5 Regression discontinuity

In this approach, participation is assumed to depend on some observed variables according to a known, deterministic rule (Heckman et al., 1999). These variables affect the employment status of participants and non-participants, and also have an indirect effect through the participation. This indirect effect is the impact of training, which can be estimated if the method is successful (Frölich, 2002).

An example of this technique is the endogenous switching regression model which was used in the Italian evaluation quoted in Appendix 7. In this approach, participation and non-participation are the result of a decision process which is modelled through a latent decision variable and a threshold value.

5.3.6 Fixed effects

This technique relies on a regression analysis correlating the change in employment status to a series of explanatory factors including participation. However, the influence of all factors other than participation is supposed to be the same for participants and the control group, and captured by a “fixed-effects estimator”. The fixed-effects estimator takes into account any confounding influence of unobserved individual characteristics that correlate with both wages and training as long as these characteristics are fixed over time. Thus, the fixed effects estimator produces unbiased estimates, when the unobserved individual effects are permanent. However, it is likely that selection


47 Another difficulty is that the adjusted models are restrictive in the sense of being based on an assumption about the distribution of the unobservable variables (parameters).
into training also has dynamic aspects, which provide an additional potential source of bias.

In general, fixed-effects estimates of returns to training are smaller than that of standard regression models (see next section). This suggests that fixed-effects estimates at least partially eliminate selection bias.

### 5.4 Historical perspective

In the review and meta-study of ALMPs by Kluve (2006), the evaluation studies before 1994 are categorised as "first-generation" evaluation studies and the subsequent studies, until 1999, as "second-generation" evaluation studies.

The first-generation evaluation studies are described as evaluations of relatively new policies at the time, applying rather new econometric techniques, often on the basis of still rudimental data.

The second-generation evaluation studies until 1999, according to Kluve (2006), were characterised by both more mature and a more extensive set of policies, by a deepened and rapidly developing methodological know-how, and frequently by much improved data.

Kluve (2006) defines the third-generation evaluation studies as those that have been conducted since the late 1990s, and that are characterised by applying a set of relatively mature and (now) standard econometric methods. These studies, reviewed by Kluve, show some disparity of evaluation design and estimation techniques. Most of them are based on non-experimental data, with a few exceptions. They use either matching estimators or duration models as identification strategies.

Kluve found in this meta-study that experimental study design was significantly negatively associated with the likelihood of finding a positive effect.

In the meta-study by Card et al. (2009) a large number of European and American ALMPs including training programmes were reviewed, but no distinction between training and other ALMPs was made.

About 9% of the programme estimates in the sample of Card et al. (2009) were based on a randomised design. A comparison of the results of these non-experimental evaluations was made while controlling for the programme type and composition of the participant group. This comparison showed that there are only small and statistically insignificant differences in the distribution of positive, negative, and insignificant programme estimates between experimental and non-experimental evaluations. This was taken by the authors as an indication that the research designs used in recent non-experimental evaluations are not significantly biased compared to the "gold standard", i.e. the benchmark of an experimental design. This is in contradiction with the conclusions of Kluve (2006) and it is not possible to say whether it also holds for training programmes alone.
6 Learning from impact evaluations

This chapter builds on the analysis of four cases of impact evaluations carried out in the four Member States associated with the study. The purpose is to explore the extent to which policy-makers and/or decision-makers learned from these exercises.

At the end of each section the study team’s findings are presented in a box.

6.1 Analysing four impact evaluations

An evidence-based approach

Four case studies were carried out in respectively Belgium, Hungary, Italy, and Poland (see 2.2 and Appendix 1). All case studies have reached the stage of the second draft case monograph at the time of writing this document. Complementary investigations may still be carried out during the next weeks.

In each country, the study team has selected: (1) one of the conclusions of the evaluation report\(^{48}\), on the basis that it had good potential to be used by policy-makers / decision-makers, and (2) a circumstance in which that finding was used (or could have been used) by a specific decision-maker for one specific purpose. The finding and its potential use is called ‘opportunity for learning’. It represents the ‘case’ to be studied.

All selected items (countries, impact evaluations, findings, and learning opportunities) tend to be biased towards good learning. This bias is taken into account in the analysis.

The analysis proceeds by testing a series of assumptions about learning mechanisms (see next section). Each cause-and-effect assumption is typically analysed through two questions: (1) ‘has the expected change / event occurred or not?’ and (2) ‘what were the contributing factors?’.

Each case study relies upon a review of relevant documents and interviews with concerned public servants, evaluators, and experts. Case monographs follow the logic of causal assumptions and contain a number of statements which are categorised in several dimensions, two of which are of special importance: (1) ‘does the statement confirm / invalidate the assumptions under test?’ and (2) ‘does the statement constitute strong / weak evidence?’.

This section is the study team’s synthesis and interpretation of the whole set of evidence gathered through the case studies.

In methodological terms, this approach is qualified as a ‘contribution analysis’ based on a series of ‘embedded case studies’\(^{50}\). It is worth mentioning that such an approach is a state-of-art technique for evaluating impacts and drawing credible conclusions. This technique is relevant in a wide range of situations, including situations where counterfactuals do not make sense (see 5.2.2).

\(^{48}\) In Italy several impact evaluations could be considered, and the study team selected one of them on the basis that it had been used knowledgeably.


\(^{50}\) http://en.wikipedia.org/wiki/Embedded_case_study
**Assumptions: impact evaluations in an ideal world**

The chain of cause-and-effect assumptions under test was developed by the study team. It is not derived from any official EC document. It did not stem from a group discussion. These assumptions show a deliberately simplistic and optimistic view of the process of learning from impact evaluations. It is consistent with the accumulated knowledge on how evidence and expertise percolates in the policy-making sphere.

There are six successive assumptions as follows:

- In the four countries, impact evaluations were launched with an aim to address policymakers’ needs in terms of issues and time;
- Robust impact evaluation methods were designed in order to deliver credible knowledge to policy-makers;
- Impact evaluations produced new, timely, and credible knowledge;
- The newly acquired knowledge was disseminated in a manner which was adequate for reaching policy-makers;
- Experts and policy advisers accepted the new pieces of knowledge as credible and changed their understanding accordingly;
- Policy-makers and/or top level managers used the new pieces of knowledge supplied in the framework of debates and/or decisions.

6.2 Evaluation mandate (testing Assumption 1)

The first assumption under test is that impact evaluations are launched with an aim to address policymakers’ needs in terms of issues and time.

This assumption is tested in three steps by examining: (1) the relevance of impact evaluations to policy-makers’ needs, (2) timeliness, and (3) contributing factors.

**Relevance to needs**

Across the four countries under study, the main issues addressed by the impact evaluations were:

- Effectiveness of labour market interventions both for national and ESF funded programmes (Hungary);
- Effectiveness at programme level, project level and client level (Flanders);
- Effectiveness of training (Italy/Poland), especially adult training (Italy);
- Justifying the focus on the disadvantaged, a public that is difficult to attract and keep in training activities (Italy);
- Contributing to a better allocation of resources between different target groups (Poland).

All interviewees express the opinion that these issues were relevant to policymakers’ needs. However, it should be noted that all interviewees are on the supply side of the evaluation.

The impact evaluations were also meant to focus on some implementation issues of interest for managers’ decision-making needs, e.g.

- Different target groups’ trajectories in the labour market (Flanders);
- Factors determining success and failure in terms of effective and efficient implementation of the programme (Flanders/Italy/Poland);
- Success of innovative services and institutional arrangements such as decentralisation and cooperation between actors (Flanders/Hungary).
The study team considers that the main issues addressed in the impact evaluations are in line with the needs of policy-makers.

**Timeliness**

Impact evaluations were mainly supposed to feed into the processes of:

- Formulating national action plans (Poland)
- Designing co-financed policies and interventions for the next programming period (e.g. Hungary);
- Adjusting implementation if necessary (Flanders, Poland).

However, there is no evidence that the date of launching the evaluations was chosen in order to match the political agendas. On the contrary, the evaluations appeared as a long and continuous processes where the political agenda was not considered as crucial.

**Contributing factors**

Significantly, all impact evaluations were launched in a context where Managing Authorities had to comply with European obligations. In the case of Emilia Romagna, a national obligation was added to the European one, and this is said to be one of the main reasons for launching the impact evaluation.

In all four cases under study there is much more evidence that impact evaluations were pushed by information suppliers rather than pulled by policy-makers. This statement is confirmed in Poland by cross-checking information from interviews and the Evaluation plan. In the case of Flanders both policy-makers and public employment services supported this study. In the case of Emilia Romagna, interviewees assess that policy-makers paid limited attention to the design and findings of the evaluation. Impact evaluations were not mandated on an ad hoc basis. They were rather carried out to comply with an administrative requirement. They consequently addressed broad questions which were not specifically tailored to meet particular needs of policy-makers at a specific time.

Moreover, the substance of the impact evaluations seems to be strongly linked to the level of administrative and monitoring capacity. Flanders launched an ambitious evaluation because a sophisticated monitoring system allowed access accurate data. On the other hand, the ambition of the Hungarian evaluation was much more limited, due to reservations about evaluability.

**Concluding statement**

It was assumed that impact evaluations were launched with an aim to address policy-makers’ needs in terms of issues and time. The test of this assumption shows that:

| The main issues addressed are actually in line with the needs of policy-makers. | The evaluations were part of a long and continuous processes in which limited attention was paid to matching precise needs of policy-makers or decision-makers. |
| Managers and technicians have promoted impact evaluation in their countries, often (but not always) using the European obligation as a lever. They have cleverly identified evaluation questions which would be of keen interest to policy-makers and top managers. However, impact evaluations were pushed by information suppliers rather than pulled by policy-makers, something which is consistent with the fact that the issues addressed are not meant to match specific needs of specific policy-makers or decision-makers. |
6.3 Evaluation method (testing Assumption 2)

The assumption under test is that impact evaluations are designed with an aim to deliver credible and timely findings to policy-makers and decision-makers.

This assumption is tested in three steps by examining: (1) the evaluation methods, (2) the contributing factors, and (3) the limitations.

Main methodological choices

First, it should be remembered that all four evaluations under study aim to analyse the impact of training activities on unemployed participants, mainly in terms of finding jobs. Therefore, the methods applied are consistent with this purpose.

Former participants were usually retrieved in the lists provided by the suppliers of training services, and interviewed some time after the termination of their training activity (e.g. 6, 12 or 24 months). Participants were usually questioned by an independent research team using a CATI approach, about their current situation on the labour market, their job search, and their training. In the case of Poland they were also questioned about their labour market situation at the time of entering the training.

Flanders is a special case since participants were retrieved from a Public Employment Service database, something which made it possible to know the full story of their unemployment benefits and participation in various training activities. By using this data set it was possible to analyse the impact in terms of unemployment duration. Moreover, this approach enabled the evaluators to rely upon factual data about people's situation on the labour market, instead of statements.

All evaluations under study assess the impact of training activities by comparing a sample of participants and a control group.

The Italian case exemplifies the least demanding approach in terms of data collection. An external control group was drawn from a database of the National Institute of Statistics providing micro-data on the labour force. Individuals in the control group were selected if unemployed at time t (time when participants ended their training) and had not undergone any sort of training until t+12. The studies took into account the initial characteristics of participants and non-participants, and the selection bias was controlled through an appropriate method. The approach was the same in Hungary.

The Polish case involves an internal control group created in two steps. First, 81 local offices of the employment service were selected at random, and were asked to provide lists of people who were unemployed at the time when participants were entering the training activities. Then non-participants were selected through a layered randomised process and surveyed roughly at the same time and in the same way as participants.

In the Belgian case, participants and non-participants were drawn from the same database and no additional information was collected through interviews. Non-participants were matched with participants in terms of socio-demographic profile as in the other cases, and also in terms of the precise duration of unemployment before the training period.

The study team considers that the methodological approaches of all reviewed impact evaluations were valid enough and likely to address the needs of policy-makers.

The credibility of the approach is confirmed by several interviewees in Poland.
However, it should be kept in mind that the four impact evaluations were selected among the handful of such evaluations which are much more an exception than a rule across Member States.

**Contributing factors**

The following lines show that a number of external factors contributed to generating these examples of good methodological practice.

In the case of Poland, the good methodological practice is explained, at least in part, by the institutionalisation of evaluation. A National Evaluation Unit (hereafter 'the Unit') was established in April 2004. This Unit coordinates the evaluation work conducted at lower levels. It is also responsible for the development of evaluation standards, so as to improve the quality of the analytical studies conducted. Having a clear methodological mandate, the Unit was legitimate enough for promoting new approaches to impact evaluation. The Unit cooperates with the managing authorities of ESF-supported programmes in designing and conducting their evaluation programmes.

A first impact evaluation was carried out in 2005, although with a method which had limited validity. The Managing Authority subsequently launched the impact evaluation which is reviewed in the present study. This was done in 2006, two years after the institutionalisation of an evaluation function, something which shows quite a steep learning curve.

In the case of Hungary, an impact evaluation using a control group was carried out at the end of the nineties, but the institutional capabilities do not seem to have been maintained. This may explain that impact evaluation developed slower than in Poland.

In the Italian case the practice of impact evaluation was initiated by a few forerunning regions before being promoted at national level by ISFOL. In one of the pioneering regions, Emilia Romagna, there was a well-established tradition of data collection and statistical analysis. Interviewees consider that this was an enabling factor.

Another enabling factor is the administrative capacity, especially because this capacity is a prerequisite for running good monitoring systems, and then for retrieving former participants.

The case of Italy exemplifies this issue through a comparison between Objective 1 regions (South) and Objective 3 ones (North). At the beginning of the 2000-2006 programming cycle, Italy adopted a common information system for monitoring the implementation of the structural funds at the national level (called Monit 2000 and later Monitweb). By 2003, all the Northern regions had started their own placement surveys. This was the case in Emilia Romagna, which is one of the reasons why this region was a forerunner in terms of impact evaluation.

In contrast, for Objective 1 regions the Ministry of Labour decided to delegate monitoring and data collection to the ESF national evaluation unit, located at ISFOL. In practice, ISFOL encountered a few obstacles in retrieving personal data records on the project beneficiaries, mainly because of failure of project holders to transmit detailed data sets to regional authorities, and lack of adequate IT systems. The subsequent delays are one of the reasons why an impact evaluation could not be undertaken in the South.

Flanders is another example where a good monitoring system facilitated the undertaking of an impact evaluation. As in the case of Emilia Romagna, the evaluation of ESF-supported activities benefited from the good regional administrative capacity (here, a very comprehensive database of the Public Employment Service covering all types of training, including ESF-supported
ones). The case monograph confirms that the demand for knowledge was not the main factor.

The importance of public service databases is also demonstrated by the case of Southern Italy where, due to the bad state of the regional administrative archives on beneficiaries, ISFOL lacked sufficient information for developing the control group as initially planned. Moreover, ISFOL encountered difficulties when building the external control group, due to a change in the quarterly Labour Force Survey by the Italian statistics institute (ISTAT).

Considering the role of the above-mentioned factors, the study team considers that the sound methodological approaches observed are explained by: (1) the strength of administrative capacity, monitoring, and statistical databases, (2) the institutionalisation of evaluation, and (3) only a need for acquiring highly credible knowledge, i.e. the assumption under test.

**Limitations**

The following paragraphs investigate the limitations which applied to the impact evaluations under study. The issues addressed are value for money and feasibility.

In the case of Emilia Romagna, the annual placement surveys and impact evaluations were eventually considered as questionable in terms of value for money. After three years, it was decided to simplify the process. The regional monitoring system now relies upon an online database through which the subsidised training centres monitor the placement of their former participants. Information is to be checked for quality through periodic light independent surveys involving small samples, although these checks have not been carried out so far. Similarly, impact evaluations are now implemented on an ad hoc basis and no longer annually.

This light option is also preferred by Hungarian interviewees who say that impact evaluation should remain targeted and occasional, and that their findings should be used in the form of learning lessons in a wide context, and not in the form of a systematic and frequent feedback.

In the case of Flanders the evaluation approach was applied to the interventions targeted at the unemployed only because data were easily accessible in this area. Other ‘priority lines’ of the programmes were left aside by the evaluators because data were inexistent or inaccessible.

The case of Poland also suggests that impact evaluation cannot be systematic. In Poland nine measures were evaluated, but only two of them were subjected to an impact evaluation with a control group. Here again, the Managing Authority has opted for targeted impact evaluations and the cost of such exercises was certainly one of the reasons supporting that choice. This kind of study requires access to personal data such as registers of unemployed persons, which in the Polish case was quite problematic due to legislation on the protection of personal data.

Another limitation to impact evaluations is certainly that of feasibility since some ESF-supported activities do not lend themselves to building control groups. For instance, the Polish measure 1.1. (Development and modernisation of labour market instruments and institutions) produces its impacts through a small number of large open systems, and such systems cannot be analysed through control groups.

Finally, it must be stressed that all evaluations under study assess the impact of training activities on unemployed participants in terms of labour market status. Consequently, this study may not identify the specific limitations that
might apply to impact evaluation of activities targeting other groups, e.g. at-risk employees.

Considering the above-mentioned limitations, the study team considers that the observed good methodological practices could not be applied to all ESF-funded activities in a comprehensive and/or frequent manner.

**Concluding statement**

It was assumed that robust impact evaluation methods were designed in order to deliver credible knowledge to policy-makers. The test of this assumption shows that:

<table>
<thead>
<tr>
<th>All reviewed impact evaluations applied valid enough methods making it possible to deliver credible knowledge to policy-makers.</th>
</tr>
</thead>
<tbody>
<tr>
<td>However, the factors contributing to such good methodological practices are, per order of assessed contribution: (1) the strength of administrative capacity, monitoring, and statistical databases; (2) the institutionalisation of evaluation; and (3 only) a need for acquiring highly credible knowledge, i.e. the assumption under test.</td>
</tr>
<tr>
<td>Moreover, cost and feasibility constraints, including privacy issues, made it impossible to apply these good methodological practices to all ESF funded activities in a comprehensive and/or frequent manner.</td>
</tr>
</tbody>
</table>

**6.4 New knowledge (testing Assumption 3)**

The assumption under test is that impact evaluations produce new, timely, and credible knowledge. The term knowledge is understood as encompassing findings, conclusions and lessons arising from the evaluations.

This assumption is tested in three steps by examining: (1) the novelty of the acquired knowledge, (2) its timeliness, and (3) its credibility.

**Findings, conclusions and lessons**

Across the four countries, the impact evaluations found that:

- Training is effective in terms of the probability of finding a job, i.e. about +10% in Hungary, Poland, and Emilia Romagna;
- It has a better impact on the long-term unemployed than on the short-term unemployed (Poland/Flanders/Italy);
- There are different levels of impact depending on gender (stronger impact of vocational training policies for women than men in Italy and poor impact on long-term unemployed women in Poland - impacts almost three times lower for women than for men);
- Positive impacts are on the most vulnerable groups such as the long-term unemployed (Italy);
- Impacts decrease with age and increase with level of schooling (Italy).
- Counselling and custom-made guidance are effective in specific conditions (Flanders);
- The quality of partnership among actors of the labour market is a success factor (Flanders).

Most of the above findings, conclusions and lessons were new, if not always in substance, at least in their quantitative nature. In some instances, these were the first valid impact estimates ever produced in the country and in the policy area (e.g. Poland). In Italy, the substance of the findings was all but a surprise, but all the interviewees stressed that the evaluation under study provided the first pieces of credible evidence confirming the positive impact of training.
Sometimes the newly acquired knowledge is said to have addressed issues of specific interest for policy-makers, such as the success of counselling in Flanders, the good results of disadvantaged trainees in Italy, or the relative failure of the Polish programme with long-term unemployed women in comparison with men.

**Timeliness**

As seen earlier, the delivery of the evaluation findings was mainly driven by institutional and technical reasons. The following sections will show that the newly acquired knowledge was not always in line with the political agenda.

**Credibility**

Those involved in the impact evaluation processes generally had a positive self-assessment of the credibility of the findings.

However, the lessons learned in the Hungarian case were not considered as fully robust and credible, as doubts emerged about data reliability.

In Poland, the evaluation method combined with the reputation of the evaluator ensured the credibility of the results. However, a problem arose from the fact that the impact analysis applied to the situation of former participants six months after the completion of their training. The Managing Authority reported that, because of such a short-term perspective, it was not possible to convert all findings into credible recommendations.

In Italy no questioning of the credibility of the findings was mentioned, probably due to the strong reputation of ISFOL (the evaluation team) and the fact that the evaluation method was designed in close coordination with the Managing Authorities.

In the case of Flanders several scientific studies were later conducted with the same data sources, and confirmed the results of the evaluation.

**Concluding statement**

It was assumed that impact evaluations produced new, timely, and credible knowledge. The test of this assumption shows that:

| All impact evaluations delivered important and relevant findings, conclusions and lessons which were new, if not always in substance then at least in their quantitative nature. |
| The evaluation findings were generally perceived as credible. No criticism was mentioned about the validity of the analyses, even in the single case (Belgium) where the analysis was replicated. |
| However, in some instances, credibility was questioned on the grounds that data were unreliable or inappropriate. This shows that credibility applies to the evaluation as a whole, and not only to the impact analysis method. |

### 6.5 Dissemination (testing Assumption 4)

The assumption under test is that the acquired knowledge was expressed in a clear-cut manner and disseminated through media and in a form that are adequate for reaching policy-makers.

This assumption is tested in two steps by examining: (1) the quality of dissemination and (2) the contributing factors.
Quality of dissemination

A first dissemination channel consists of meetings and seminars that are part of the evaluation process. In the case of Poland, quarterly reports from the impact analysis were discussed in an Evaluation Steering Group. An equally important role is played by the monitoring committees of the ESF-supported programmes which discuss the recommendations arising from the evaluations. They take a formal position on how to implement them within the concerned institutions.

In the case of Hungary, the main dissemination event was the closing session of the evaluation project. The consultants presented their overall findings as well as the 7 analysed measures. The following discussion covered methodological issues (nature and complexity of the evaluated measures) rather than substantial ones. Considering the attendance, it is doubtful that substantial messages could be conveyed to decision-makers. For instance, one of the key institutions delegated a brand new colleague, for whom the subject of the evaluation was as new as the evaluation itself.

A second dissemination channel consists of presentations in the ‘monitoring committee’ of the programme. In the case of Emilia Romagna, it is said that limited attention was paid to such presentations. On the other hand, there is sound confirmation of the assumption in the case of the evaluation finding selected for deeper investigation in Poland. The finding is that long-term unemployed women succeed three times less than do men in finding a job after their training. This finding was discussed in a meeting of the Monitoring Committee in June 2008.

In addition, the evaluation reports tend to be published on the Internet (e.g. Hungary). In the case of Italy, it was proposed that not only the report but also the data sets be made publicly available, with an aim to encourage further academic research complementing the evaluation. However, this idea has not yet been implemented.

Overall, the case studies show that the quality of dissemination ranges from disappointing to very good.

Contributing factors

The high/poor quality of dissemination is certainly explained by the existence of a well/poorly established evaluation function. This is shown by the contrasting examples of Poland and Hungary. In the case of Poland, the evaluation findings were presented to and discussed by decision-makers in a Monitoring Committee meeting. The following sections will show that this had an influence on subsequent decisions. One of the reasons for this success is that the National Evaluation Unit was sufficiently well-established for managing the dissemination process. In Hungary, the dissemination process was both unsuccessful and weakened by an institutional change. The responsibility to run evaluation projects, which was centralised at the level of the National Development Agency, was eventually transferred to the Department for Evaluation and Strategy. A part of the institutional memory and capacity was thus temporarily lost.

The second main contributing factor is the above-mentioned balance of political demand versus technical supply. Of course, this factor may enhance (e.g. Poland) or weaken (e.g. Hungary) the interests of those who are the assumed recipients of dissemination efforts, but another and rather perverse phenomenon is also at play, i.e. heavy/light weight given to substantial issues in comparison to technical ones. This problem was already quoted in the case of the final event which ended the Hungarian evaluation. It is also quite visible in the case of Italy where interviewees quoted a number of dissemination efforts...
(meetings, presentations in conferences, papers), all of which had a strong methodological dimension.

**Concluding statement**

It was assumed that the newly acquired knowledge was disseminated in a manner which was adequate for reaching policy-makers. The test of this assumption shows that:

- The quality of dissemination efforts ranges from disappointing to very good.
- Just releasing credible findings to policy-makers contributes very little to the uptake of such findings. On the other hand, other factors make a greater contribution, i.e. (1) institutionalisation of evaluation, (2) political demand for evaluation, and (3) priority given to communicating on substantial issues rather than technical ones.

### 6.6 Accumulation of knowledge (testing Assumption 5)

The assumption under test is that experts and policy advisers accept the new pieces of knowledge as credible and change their understanding accordingly.

At this stage of the analysis, the available material can be examined in just one dimension, but quite an important one, i.e. the accumulation of knowledge through streams of successive studies. This part of the analysis builds upon two converging stories which reinforce each other.

In the case of Emilia Romagna, one of the main lessons drawn from the impact evaluation was that women did better than men in finding jobs after their training. In fact, the lesson was reinforced by the subsequent EQUAL evaluations, which concluded that the impact of training for women is higher where a network connects the actors in the field and guides participants from the educational activities to the labour market. The EQUAL evaluations were published around 2006, while the impact evaluations under review were issued in 2004 and 2005. In the case of Poland, a major and surprising finding was that long-term unemployed women succeeded three times less than men in finding jobs after their training. Interviews conducted with the Managing Authority representatives revealed that the findings of the impact evaluation led to the decision to conduct two subsequent qualitative analyses ("The Study on a perspective of equal gender opportunities within SOP 2004-2006" and "OP HC pilot projects study"). Both have confirmed the lesson learnt and provided an explanation for the findings.

In both countries it can be seen that knowledge was accumulated through successive evaluations and studies before being ready for use in the policy-making process. This actually occurred, as seen in the next section.

**Concluding statement**

It was assumed that experts and policy advisers accept the new pieces of knowledge as credible and change their understanding accordingly. The test of this assumption is not comprehensive, but it shows that:

- The newly acquired knowledge was often used indirectly, by feeding into a stream of subsequent evaluations and studies before being ready for use in the policy-making process.
6.7 Knowledge use (testing Assumption 6)

The assumption under test is that policy-makers and/or top level managers used the new pieces of knowledge supplied by the impact evaluations in the framework of debates and/or decisions.

This assumption is tested in three steps by examining: (1) the affected decisions, (2) the contributing factors.

Affected decisions

In two instances (Poland and Emilia Romagna), there is factual evidence of policy-making decisions taken after the supply of new knowledge and in line with it.

In Emilia Romagna, the new ESF programme for 2007-13 has maintained a strong focus on social inclusion and gender equality, despite a significant decrease in financial resources. This is in line with the knowledge accumulated over the previous three years, which indicates that women and vulnerable groups greatly benefit from training, provided that some assistance is combined with training.

In Poland, a range of actions were taken in order to better comply with the principle of equal opportunities, including inter alia working groups involving intermediary bodies, setting standards on equal opportunities, raising awareness among implementing agents and potential project applicants, recommendations to evaluators, etc. These actions are fully in line with the knowledge accumulated over the previous two years, which shows that long-term unemployed women do not benefit enough from training.

In both cases, the interviewed public servants assess that the impact evaluation contributed to the decision-making, although in an indirect way through a stream of studies rather than by their own merits. However, in the Italian case the EQUAL evaluations had a much stronger influence on decision-makers because they were delivered at the beginning of the design phase of the following programme.

In Flanders the interviewed civil servants assessed that the findings of the impact evaluations were used by managers to steer the programme in a short-term approach, much more than by policy-makers in the framework of political debates and/or policy-making.

In Hungary, no evidence could be found for supporting the assumption that the Managing Authority and/or the Ministry of Labour and Social Affairs used the newly acquired knowledge.

Overall, the case studies show a relatively good alignment between the newly acquired knowledge and the subsequent decisions taken at either political or management level.

Contributing factors

In the most obvious cases of connection between evaluation and politics, i.e. Poland and Emilia Romagna (see above), the issues at stake were gender and equal opportunities. By then, these EU goals were promoted highly actively, which is probably the main reason for the political decisions taken in both countries.

In the same two cases of Emilia Romagna and Poland, there is evidence of a lasting close relationship between policy-makers and a stable group of technical and administrative people in charge of monitoring and evaluation. This may have facilitated the take-up of evaluation findings, even if the interviewees are doubtful (Italy) or silent (Poland) about the political ownership of the
evaluation process. The same factor may have been an obstacle in the case of Hungary.

Finally, two case studies suggest that the capacity of policy-makers to absorb evaluation findings is limited. As noted above, the authorities of Emilia Romagna decided to stop the annual placement surveys and the annual impact evaluations because their cost was high and the supply of new knowledge became smaller and smaller over years. In Flanders the impact evaluation was designed primarily to serve the needs of managers. An issue of major political importance could have been the so-called ‘creaming effect’, i.e. the fact that training institutions recruit the most employable candidates in order to show good performance indicators. However, it seems that there was no political demand for learning in this area, which is not covered by the impact evaluation.

**Concluding statement**

It was assumed that policy-makers and/or top-level managers used the new pieces of knowledge supplied by credible impact evaluations in the framework of debates and/or decisions. The test of this assumption shows that:

<table>
<thead>
<tr>
<th>There is a relatively good alignment between the newly acquired knowledge and the subsequent decisions taken at either political or management level.</th>
</tr>
</thead>
<tbody>
<tr>
<td>However, the main contributing factor (1) for this apparent success is the fact that such decisions were in line with one of the top priorities of the EU. At best, the availability of new credible knowledge comes second (2) in the contributing factors. In addition, other factors play a significant role, such as (3) a lasting close relationship between policy-makers and a stable group of technical and administrative people, and (4) the limited capacity of policy-makers to absorb evaluation findings.</td>
</tr>
</tbody>
</table>
7 Recommendations

This section starts with a reminder of the findings of the previous chapter. It then introduces a series of other relevant findings arising from the study. On this basis, the study team suggests a series of recommendations.

In their draft version, these recommendations were discussed extensively in a focus group meeting between participants from several Managing Authorities (see Appendix 10). Another presentation was also made in a workshop of the Conference ‘Shaping the future of the ESF’. This chapter takes stock of the comments received during these two events.

7.1 Findings from the four case studies

This section is a slightly edited compilation of the step-by-step conclusions of the previous chapter (boxes).

The reviewed impact evaluations addressed a series of issues that were in line with the needs of policy-makers, despite the fact that limited attention was paid to matching the political agendas. This is understood to be the result of clever efforts by public managers to promote impact evaluation, using the European obligation as a levy, and to identify evaluation questions which would be of particular interest to policy-makers and top managers. Overall, impact evaluations were pushed by information suppliers much more than pulled by policy-makers.

All reviewed impact evaluations applied sufficiently valid methods enabling the team to deliver credible knowledge. However, the main factors contributing to such good methodological practices are: (1) administrative capacity, good monitoring, and statistical databases, and (2) institutionalisation of evaluation. The need for acquiring highly credible knowledge (assumption under test) is only the third contributing factor per order of importance.

All impact evaluations delivered important and relevant findings, conclusions and lessons which were new, if not always in substance, then at least in their (quantitative) form. The validity of impact analyses were generally assessed as satisfactory. However, credibility was sometimes questioned on the grounds that data were unreliable or inappropriate, showing that credibility applies to the evaluation as a whole and not only to the impact analysis method.

The quality of dissemination efforts ranges from disappointing to very good. The study shows that releasing credible findings to policy-makers contributes very little to the take-up of such findings. On the contrary, other factors make a greater contribution, i.e. (1) institutionalisation of evaluation, (2) political demand for evaluation, and (3) priority given to communicating on substantial issues rather than technical ones.

The newly acquired knowledge was often used indirectly, by feeding into a stream of subsequent evaluations and studies before being ready for use in the policy-making process.

There is a relatively good alignment between the newly acquired knowledge and the subsequent decisions taken at either political or management level. However, the main contributing factor for this apparent success is the fact that such decisions were in line with one of the top priorities of the EU. At best, the availability of new credible knowledge (assumption under test) ranks second among the contributing factors. In addition, other factors play a significant role such as a lasting relationship between policy-makers and a stable
group of technical and administrative people, and the limited capacity of policy-makers to absorb evaluation findings.

From the above conclusion, the study team understands that the main problems restricting the use of impact evaluations are:

- Limited political ownership;
- Uneven dissemination efforts with an excessive focus on methodology;
- Impact estimates not associated with sufficient explanatory studies;
- Cost and feasibility constraints making it impossible to apply impact evaluations in a comprehensive and/or frequent manner;
- Limited interest of individual impact evaluations in comparison to clustered cross-country exercises.

7.2 Other problems identified in the study

Through a synthesis of the research findings available in the literature, the study team has identified a number of "knowledge gaps", i.e. areas of knowledge that are highly relevant for the ESF, but where no or almost no lessons have been learned (see Box 2 and Box 4). The most serious problems arise from the lack of knowledge in the following areas:

- Skills obsolescence and the effects of training on aged employed people;
- Concrete calculations of costs and benefits of return-to-work training, including possible positive externalities associated with leaving unemployment;
- Effects of various combinations of return-to-work training, coaching and counselling services, and the overall incentive framework for finding jobs;
- Effects of the various combinations of training and assistance targeted at young job-seekers;
- Effects of training and other associated services on people at risk of losing their job;
- Substitution effects and other macro-level effects.

Through various sources, including a test of several impact analysis techniques in four countries, the study team has identified a series of recurrent methodological difficulties or risks:

- Data may be biased towards success if they have been collected by training service providers;
- Participants and control groups may be difficult to match if the dataset is not rich enough in terms of observed variables, something which is frequent in the case of an external control group derived from a statistical database;
- Privacy problems are often raised as a constraint to accessing data, even where privacy is not a real issue;
- Impact estimates may not be credible enough if the employment status of participants is measured after six months.

7.3 Suggestions for new rules, incentives and guidance

The following recommendations are made in the perspective of designing the evaluation requirements and guidelines for the next programming cycle of the ESF. They derive from the problems identified above.
Introduction
The views expressed in this text derive from the author’s expertise and from a recent study launched by DG Employment and Social Affairs in October 2009.

The study was carried out by a consortium of three firms: Euréval, Ecorys and Ramboll-Management. It applied to human capital investments understood as public expenditures in: (1) education and training of employees, job-seekers and/or inactive people, (2) internships and temporary work placement associated with education or training, and (3) strengthening the capacity of education and training systems and structures. Overall, such investments account for over 50% of the ESF resources over the current programming period.

The study included three clusters of tasks as follow:

- A synthesis of available research findings related to the return on human capital investments, with a focus on: (1) impact of training programmes targeting employed people, and (2) impact of return-to-work programmes targeting job seekers. Based on a first extensive screening of peer-reviewed publications, the research synthesis was improved with the assistance of a panel of seven experts from EU and USA.
- A series of impact analyses based on individual data provided by four volunteering Authorities in respectively Belgium (Flanders), Italy (northern regions), Hungary, and Poland. Through these analyses, the study team tested a series of impact analysis techniques, all including control groups. Due to the nature of the available data, this part of the study applied to unemployed people only.
- An in-depth investigation into four success stories of impact analyses that had been carried out in the same four countries between 2005 and 2009. This investigation focused on the process of learning from impact evaluations. The study team discussed the recommendations stemming from this investigation in a focus group meeting attended by several managing authorities.

Accumulated knowledge

Training of employed people

Recent studies have shown a significantly positive impact of training on productivity at industry level, given that the impact for individual trainees or companies is not clearly demonstrated. Impact at industry level is sufficiently high to argue that the returns on human capital investments are profitable and competitive with other investments.

The problem is that only part of these returns go to the company that pays for the training. Most of the benefits go to other companies in the industry and to the employees who received the training, especially when they move to other companies. This raises the question of underinvestment by the employers who do not have sufficient incentives to invest in the training of their staff. Since the larger part of the returns to human capital investments actually goes to the industry rather than to the individual firms, there is a case for investing public money in education and training systems and structures at this level.

Off-the-job training is known to be more effective than on-the-job training, provided that trainees have the opportunity to apply their newly acquired skills when returning to the work place. This calls for a close connection between training and human resource management.
**Training of unemployed people**

Return-to-work training has a positive employment effect only in the medium and long term, after an initial period needed for balancing negative lock-in effects, i.e. the fact that people stop seeking jobs when they participate in a training programme.

Moreover, training has no significant impact unless it is associated with job-search assistance and or a policy framework that creates incentives for seeking jobs. This calls for a closer connection between ESF investment in return-to-work training and the reform of national labour-market policies.

**Knowledge gaps**

There is a specific knowledge gap as regards at-risk industries and firms, where human capital investments might be particularly valuable, but where the shortage of resources may generate under-investments.

Skills obsolescence and the impact of training on aged staff members is another field where more knowledge should be accumulated. This is particularly the case in the context of the development of greying economies in Europe.

Very little research exists on the economic returns on training unemployed people, because: (1) costs and benefits tends to be studied separately, and (2) limited efforts have been made at valuing the many benefits of social inclusion.

There is insufficient evidence on the important issue of “quality of training”, in terms both of how to assess quality and of whether the benefits of higher quality are worth the cost.

Similarly, little is known about the patterns of employment effects in relation to specific target groups, except that young job seekers are systematically assessed as less successful with labour market training.

No clear conclusion can be drawn concerning the possible pattern of training effectiveness over the business cycle. There are some indications that the training effect is pro-cyclical in so far as the impact of return-to-work training is higher in periods with good job opportunities and low unemployment, but the picture is not clear.

Finally, the whole issue of secondary impacts remains a black box. The main concern is about substitution effects, i.e. the fact that an unemployed trainee finds a job at the expense of someone else. Substitution is assumed to be limited where a training programme is designed to match some skills gap. Virtuous circles are also assumed to take place through improved productivity and innovation, reinforced competitiveness, large market shares, and secondary impact on employment.

**Lessons from the methodological tests**

The study confirms that there is a large discrepancy between actual impact and what is measured with impact indicators such as “proportion of trainees having found a job after one year”. Where control groups are used for analysing causes and effects, impact estimates may be divided by a factor of 2 or even more. Moreover, additional econometric analyses tend to further reduce impact estimates significantly. Changes from impact indicators to impact estimates do not follow the same pattern for all publics. This raises the question of the relevance of monitoring impact indicators in the absence of a sound impact analysis.
The study has shown considerable cross-country differences in terms of impact estimates. Such differences call for a cautious interpretation, and could create opportunities for learning, for instance if they can be attributed to variations in the quality of training or to differences in the economic environment. However, there is at least one instance where a highly positive impact had to be attributed to unreliable data, i.e. data collected by training providers and suspected to be positively biased. The lesson is that individual data used in impact analyses should preferably originate from specific surveys or external databases rather than from internal management systems.

Depending on the analysis technique chosen, the evaluation should be launched up to one year before training. The study has shown that impact should be analysed over a period of one to two years after the training. An additional year should be added for the analysis itself and for learning the lessons. Overall, the duration of the learning cycle may be in a range of 2.5 to 4 years. This shows that impact analyses cannot be included in annual management cycles.

Many findings of the study called for interpretation and explanations, showing that impact estimates seldom speak for themselves. Most often, they need to be included in some evaluation process where a common understanding could be developed in an interactive manner.

Finally, the findings of the study were difficult to interpret in cases where a wide variety of training approaches were analysed altogether. On the other hand, more lessons could be learned from narrower studies focusing on specific training approaches, specific target groups, and/or specific regions, especially if the study had a comparative perspective.

Learning about the learning processes

The study team investigated four success stories of impact analyses and came up with mixed findings. On the positive side, it was found that the impact analyses delivered new, credible, and relevant lessons. On the negative side, it must be said that actual learning from these lessons was quite limited. This is explained in more details in the following paragraph.

The reviewed impact evaluations addressed a series of issues that were in line with the needs of policy-makers, despite the fact that limited attention was paid to matching the political agendas. This is understood as the result of clever efforts by public managers to promote impact evaluation, using the European obligation as a lever, and to identify evaluation questions which would be of particular interest to the policy-makers and top managers. Overall, impact evaluations were pushed by information suppliers rather than pulled by policy-makers.

With limited exceptions, the reviewed impact evaluations applied sufficiently valid methods enabling the team to deliver credible knowledge. However, the main factors contributing to such good methodological practices are: (1) administrative capacity, good monitoring, and statistical databases, and (2) institutionalisation of evaluation. The need for acquiring highly credible knowledge (assumption under test) is only the third contributing factor per order of importance.

All impact evaluations delivered important and relevant findings, conclusions and lessons which were new, if not always in substance then at least in their (quantitative) form. The validity of impact analyses were generally assessed as satisfactory. However, credibility was sometimes questioned on the grounds that data were unreliable or inappropriate, showing that credibility applies to the evaluation as a whole and not only to the impact analysis method.
The quality of dissemination efforts ranges from disappointing to very good. The study shows that releasing credible findings to policy-makers contributes very little to the take-up of such findings. On the contrary, other factors make a greater contribution, i.e. (1) institutionalisation of evaluation, (2) political demand for evaluation, and (3) priority given to communicating on substantial issues rather than technical ones.

The newly acquired knowledge was often conveyed to policy-makers in an indirect way, by feeding into a stream of subsequent evaluations and studies before being ready for use in the policy-making process.

There is a relatively good alignment between the newly acquired knowledge and the subsequent decisions taken at either political or management level. However, the main contributing factor for this apparent success is the fact that such decisions were in line with some top priority of the EU. At best, the availability of new credible knowledge (assumption under test) ranks second among the contributing factors. In addition, other factors play a significant role such as a lasting relationship between policy-makers and a stable group of technical and administrative people, and the limited capacity of policy-makers to absorb evaluation findings.

From the above conclusion, the study team understands that the main problems restricting the use of impact evaluations are:

- Limited political ownership;
- Uneven dissemination efforts with an excessive focus on methodology;
- Impact estimates not associated with sufficient explanatory studies.

In addition to the above conclusion, this study shows that, even in exceptionally favourable circumstances, cost and feasibility constraints make it impossible to undertake impact analyses in a comprehensive and/or systematic manner.

**Proposals for addressing the identified problems**

Here are a series of proposals made to solve the identified problems. This set of proposals builds upon the idea that impact analyses should no longer be restricted to ex-post evaluations of individual programmes, but should rather be part of an EU-wide permanent process involving communities of researchers, evaluators, and practitioners in a context of strong political demand.

**Engage policy-makers in a culture of learning**

The expected consequence of the following recommendations is that policy-makers become interested in evidence from previous successes or failures, and engage in a culture of learning.

For that purpose, the set of regulations applying to the next programming cycle might include the following rules:

- Every year, Managing Authorities would have to approve a rolling multi-annual evaluation plan. Over a five year period, each plan should make at least one contribution to bridging one of the identified knowledge gaps through a sound impact analysis.
- Every year at the same time, Managing Authorities should approve an annual follow-up report identifying the main lessons learned on the programme, including an assessment of the reliability of these lessons, and the actions taken.
• A synthesis of relevant knowledge (evidence base) should be included in: (1) all evaluation reports, and (2) all decisions having a significant financial impact.
• By implementing the above three rules, Managing Authorities should demonstrate actual progress towards learning from impact analyses and towards evidence-based policy making. A part of the disbursements should be contingent on the ability to demonstrate such progress.

The above set of recommendations should be considered as a whole and not as a menu. During the focus group meeting, the last and most difficult suggestion (making disbursements contingent on learning) was assessed as the only way of ensuring that the other suggestions were not implemented in the form of bureaucratic compliance.

Of course, the above suggestions involve costs and burdens, and they are acceptable if and only if other costs and burdens are alleviated, for instance in terms of reporting or controls.

Managing knowledge through incentives
The expected consequence of the following recommendations is that knowledge communities connect users and information providers across programmes and across programme cycles, with a strong focus on knowledge gaps.

For that purpose, attractive financial incentives should be used in a strategic manner at EU level. Specific funds should be earmarked for evaluation and allocated for, and only for:

• Evaluations involving a sound impact analysis aimed at bridging one of the identified knowledge gaps;
• Clustered evaluations addressing the same cause-and-effect question in several regions or several countries;
• Syntheses, quality assessment, and discussion of research and evaluation findings within knowledge communities (e.g. networks, conferences) involving practitioners.

Providing technical guidance
The Commission or any relevant body with the Commission’s support should create publicly available guidance on:

• Identified knowledge gaps and sound impact analysis techniques that are relevant for evaluations aimed at bridging such gaps;
• Assessing the quality of impact analyses;
• Effective approaches to translating impact analyses into usable knowledge.

Mutual learning on evidence-based policy-making
During the events bringing together those in charge of evaluation and programming in the Managing Authorities, sufficient time should be devoted to learning from one another about evidence-based policy-making.
Appendix 1 – Study method

Return on human capital investments

Which types of human capital investments present the highest return for the individual participant or firm involved and for society at large?

Purpose

To address the first main expectation of this study, which is to feed the next round of policy-making with evidence-based information about the benefits of human capital investments in general, and of specific types of investments in particular.

Nature

A question which involves a value judgement, ideally encompassing four criteria, i.e. (1) cost and (2, 3, 4) value of impacts on respectively: concerned firms, concerned individuals, and the rest of the society. The expected assessment is both relative (comparing returns across various types of investment), and absolute (to what extent is a given type of investment worth at all?).

Method

This question was answered on the basis of the available literature at first, plus in-depth investigations in a few Member States. Several methodological points deserve to be highlighted:

- **Defining human capital in general.** The term is defined in 1.1.2. The expert panel pointed out that this concept should be understood in a broad way as the accumulation of an individual’s experiences, including both education, training, and a wide range of other factors (see page 127, footnote 53). Most research works, including this study, reflect on marginal changes in human capital, often measured through a proxy like training duration. No or almost no information is available about the stock of human capital and about the important issue of depreciation over time.

- **Types of human capital investment.** This study focuses on the types of investments that are relevant in the context of the ESF, as explained in 1.1.2. In Chapter 3, such investments are categorised in two main clusters: training of employees (general or firm specific, formal or informal), and training of the unemployed (associated or not with job search assistance or incentives for taking job opportunities). Chapters 4 and 6 concentrate on the second category only, since ESF evaluations have most often focused on the impact on the unemployed. The concept of ‘type of investment’ is also understood to cover the targeted groups such as women, younger or older people, and so on. As shown in 3.3, an important issue is that of ‘at-risk’ workers. This issue is quite important in the context of the ESF, and it is almost not addressed in this study.
• **Return on investment.** This concept is introduced and explained in 5.1. However, the study shows that there are severe limitations in the available knowledge in this area (see Box 2 and Box 4). For this reason, the study team answered the question through a series of scenarios (see 4.5.2). The study team’s assessment is that this part of the study is subjected to methodological limitations (see 2.3) and should be subjected to further rounds of discussions, comments and improvements.

**Impact analysis methods**

*What methods are available to estimate the impacts of human capital investments and to what extent are they applicable to ESF supported interventions?*

**Purpose**

To recommend evaluation methods enabling to learn valid and transferable lessons, and applicable in all Member States.

**Nature**

A technical question which does not involve any value judgement related to ESF.

**Method**

This question was answered on the basis of the literature review, plus a test in four Member States. Several methodological points deserve to be highlighted:

**Available methods:** A number of impact estimation techniques are briefly introduced by the study team in 5.3. In addition, Section 5.4 discusses the issue of available methods in a historical perspective on the basis of the literature review.

**Applicable methods:** The expert panel held a controversial discussion on whether or not a hierarchy of methods could be established (see Box on page 129). The discussion concluded that technical choices are always relative and contingent to a given situation (including evaluation questions and data collection constraints) and that nobody knows the extent to which a technique over- or under-estimates impacts in comparison with another one. Notwithstanding these prudent views, the study team took the liberty to suggest a procedure intended to assist in the choice of an impact analysis approach in the context of a given ESF evaluation (see 5.2.2).

**Learning from evaluations**

*How could impact evaluation and dissemination of learning be incorporated in the design of the ESF supported interventions from their outset?*

**Purpose**

To suggest evaluation and monitoring arrangements enabling policy-makers to learn more and faster from experience.
Nature

An organisational question which did involve some value judgement about the 'good use' of evaluative information, i.e. use of impact related knowledge for policy-making purposes (see 6.1 and 6.7).

Method

This question was answered on the basis of four case studies of past ESF evaluations having included impact estimates. The findings were discussed and improved through an interaction with managing authorities in a few Member States. Several methodological points deserve to be highlighted:

- **Impact evaluation.** The term impact is defined in Appendix 11. An 'impact evaluation' is an evaluation which is intended to produce a quantitative impact estimate and/or to learn lessons from evidence about impact, either through quantitative or qualitative approaches. Impact analysis is often said to be the main value that evaluation adds to audit, monitoring, or quality management. In this respect 'impact evaluation' and 'evaluation' should be considered as synonymous expressions. However, so many evaluations fell short of analysing impacts that it is worth speaking of an 'impact evaluation' as an evaluation that answers at least an impact related question in a professional manner.

- **Learning.** In this study, the term 'learning' is understood as 'learning lessons' from impact analysis in a given context for the sake of future policy-making in the same context or elsewhere. Examples of lessons learned are visible in e.g. Box 1, Box 3, Section 4.3.4, and Section 4.4.6. As regards learning, an issue of particular importance is that of the transferability of lessons. This issue was addressed by the expert panel (see Box at Page 132) with the conclusion that some lessons are transferable, although with careful caveats, whilst impact estimates cannot be generalised.

- **Design of ESF interventions.** This term is understood in a broad way as encompassing the design of ESF supported programmes, but also that of rolling multi-annual evaluation programmes, and activities related to learning across programmes.
Appendix 2 – References of the literature review

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Title</th>
<th>Cited by(^{51})</th>
<th>Journal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barrett and O’Connell (2001)</td>
<td>Does training generally work? The returns to in-company training</td>
<td>171</td>
<td>Industrial and Labor Relations Review</td>
</tr>
</tbody>
</table>

\(^{51}\) Number of authors having cited the articles. Source: Google Scholar.
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Title</th>
<th>Cited by</th>
<th>Journal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caponi et al. (2009)</td>
<td>The Impact of Aggregate and Sectoral Fluctuations on Training Decisions</td>
<td>-</td>
<td>APER SERIES Forschungsinstitut zur Zukunft der Arbeit Institute for the Study of Labor</td>
</tr>
<tr>
<td>Cockx (2003)</td>
<td>Vocational Training of Unemployed Workers in Belgium</td>
<td>12</td>
<td>IZA Discussion Paper</td>
</tr>
<tr>
<td>Conti (2005)</td>
<td>Training, productivity and wages in Italy</td>
<td>34</td>
<td>Labour Economics</td>
</tr>
<tr>
<td>Fitzenberger &amp; Völter (2007)</td>
<td>Long-Run Effects of Training Programmes for the unemployed in East Germany</td>
<td>21</td>
<td>Labour Economics</td>
</tr>
<tr>
<td>Fitzenberger et al. (2006)</td>
<td>Get Training or Wait? Long-run Employment Effects of Training Programmes for the Unemployed in West Germany</td>
<td>40</td>
<td>ZEW Discussion Paper</td>
</tr>
<tr>
<td>Frölich (2002)</td>
<td>What is the Value of Knowing the Propensity Score for Estimating Average Treatment Effects</td>
<td>5</td>
<td>Department of Economics, University of St. Gallen</td>
</tr>
<tr>
<td>Gerfin and Lechner (2000)</td>
<td>Microeconometric evaluation of the Active Labour Market Policy in Switzerland</td>
<td>271</td>
<td>The Economic journal</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Title</td>
<td>Cited by</td>
<td>Journal</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>----------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>Hujer, Thomsen and Zeiss (2004)</td>
<td>The Effects of Vocational Training Programmes on the Duration of Unemployment in Germany</td>
<td>49</td>
<td>Allgemeines Statistisches Archiv</td>
</tr>
<tr>
<td>Jespersen et al. (2008)</td>
<td>Costs and benefits of Danish active labour market programmes</td>
<td>34</td>
<td>Labour Economics</td>
</tr>
<tr>
<td>Lechner et al. (2005)</td>
<td>The curse and blessing of training the unemployed in a changing economy: The case of East Germany After Unification</td>
<td>75</td>
<td>Discussion Paper</td>
</tr>
<tr>
<td>Leuven and Oosterbeek (2002)</td>
<td>A new approach to estimate the wage returns to work-related training</td>
<td>20</td>
<td>IZA Discussion Paper</td>
</tr>
<tr>
<td>Leuven et al. (2005)</td>
<td>Worker Reciprocity and Employer Investment in Training</td>
<td>27</td>
<td>TI Discussion Paper</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Title</td>
<td>Cited by</td>
<td>Journal</td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>Lorenzen and Dahl (2005)</td>
<td>Active labour market programmes in Norway: are they helpful for social assistance recipients</td>
<td>19</td>
<td>Journal of European Social Policy</td>
</tr>
<tr>
<td>Ludwig et al. (2010)</td>
<td>Demographic change, human capital and welfare</td>
<td>3</td>
<td>University of Mannheim Discussion Paper</td>
</tr>
<tr>
<td>Mincer (1958)</td>
<td>Investment in human capital and personal income distribution</td>
<td>1132</td>
<td>The Journal of Political Economy</td>
</tr>
<tr>
<td>Parent (2003)</td>
<td>Employer-supported training in Canada and its impact on mobility and wages</td>
<td>16</td>
<td>Empirical Economics</td>
</tr>
<tr>
<td>Pavlopoulos, Muffels and Vermunt (2009)</td>
<td>Training and low-pay mobility: The Case of the UK and the Netherlands</td>
<td>1</td>
<td>Labour Economics</td>
</tr>
<tr>
<td>Raaum, Torp and Zhang (2002)</td>
<td>Do individual programme effects exceed the costs? Norwegian evidence on long run effects of labour market training</td>
<td>21</td>
<td>Memorandum</td>
</tr>
<tr>
<td>Richardson and van den Berg (2001)</td>
<td>The effect of Vocational employment training on the individual transition rate from unemployment to work</td>
<td>53</td>
<td>Working Paper</td>
</tr>
<tr>
<td>Røed and Raaui (2003)</td>
<td>The Effect of Programme Participation on the Transition Rate from Unemployment to Employment</td>
<td>8</td>
<td>Memorandum</td>
</tr>
<tr>
<td>Stenberg (2005)</td>
<td>Comprehensive Education for the employed – Evaluating the effects on unemployment of the adult education in Sweden.</td>
<td>10</td>
<td>Labour</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Title</td>
<td>Cited by</td>
<td>Journal</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-----------------------------------------------------------------------</td>
<td>----------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>Stephan (2008)</td>
<td>The effects of active labor market programs in Germany</td>
<td>?</td>
<td>IBA discussion paper</td>
</tr>
<tr>
<td>Winkelmann and Winkelmann (1998)</td>
<td>Why Are the Unemployed So Unhappy? Evidence from Panel Data</td>
<td></td>
<td>Economica</td>
</tr>
<tr>
<td>Wunsch &amp; Lechner (2007b)</td>
<td>What did alle the money do? On the General Ineffectiveness of Recent West German Labour Market Programmes</td>
<td>15</td>
<td>CEPR Discussion Paper</td>
</tr>
<tr>
<td>Zhang (2003)</td>
<td>Identifying treatment effects of active labour market programmes for Norwegian adults</td>
<td>7</td>
<td>The Frisch Center for Economic Research</td>
</tr>
<tr>
<td>Zwick, 2002</td>
<td>Continuous training and firm productivity in Germany</td>
<td>42</td>
<td>ZEW Discussion Paper</td>
</tr>
<tr>
<td>Zwick, 2005</td>
<td>Continuing vocational training forms and establishment productivity in Germany</td>
<td>38</td>
<td>German Economic Review</td>
</tr>
</tbody>
</table>
Appendix 3 – Members of the expert panel

**René Böheim**

René Böheim is an associate professor at the Department of Economics, Johannes Kepler University, Linz. He studied Economics at the University of Economics and Business Administration, Vienna, and at the Institute for Advanced Studies, Vienna. He received his PhD from the University of Essex, Colchester, in 2002. Prior to working at the Johannes Kepler University in Linz, he was an economic researcher at the University of Munich. Recently, he was a visiting scholar at UC Berkeley and Harvard University.

His research focuses on labor economics, population economics, and applied microeconometrics. Current research includes the analysis of the gender wage gap, the effects of temporary employment on subsequent careers and the economic consequences of divorce.

**Kevin Hollenbeck**

Dr. Hollenbeck is the vice president of the Upjohn institute for employment research and director of publications. He is currently principal investigator and project director for an evaluation of the State of Indiana’s 21st Century Workplace Skills Initiative and for the State of Virginia Workforce Council’s Integrated Performance Information system. He has recently conducted an evaluation of the Massachusetts Workforce Training Fund grant programme and an evaluation of the Focus: HOPE education programmes in Detroit, Michigan. He recently presented papers on “The Effectiveness of Public Subsidies of Incumbent Worker Training” and “On the Use of Administrative Data for Workforce Development Programme Evaluation.”

He is conducting research in the areas of programme evaluation, return on investment and benefit-cost, performance measures, education, school-to-work transitions, formal and informal training, and federal job training policy.

**Jasper van-Loo**

After finishing his master in general and quantitative economics, Jasper van Loo started his career at the Research Centre for Education and the Labour Market (ROA) in March 1997. As a student he was involved in research activities for ROA on a part time basis. His main projects then were setting up a labour market information system for a transnational border region (the Euregion Maas-Rhine project) and a technology indicators project, for which he was responsible for the data collection part. In his first position as a researcher at ROA, he was involved in making replacement demand labour market forecasts for the information system on the Dutch labour market and the employment and schooling observatory project commissioned by the Dutch ministry for employment and social affairs. As a project manager at ROA, Jasper completed a dissertation research project on the relation between training, self-management and labour market outcomes and was involved in projects focusing on learning, training, skills obsolescence, employability and competences. In addition, he led a labour market information system commissioned by the Netherlands federation of Metal-Electronics Industry.

In 2007 Jasper transferred to CEDEFOP, the European agency for the development of vocational training. At this centre, he is involved in labour market research on skill mismatch issues and ageing workers. The main challenge in his work is to support European policies in Vocational Education and Training by sound information and analysis. Jasper is a member of the U.S. Academy for Human Resource Development and has, through this membership, co-
chaired a preconference on continuing professional education for several years. He is currently on the editorial board of two academic HRD journals and acts as an ad-hoc reviewer for journals and books. He has a number of book publications, academic journal articles and popular press articles to his name. His publications deal with human capital and competence development, key skills, ageing workers, skill mismatch, human resource development policies and continuing professional education.

**Alberto Martini**

Alberto Martini is associate professor of Statistics and Public Policy Evaluation at the Università del Piemonte Orientale in northern Italy, where is conducting research of the methodology for evaluating the effectiveness of public policies. He is also directing Progetto Valutazione, a non profit research center whose mission is to disseminate the use of rigorous evaluation methods in Italy. In this capacity he collaborates with the DG-Regio of the European Commission for revising the sections on impact evaluation of the Evalused Guide.

Professor Martini has over 20 years of experience in the application of quantitative methods to policy issues. After obtaining a Law Degree in Italy and a Ph.D. in Economics from the University of Wisconsin-Madison in 1988, he joined Mathematica Policy Research in Washington, to work on the evaluation of social policies by means of experimental and non-experimental methods. In 1993 he joined the Urban Institute as senior research associate to work on microsimulation models to estimate the impact of changes in welfare programmes.

Between 1993 and 1999 he was a consultant of the World Bank for the design of new income and expenditure surveys in transitions countries. He designed and oversaw the implementation of the Income and Expenditure Survey of Belarus and Ukraine.

Since moving back to Italy in 1998, he has conducted evaluations in a variety of fields, including enterprise support, employment policies, and education. He is directing a random assignment evaluation of an employment programme for persons with mental illnesses, and a non-experimental evaluation of active labor market policies. He is member of several advisory boards and recently co-authored a monograph on quantitative methods for programme evaluation. Between 1999 and 2003 he was President of the University Evaluation committee. In 2001-02 he was President of the Italian Evaluation Association.

**Deborah Roseveare**

Ms Deborah Roseveare was appointed Head of the Education and Training Policy Division in the Directorate for Education in June 2007. She takes the lead in providing policy analysis and advice to help governments develop effective and efficient policies for education and learning. She manages a team of policy analysts, economists and support staff as well as being a member of the Management Group of the Directorate for Education.

A dual New Zealand and British national, Ms Roseveare has held several positions in the Economics Department since she joined the OECD Secretariat in 1993 as a senior economist. In the Policy Studies Branch, she has worked on public economics issues including fiscal policy and ageing populations. In the Country Studies Branch, she was Structural Issues Co-ordinator, where she worked on policies to foster entrepreneurship and on structural surveillance, before becoming Head of the Denmark and Sweden desk and then Head of the Canada and New Zealand desk. In the context of EDRC reviews, she has worked on a wide range of education policy and human capital development issues.
Between 1990 and 1993, she was the Economic Counsellor in the New Zealand Delegation to OECD and was vice-chair of the OECD's Economic and Development Review Committee in 1992 and 1993. Before that, Ms Roseveare was a manager within the Budget Management Branch of the New Zealand Treasury and also held posts dealing with primary health care policy, macroeconomics, trade, and economic statistics after joining the NZ public service in 1976.

**Torben Schewe**

Torben Schewe received his diploma in economics from Gottfried Wilhelm Leibniz University Hannover, Germany, in 1999. His major study topics were labour economics, economic growth, income distribution and mathematical economic theory. During his studies he worked as a student research assistant and co-ordinated the EU-students' exchange programme Socrates (formerly Erasmus) for the Department of Economics. From 1999 to 2004 he was affiliated to the Institute for Economics, Chair of Economic Growth and Income Distribution, also University Hannover, and was teaching economic theory. His research work on economic growth and labour market imperfections has been supported by a post-graduate grant of the Federal State of Lower Saxony. During this time, he also conducted several studies as a freelance researcher for a German bank.

Since 2005 he is working for the German Public Employment Service headquarters. After having served as deputy head of the Department of Product- and Programme Analysis from 2006 to 2008, his current position is head of the unit. The department is in charge of the provision of a scientific-based information system, which depicts the effectiveness of active labour market policy programmes in a highly differentiated manner.

**Hilmar Schneider**

Hilmar Schneider is Director of Labour Policy at Germany’s Institute for the study of Labour (IZA). He studied Social Sciences and Economics at the university of Frankfurt/Main. After finishing his diploma as a Social Scientist he worked from 1983 to 1987 as a research assistant in the special collaborative programme ‘Microanalytic foundations of social policy’. He received his doctoral degree in 1987 with a dissertation on the determinants of unemployment duration. Then he held an assistant professorship from 1983 to 1987 in the department of economics at the university of Frankfurt/Main, giving lectures in statistics, econometrics and labor economy. In 1994, he became head of the labor market department of the Halle institute for economic research. Among else within this activity, he worked as a visiting scholar at the World Economy Laboratory at the Massachusetts Institute of Technology in 1998. In July 2001, Hilmar Schneider moved on to the IZA as director of labor policy. Since 2002, he is also a research affiliate of the German Institute for Economic Research (DIW Berlin). From 2006 to 2008 he acted as a member of the National Council for Social and Economic Data, and since 2007 he is a member of the German Census Committee.

Besides labor policy, his main research emphases comprise problems of social protection, wage policy and demography. His most important publications cover papers about the labor market effects of replacement wages, the labor market perspectives of East Germany, the efficiency of active labor policy in the transformation process, and the welfare state perspectives of Europe.
**Appendix 4 – Expert panel report**

**The expert panel**
The expert panel (see membership in Appendix 3) convened on March 26th for a one-day meeting in Brussels and reflected on the basis of a draft synthesis of a selection of 60 pieces of research works (hereafter research synthesis).

The meeting was mainly devoted to:

- Delineating the available knowledge ("lessons learned")
- Identifying knowledge gaps

In addition, the panellists were asked to suggest improvements in the literature review as far as relevant52. A part of the meeting was devoted to methodological issues.

A first version of this report was circulated to all experts, together with a list of questions which have been almost fully answered by all experts. The current version takes stock of the answers and comments received.

**Assessing the draft research synthesis**
The experts suggest the following *improvements* in the report presenting the research synthesis:

- Explaining that, despite ROI was the initial subject of the study, impact of training is de facto the main issue addressed, and also cost-effectiveness to a more limited extent;
- Stressing the fact that general equilibrium effects are not covered in a systematic manner;
- Explaining why adult training is covered and not formal education; mentioning the absence of a broader human capital formation model53 within which training is embedded;
- Mentioning the case of targeting people at risk of quitting the labour force (e.g. ageing workers54), even if no or almost no knowledge is available in this area;
- Assessing the credibility of each reviewed piece of research, and making assessment criteria explicit;
- Clarifying whether there is a rationale for a public sector intervention, and not just a return on private investment and addressing the issue of deadweight, i.e. public sector interventions replacing private initiatives that would have taken place otherwise;
- Recalling that a strategy for identifying causal relationships is required in any impact evaluation; an identification strategy allows for controlling co-influencing factors in order to isolate the impact of the intervention under evaluation;

52 Considerable advice was received in this respect and this is being integrated into the final version of the research synthesis. Some of these references are mentioned in the next footnotes, but not all.

53 In such a model, human capital would be understood as the accumulation of an individual’s experiences to date, including both education and training and a wide range of other factors.

54 In the context of the threat of demographic change and skill shortages, the qualification and activation of people out of the labour force should be an important issue for ESF, and ESF evaluations.
• Extending the coverage of the research synthesis to the USA, mostly to benefit from their experience in generating rigorous evidence;
• Including research studies based on RCT and modelling;
• Referring to works in Eastern and Southern Europe as far as possible.

Methodological issues

Conceptual issues

Most of the reviewed research works produce estimate of the average benefits of “average training” for an average participant\(^{55}\), but this approach may be incomplete.

First, the distribution of impacts across subgroups of targeted people deserves special attention.

Second, impacts might depend on the structure of the programmes, and especially:

• The way participants are approached (e.g. global, individual);
• Whether the programme validates competencies to a larger or smaller extent;
• Whether the training ends with a qualification or not;
• Whether the training programme is sector specific or not.

Third, the benefits accruing from a given training programme should always be seen as a part of a larger picture since the newly acquired skills are just a marginal change in participants’ human capital. Human capital results from previous education and lifelong learning, and it depreciates through skill obsolescence. However, the obvious fact that other things matter should not diminish the desirability of rigorous impact evaluations.

Fourth, the social dimension of training (integration, networking, culture building, cohesion), should be considered as (1) valuable benefits per se, and (2) one of the pathways to economic impacts.

Credibility of micro-studies

Data

The best method is useless if it applies to a dataset of weak quality. Methodological improvements should therefore be conditional on quality assurance in the data collection process.

However, dealing with micro data has always been difficult for the European Commission. In fact, this study is the first opportunity for processing individual data in a cross-national manner at the level of the ESF. This practice cannot be extended at EU level since some national governments refuse to release individual information on labour market status, age, skills, etc, whether it originates from administrative databases or surveys, for reasons of privacy. However, the panellists suspect that “privacy” is conveniently invoked when there are much less commendable reasons to restrict access to the data (e.g. low quality). In any case, micro-datasets should always be available as long as single individuals are not identifiable. To this end, some items of information are suppressed (e.g. identifiers, social security number, address), while others are partly hidden (e.g. displaying annual income information in 1000 euros intervals).

\(^{55}\) In econometric terms, this is called ‘average effect on the treated’
Methods

Evaluating impacts is not just a matter of choosing an appropriate method, but also one of applying this method correctly. A recent research synthesis carried out in Germany\(^\text{56}\) showed that several research studies produced contradictory findings that resulted from control groups designed differently, and not from substantial causes. In this respect, the panel highlights that correct application is an issue that is much larger than just one of sample size. A very important point is that evaluations questions and evaluation methods need to be discussed and selected far in advance in order to collect appropriate data in a timely manner.

Assuming that evaluation methods are implemented correctly, it remains that they have different merits\(^\text{57}\). Applying different methods to the same dataset may lead to different findings and different levels of credibility.

Is there some hierarchy of methods that would have an absolute value, irrespective of the context? –This was suggested in the 1980s, when a number of papers called into question the reliability of econometric and statistical methods for estimating causal effects in observational studies. In particular, LaLonde (1986) and Fraker and Maynard (1987) used data from the National Supported Work (NSW) programmes concluding that widely used econometric methods were unable to replicate the results from experimental evaluations. These influential conclusions encouraged government agencies to insist on the inclusion of experimental evaluation components in job training programmes such as JTPA, Job Corps, and JOBS. One of the best known of such studies is the National JTPA evaluation, which concluded on the impact of a large ongoing policy having some similarities with the ESF.

However, in a later paper, Heckman et al. (1999)\(^\text{58}\) took a more open-minded position and suggested that the best method is the one that best fits in with the context. For instance, they suggested that quasi-experimental approaches are reasonable if the comparison group comes from the same labour market (something that Lalonde was not able to check). From this standpoint, it could be understood that the hierarchy of methods is always relative and contingent to a given situation (including research question and data collection constraints) and that nobody knows the extent to which a method over- or under-estimates impacts in comparison with another one.

Questioned after the meeting on this issue, the panellists express a range of different views.

Hierarchy of methods: relative or absolute?

Do you agree with the statement that “the relative merits of various methods are always contingent to a given situation”?

- I strongly agree


- Any technique needs to be “fit for purpose”. Policymakers are often faced with having to take policy decisions in the absence of high quality evidence, and can’t – or won’t – wait until “gold standard” evidence has been collected. What is important in my view is that the evidence available is presented in such a way that the limits and caveats are not forgotten.

- From a theoretical point of view one would expect a higher internal validity of RCTs than of observational studies. But on the other hand RCTs often lack external validity, are complex and costly and they can generate new unavoidable observation effects.

- Methods need to be chosen with respect to unobserved heterogeneity, non-random selection, attrition, and so on. However, all such problems could be avoided in a perfect world where strong evaluation designs would supply high quality data.

- I see RCT studies on the top of the hierarchy, and simple methods that do not consider selection effects at the bottom of the hierarchy. In between, however, one might have reason to give equal weight to different methods depending on issues like data availability and focus of interest.

- I disagree. RCT is to be preferred if feasible and economical. Natural experiment and Regression Discontinuity are stronger methodologically than non-experimental designs. If one does a non-experimental design, then having observations from the same labour market is preferable to the opposite. Longitudinal designs dominate cross-sections. In short, the evaluation profession has a reasonable pecking order, ceteris paribus.

- I disagree

**General equilibrium issues**

What would happen if everybody was sent to training? Would the benefits identified in micro-level studies be multiplied or divided? When shifting from micro-level impact to macro-level impact, i.e. from impact on individual participants to impact on the whole labour market, a huge number of cause-and-effect loops need to be considered and the general economic equilibrium should be considered.

If it were possible to produce valid impact estimates at this upper level of complexity the findings might be quite different from what is found in micro studies. Positive value judgements might become negative or the reverse

Among the chains of cause-and-effect assumptions that could be considered at this level are:

- A ‘return-to-work’ training programme (1) provides participants with skills that are lacking on the market (skill mismatch) and (2) the quality of the training is sufficient. In such a case, positive impacts at micro-level will be associated with even more positive impacts at macro-level.
- If these assumptions are not true, there is a risk that successful trainees get their jobs at the expense of other people who would have been employed otherwise. This is the so-called substitution effect.\(^{59}\)

---

\(^{59}\) For an overview of the skill matching issue, see the CEDEFOP report "The skill matching challenge, Analysing skill mismatch and policy implication", 2010; On skill obsolescence see De Grip A.(…) "Evaluating Human Capital Obsolescence", Research Centre for Education and the Labour Market, Maastricht University.
A training programme improves the productivity of workers. Due to productivity enhancement the competitiveness of firms increases, and the economy as a whole grows. Then, labour demand increases and it becomes more likely that non-trained workers also get a chance of being hired.

An education programme generates new skills that are not yet on demand on the labour market, but participants convey these skills in their new jobs or create new firms. In turn, this entails innovation, economic growth, and labour demand. This chain of assumptions is often described as like "new skills create their own demand on the labour market"\textsuperscript{60}.

**Transferring lessons**

**Transferability of knowledge**

There is virtually no likelihood that any aggregate impact estimates will hold true for all investments and all target groups. Rather empirical evidence may be used to draw transferable lessons. Such lessons should be expressed in the form of ‘laws’ like “a given type of training benefiting a given public generates a larger return than ...”

For a ‘lesson’ to be drawn, there should be:

- Comparable impact estimates applying to at least two different types of investments / target groups;
- Well specified types of investment applying to well specified target group;
- A shared understanding of where observed differences come from and how.

The transferability of evaluation findings can be assessed on a four-level scale as follows:

- **Internal validity:** the finding is fully credible with regards to the samples under study;
- **External validity:** the finding can be generalised to the population from which the samples were drawn (all participants or the whole targeted population);
- **Scalability:** the finding can be generalised to a wider population, e.g. from pilot regions to the whole country;
- **Transferability,** i.e. the finding enables to draw a lesson\textsuperscript{61}, and this lesson may be used in the context of another population which is similar enough.

**Learning across countries**

Members of the expert panel expressed several views as regards the transferability of lessons learned, ranging from pessimistic to optimistic ones:

\textsuperscript{60} Much of the current policy discussion around education, training and skills is based on the assumption that additional education and training will, by itself, drive the development of jobs using the skills they have acquired in the education and training. This assumption is rarely made explicit or even challenged in the policy debate and yet there is certainly plenty in the economic literature that would lead one to challenge this assumption. At the very least, the validity of the assumption that education and training creates its own demand should be rigorously tested.

• Careful caveats should be stated before applying the lessons learned in a given country (e.g. Germany, USA) in another one.

• The effects of training depend on labour market institutions and this is an obstacle to transferability. However, this is not a black-and-white matter but rather one of degree.

• We can generalise in some instances but we have to identify the contextual factors that may limit transferability. However, papers written for academic purpose tend to emphasise internal validity and to tell little about transferability.

• Using lessons learned elsewhere is not just taking good idea/practices from the shelves, but making intelligent use of the accumulated knowledge about what works better or worse in similar contexts.

The panellists quoted several examples of contextual factors that could restrict transferability:

• The fact that people tend to combine education and work in a higher or lower proportion may entail quite different job search strategies, something which affects the outcomes of the education programmes.

• The balance of incentives/sanctions associated with return-to-work training. This balance is quite different across countries (e.g. DK/DE), and it affects the outcomes of training programmes in many ways, including the so-called ‘threat effect’ (people take a job in order to escape activation measures).

Questioned after the meeting on the issue of transferability, the panellists express relatively convergent views which range from “optimistic but ...” to “pessimistic but ...”. A typical view is that there are some transferable lessons, although with careful caveats, whilst impact estimates cannot be generalised.

<table>
<thead>
<tr>
<th>Should we be optimistic about the possibility of transferring lessons?</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Optimistic: however, this requires a more than mechanical assessment of existing findings.</td>
</tr>
<tr>
<td>- Optimistic: we can generalise sometimes, but we have to identify those factors that limit transferability and we need not to take the results too literally. Generalisation may apply to certain tendencies and causal relationships, not to absolute answers to particular questions.</td>
</tr>
<tr>
<td>- Optimistic: transferring knowledge geographically or across programmes or dynamically is one of the reasons why we evaluate them. Impact estimates are not transferable, but general lessons are.</td>
</tr>
<tr>
<td>- Optimistic on transferring ways of achieving impacts, but pessimistic on transferring impact estimates,</td>
</tr>
<tr>
<td>- Cautiously optimistic</td>
</tr>
<tr>
<td>- Cautiously optimistic: countries can and do learn a great deal from looking at what others are doing, but sometimes do not exercise enough caution in doing so. Some fads in education can travel round the globe at alarming speed! While we can draw broad conclusions for the nature of policy development, we only make country-specific recommendations after looking carefully at the country context. Even then, we stay at a very general level and suggest them as &quot;pointers&quot;</td>
</tr>
</tbody>
</table>
Learning from experience

The whole issue of learning, using and transferring lessons relies upon the assumption that there is sufficient willingness to learn at the level of the authorities managing ESF programmes in the Member States. Some of the panellists question that assumption on the grounds that managing authorities feel under ‘pressure for spending’, something that does not create an incentive for learning. Furthermore, panellists worried that the emphasis put on indicators in the context of the ESF may be taken as an excuse for not doing sound impact analyses.

In the case of the next generation of the ESF, there needs to be a shared willingness (EC and Member States) to improve the learning process by (1) inserting strict requirements in the legislation, and (2) incorporating recommendations in the agreements made between the Commission and managing authorities. However, the Commission does not have any direct power on project promoters and cannot require them to produce specific data or to apply specific evaluation methods.

The context of learning is also quite different across EU countries:

- In some Member States (e.g. DK), ESF funding amounts to only 2% of national spending in the area. In this situation, ESF is used for introducing new ideas, supporting pilot projects, or bringing labour market institutions to work in innovative ways. Evaluating and accumulating knowledge are quite important issues in such a context.
- In some Member States (typically newer ones), ESF funding represents the bulk of active labour policy of the whole country. In such a context, it is very important to better understand what works and what does not work for whom.

Some ideas were raised by panellists as regards the requirements and recommendations to be developed for the next programming cycle of the ESF:

- Requiring that the rationale for public intervention be made explicit;
- Promoting tailor-made solutions in the design of the next programmes;
- Orienting the efforts of monitoring and evaluation towards long-term impacts (what is expected to be achieved in 5 years time) rather than immediate results;
- Creating institutional arrangements that can accumulate valid knowledge and translate it into the language of policy-makers;
- Disseminating information on what does work or does not work and why, types of training that are particularly (un)successful, and things that can help to go forward without negative effects;
- Creating incentives that would encourage managing authorities to launch well designed impact evaluations (not necessarily relying on any specific method), especially where innovative approaches are tested; Considering some reallocation of resources from monitoring to impact evaluation.
Identifying knowledge gaps and suggesting that future evaluations focus on such areas in priority.

Questioned after the meeting on the issue of recommendations, the panellists raised various ideas which are listed in the following box.

**How could we learn faster?**

- Large-scale funding needs to be made conditional on a positive pilot study. Each fund needs to pass a carefully designed and scientifically accompanied study, preferably involving randomised assignments, before rolled-out to the general public.
- Different policy options should be assessed on their relative marginal benefits.
- Political orientations (e.g. supporting disadvantaged groups) should be implemented through evidence-based decisions.
- Due to the nature of the investment (training), one would expect that positive results may occur over a longer time span than the short-term perspective considered in many studies. Therefore, we should think about ways to gather data and results before a particular programme starts and to keep collecting data after a programme has finished.
- Next to generating valid impact estimates, evaluations should also consider intermediate causes and effects, e.g. competence development, transfer of training benefits to the workplace. Such issues are well known in the Human Resource Development literature, but are not considered enough in impact evaluations.
- Careful investigation of validity and credibility of existing studies. This means going beyond stating contradictions by trying to identify the sources of contradiction. This will lead to a systematic assessment of knowledge gaps.
- The Commission cannot require, but can provide: 1. strong encouragement to take the “what works” question seriously/2. training in proper methods/3. dissemination of correctly done impact studies/4. “naming and shaming” of lousy impact evaluations.

**Investing in the employed**

This section and the next one are intended to delineate the available knowledge (lessons learned) and the knowledge gaps.

**Cost of training**

Most of the reviewed studies focus on benefits and only a few articles address the issues of (1) cost of the various types of training and (2) sharing of this cost between enterprises, individuals, and public institutions.

This issue is assessed as a **knowledge gap**, not that much because costs would be difficult to measure, but mainly because impacts and costs tend to be measured and analysed separately, thus rendering cost-effectiveness lessons extremely rare.
**Trainees’ benefits**

Many studies focus on the short term impact on salaries, with convergent findings that it is *not that important*. Questioned after the meeting on this issue the experts broadly agree with this finding.

<table>
<thead>
<tr>
<th>Is it true that short term impact on salaries is limited?</th>
</tr>
</thead>
<tbody>
<tr>
<td>- partly agree</td>
</tr>
<tr>
<td>- agree on the ground that wage setting mechanisms are not flexible; that salaries as well as other outcome variables are more important in the mid to long term perspective</td>
</tr>
<tr>
<td>- agree but underline that the focus on short-term financial gains in terms of salaries implies that other important impacts are not considered. In the case of training to combat skills obsolescence, training makes up for loose of human capital and results are likely to materialise in terms of sustained labour market participation, not salary increases.</td>
</tr>
</tbody>
</table>

From this point of view, it would be useful to know about comparative impacts of programmes having different structures, different ways of reaching their public, different training approaches. However, little is known about these issues.

Other types of benefits would deserve to be studied such as job satisfaction, career management possibilities and mobility. This is another knowledge gap.

Of particular interest for the ESF is the case of *people at risk* of losing their job, not only because EU money is massively investing in this area, but also because there is a strong argument supporting such public interventions in order to compensate for enterprises’ under-investment.

However, the panellists consider that this category is not yet defined in a stable enough manner, and that a definition of workers at risk of losing their job is probably impossible to construct. The panellists recommend analysing this issue on a continuum from secure employment to unemployment through various degrees of risk associated with factors such as obsolescence of sector specific skill, or fragile contract employment. However, information on employability is very difficult to collect for employed individuals.

Unfortunately, this ESF priority has not been one for researchers and little is known on this issue. For instance, the US programmes targeted at dislocated people include people at risk, but no study focuses on them specifically. An IZA study deals with the danger of losing one’s job and shows that the results of supported training are not better than if participants were assisted by employment services. Moreover, the programmes that are supposed to prevent risks are often used as early retirement schemes.

**Employers’ benefits**

Many studies focus on the impact of training on productivity, with convergent findings that it is *important*. Questioned after the meeting on this issue the experts broadly agree with this finding.

<table>
<thead>
<tr>
<th>Is it true that impact on productivity is important?</th>
</tr>
</thead>
</table>

135

Euréval / Ecorys / Rambøll Management
- Fully agree, and add that this result also provides some rationale why employers are inclined to invest in training in the first place.
- Partly agree, as according to the evidence shown but stay sceptical on the basis of other studies.
- Agree, but consider that this is a second order issue. The main aim should be to prevent unemployment. Productivity enhancement may only be an indirect vehicle and is hard to measure.

However, the credibility of these findings may be questioned because most of the benefits of training are conditional to some transformation in the companies. Research works in the area of human resources\textsuperscript{62} on High Performance Work Places (HPWP) show that training is just one of the factors that have to be combined altogether for a firm to be successful.

Human resources research also includes a stream of cost benefit studies focusing on the private return on training\textsuperscript{63} and involving before-after comparisons. However, the panellists are sceptical about the reliability of these works.

Instead of studying the impact of training on productivity, it would be preferable to refine the analysis and to consider successively:

- The purpose of training which may range from literacy and other basic skills\textsuperscript{64}, to highly technical competencies;
- The results of training in terms of knowledge, skills and competencies and the transfer of training to the workplace (a fundamental causal link which is currently missing);
- The connection between competencies and second order benefits such as productivity, innovation\textsuperscript{65}, social relationships, competitive advantages\textsuperscript{66}.

As already said earlier (Conceptual issues, p128) competencies depend on the overall human capital of people, which results from education, on-the-job training, off-the-job training and informal learning. It would be wise to consider the impact of training in this broad perspective.

**Social return on investment**

Public institutions may wish to subsidise the training of employees in case of under-investment in the private sector something which may happen for two reasons:

---


\textsuperscript{65} See Canadian study showing on skills development and firm needs showing that skills are a component of innovation strategies.

\textsuperscript{66} See Barron, Black, and Berger, (1997) On-the-Job Training, Kalamazoo: MI: W.E. Upjohn Institute, especially the chapter on productivity. Similar works were done in Australia.
Companies facing economic difficulties tend to fire vulnerable workers or to use early retirement schemes. Investing in training would improve the employability of such people and generate social benefits in the long term, but short term private returns are not worth the investment (see above Trainees’ benefits, p135).

Healthy companies do not gain the full benefits of a training investment because some of the participants move to other companies where they sell their human capital at a higher price (poaching risk)\(^67\).

These assumptions have to be put on the table, clearly identified, and reviewed in a detail per sector (e.g. emerging technology, structural change), type of company (especially SMEs in the context of the ESF), type of target groups (at-risk and others), type of training, etc. Questioned after the meeting on the issue of underinvestment the experts express diverging views.

### Does under-investment in the private sector justify that public money is invested in training?

- **Disagree**, as there is not enough evidence on this matter. Ideally there should be some assessment of the extent to which public training investments replace private investments that would have taken place in the absence of public training programs. In addition, the newer literature on employability puts forward that employers can attract and keep human resources by investing in their general training.

- **No opinion**, as while the argument for under-investment may be quite strong, the degree of under-investment may be small and not be enough to offset the distortions induced by taxation or borrowing to finance any public support. This needs to be tested empirically.

- **Fully agree**, if under-investment is a fact. However, it may be difficult to provide evidence for such a statement. The fact that firms are making use of public support is not a sufficient proof for private under-investment. In fact, public support may generate deadweight and crowd out privately funded training. It could also be the case that publicly funded training is abused for buffering under-utilised training capacity. A careful examination of the under-investment hypothesis is therefore necessary.

- **Agree**, but think that it should be verified, and that it may be difficult to identify the companies where public support is advisable.

Business cycle is also an issue since companies are assumed to invest more in training in case of skills shortage, i.e. at the high of the cycle. In the USA, a study\(^68\) examined the public expenditure in training (2% of total training investment) in 50 states. The targeted firms were facing economic difficulties. Impact estimates ranged from null to positive.

### Investing in the unemployed

For many years, an important part of ESF support has been targeted at vulnerable unemployed people such as younger and elder unemployed, long-

---

67 There is a stream of studies on trainees quitting their enterprise. Becker and Mincer are seminal works.

term unemployed, or migrants. Vocational (re)training is assumed to be a pathway to (re)accessing the labour market and to social inclusion. Both active unemployed people and inactive people have been reached in a proportion that depends on the business cycle.

**Training or not training?**

Public support to training is often a part of a broader active labour market policy including job-search assistance, placement, incentives and sanctions.

There are convergent findings that suggest that **training alone has lower impact** than a mix of training and other instruments such as job search assistance or sanctions.

For instance, a study by the Northern Ireland Employment Agency shows that interventions involving job-search incentives have a stronger impact than other ones.\(^69\)

In the USA, a long series of studies has addressed this issue. A recent piece of research covered many states and concluded that training through the Workforce Investment Act (WIA) was positive for disadvantaged and dislocated people. Other studies showed the contrary as regards dislocated workers. The discussion continued by comparing job search assistance and training since the first instrument costs far less than training and is known to be effective whilst training doesn't pay for everybody. The relative ineffectiveness of training may be tied to the severe budget cuts in public labour market institutions and the use of a voucher system for training without adequate counselling.\(^70\)

Some studies apply to programmes and schemes that mix training and incentives / sanctions. They investigate the issue of "threat effect"\(^71\), i.e. people looking more effectively for a job in order to avoid being enrolled in the programme. Threat effect was confirmed in Denmark where the system of incentives / sanctions is particularly developed\(^72\), and in the USA.\(^73\)

Finally, it is suggested that this finding be investigated in more depth in order to better understand how instruments interact with one another. For instance, in Germany, participation in long and expensive qualification programme is often conditioned by the successful finalisation of a shorter programme that assesses skills, motivation, and availability. Some instruments may also deteriorate the employment prospects of participants, e.g. by generating stigma effects.

\(^{69}\) Mc Vicar (2006) "Job Search Monitoring Intensity and Unemployment Exit: Evidence from a Jobseeker’s Allowance Natural Experiment"

\(^{70}\) In Germany, the effectiveness of training seems to have improved since vouchers were introduced in 2003


\(^{72}\) Brian Krogh Graversen, Brian Larsen (2008) "Is There a Threat Effect of Mandatory Activation Programmes? New Evidence from Denmark", 22nd conference of the European Society for Population Economics, London (this is a presentation – Is there a published paper?)

**Lock-in effect**

While participating in a training programme, it is harder for people to look for a job. This is the lock-in effect that applies to unemployed trainees and that mirrors the effect of ‘foregone productivity’ pertaining to employed trainees.

In some programmes, the lock-in effect is minimised by providing a mix of training, placement, and incentives for job search. Such programmes encourage ‘positive drop-outs’, i.e. people abandoning the training for getting a job.

Of course, it is expected that lock-in effects be compensated by long term employment impact, and this means that the balance of benefits and drawbacks should be established on a long enough time period in order not to overemphasise lock-in effects. There was convergence among panellists that lock-in effect is an unavoidable part of the training process and does not deserve major attention.

**Which benefits?**

As seen above (Training or not training?, p138) training alone may have limited impact but programmes that mix training and other instruments are more successful. For unemployed people, success is most often assessed in terms of finding a job, sometimes in terms of finding a stable or a well paid job. In a perspective of return on investment at the level of the society as a whole, it might be worth looking at the subjective value of employment for individuals, e.g. measured by the happiness index.

**Groups in need of training**

Considering the questionable impact of pure training, it is of utmost importance for public institutions to target those participants for whom training truly makes a difference. In other words, public programmes should strive to prevent the well-known ‘creaming effect’ in which field level operators enrol participants who have the best chances to finding a job.

This problem was exemplified in a recent study covering a number of German labour market districts74. The study found that training measures targeted the wrong people in recruiting the most employable candidates whilst the benefits were found to be much stronger for old age trainees who were under-represented among selected participants.

The panellists highlighted some specific groups who may gain special benefits from training programmes:

- Young people having faced school failures;
- Educated young people suffering from skill mismatch
- Old people and migrants suffering from skill mismatch75

**Women**

There are convergent findings that women benefit more from training

An explanation may be that women interrupt careers for birth and tend to participate in training programme in the context of their return-to-work strategy. Occupational effects might also play a role.

---


A counter example can be found in an Australian study\textsuperscript{76} which show that men had a higher return than women because job perspective were lower for women. One of the impact estimates made within the current study (Poland) also show an exception where long-term unemployed women succeeded three times less than men in finding a job.

Questioned after the meeting on this issue the experts expressed mixed views.

\begin{tabular}{|p{\textwidth}|}
\hline
**Are unemployed women benefiting more from training than men?**
\hline
- No opinion, as the evidence seems mixed. In any case there could be some more background to this finding. What are the mechanisms behind the larger benefits from training for women?
\hline
- If so, this is not so much a matter of gender but rather a matter of specific circumstances, which may be more likely to occur with women.
\hline
- Disagree, as this statement is very broad and policy interventions should not solely be based on gender. It is important to consider additional issues that women are typically confronted with as single parents, child care, working time etc. We cannot observe that women generally benefit more from training than men. The heterogeneity of individuals could be higher in the group of women than between women and men, and the effects of training depend on many additional factors. Hence it is important to see the whole set of indicators and make investments dependent of the individual need for qualification.
\hline
\end{tabular}

**Targeting groups or individuals?**

The management of the ESF has paid increasing attention to targeted groups, especially women and vulnerable groups. This approach is considered as the single best way to prevent the creaming effect and to ensure redistribution.

There are voices saying that Member States should be left with the full responsibility of identifying the groups to be targeted, but the Commission considers that the focus on specific needs is part of the value added Europe can provide in this area. Incidentally, it can be observed that the European focus on vulnerable groups has changed the culture of monitoring and evaluation in the Member States.

However, some panellists express doubts about the prospect of preventing the creaming effect by targeting large groups, and consider that this goal could be better achieved by lay level operators having both discretion and more information. An alternative strategy would be to focus on individual employability instead of target groups. However, several Member States argue that they cannot select participants on the basis of employability because this would breach ethical principles and privacy.

It is interesting to see that this alternative strategy was chosen in the USA. Unemployed candidates to public training programmes have to be profiled by law. This is done by completing an employability scorecard, which is converted into an employability index on the basis of coefficients derived from regression analysis. The issue of ethics and privacy was discussed with the result that gender and race were excluded from the calculation. The whole

system is said to involve a low administrative burden. Some European countries (e.g. NL in the past, DK at present) also apply a system where job-seekers are allocated to different groups according to their employability. This suggests that **individual profiling is a feasible alternative to targeting social groups**.

Questioned after the meeting on this issue the experts broadly agree with this statement.

<table>
<thead>
<tr>
<th><strong>Is individual profiling a feasible alternative to targeting social groups?</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Partly agree, as this is not a contradiction, as those who have a low score constitute a social group, namely persons who have a low employability score.</td>
</tr>
<tr>
<td>- Fully agree, and find it desirable. As, taking social groups can be misleading as typically there are large differences between individuals between these groups in terms of their employability and distance to the labour market; also as individual profiling is superior to targeting social groups. Defining social target groups will never be able to capture complex combinations of certain degrees of indebtedness, partnership trouble, burn-out and the like.</td>
</tr>
<tr>
<td>- No opinion</td>
</tr>
</tbody>
</table>

**Business cycles**

There seems to be converging findings that the benefits of training unemployed people depend on the business cycle. However the **magnitude of this relationship and even its direction remains uncertain**. The research synthesis concludes that the benefits from training are higher at the height of the business cycle when the demand for skilled workforce is high. On the contrary, a very well-done and credible Canadian study\(^77\) shows that that training was more effective during recession periods. Finally, another study\(^78\) suggests that the opposite may be true for underprivileged groups.

<table>
<thead>
<tr>
<th><strong>Is training yielding larger benefits at the high of the business cycle?</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Disagree, as the evidence is not conclusive</td>
</tr>
<tr>
<td>- Disagree, and are more convinced by German and Canadian studies that say training is countercyclical</td>
</tr>
<tr>
<td>- Question the usefulness of this finding, as it may be difficult to reproduce results on the same training scheme in another business cycle. In a long term perspective, with cyclical patterns over a five to ten years period it is likely that several economic, legal, and political factors may also have changed, which could be influential for the impacts of training activities.</td>
</tr>
</tbody>
</table>

---


A connected issue is that of public finances. In case of tight budgetary constraints on the labour market policy and employment services, there are changes in the implementation of the employment policy, with potential consequences on the impacts.

**Training delivery**

In the context of the ESF as well as in many Member States, training is delivered through a contracting out process. Therefore a very important issue is that of the effects of various types of contracting processes, depending on whether they focus on resources or performance.

A recent research work carried out in the Netherlands\(^7\) shows that a contracting process was driven by traditional public procurement principles, and created a strong advantage for large national competitors offering low prices. Small local suppliers capable to adapt training to local needs had no chance in the competition. Contracting out is therefore to be considered carefully in order to achieve quality and results.

**Main conclusions**

**Substantial issues**

The panellists reached a quasi unanimous agreement on the following points:

- General equilibrium effects matter and need to be considered;
- It is likely that there are some specific areas of under-investment in training by the private sector, something which creates social justification for public investment;
- In the short run, training is likely to produce productivity gains for employers, and unlikely to pay off in earnings for workers;
- Training for unemployed individuals may be ineffective, unless it is used in conjunction with other instruments within a consistent activation policy.

There was a relatively good consensus on the following points:

- In order to avoid the ‘creaming effect’, profiling of the unemployed on an individual basis is a feasible alternative to targeting specific social groups.

The following issues were considered as knowledge gaps:

- Nature and magnitude of the main general equilibrium effects (e.g. substitution effect)
- Specific areas of under-investment in training by the private sector, especially as regards at-risk employees
- Comparing the impact of training on unemployed women and men respectively.
- Understanding the relationship between the impact of training and the business cycle.

**Methodological issues**

The panellists agreed on the following points:

\(^7\) P. Koning, C. J. Heinrich (2010) “Cream-Skimming, Parking and Other Intended and Unintended Effects of Performance-Based Contracting in Social Welfare Services” Iza DP No 4801
Data quality is of the utmost importance;
Privacy issues should be manageable and should not restrict the development of sound impact evaluations;
RCT is likely to produce strong evaluation evidence and there is no justification for neglecting this method as it has been almost the case up to now in Europe;
The impact of training should be assessed in the long run, knowing that there are a range of short term negative effects (e.g. foregone productivity, lock-in);
The transferability of lessons learned in evaluation is likely to be possible, although it must be done carefully with caveats;

Some methodological challenges are still ahead:
- Nature and magnitude of the main general equilibrium effects (e.g. substitution effect)
- Specific areas of under-investment in training by the private sector, especially as regards at-risk employees
- Connecting the costs and impacts of training into some cost-effectiveness or cost benefit analysis.
Policy and financial framework

Implementation of the European Employment Strategy in Flanders

Objective 3 forms the most important part of the European Social Fund budget in Flanders. During the period 2000-2006, 60 million Euros a year were available for the implementation of the European Employment Strategy in Flanders. In order to steer the implementation of the programme, Flanders has developed one single programme ("Enkelvoudig Programmeringsdocument" or EPD3). The European Employment Strategy served as a guideline for the Flemish EPD3, based on four axes: employability, entrepreneurship, adaptability and equal opportunities.

The axes for Objective 3 in Flanders were:
- Axes 1&2: employability (preventive and curative approach);
- Axe 3: entrepreneurship;
- Axe 4: adaptability;
- Axe 5: equal opportunities – gender;
- Axe 6: experiments and pilot projects.

Implementation of ESF Objective 3 - Axes 1&2 in Flanders

In this qualitative study, we will focus on Axes 1&2 because of the importance of these axes for the Flemish Labour Market Policy.

Axes 1&2 aimed to improve the employability of jobseekers, through guidance and training. Each unemployed received a 'universal' provision of services. Within this service, route counselling and custom-made guidance was organised, adjusted to the individual needs of the unemployed on the labour market. A distinction was made between short term unemployed and long term unemployed persons.

During the previous Objective 3 programme 1994-1999, the principles and method of route counselling were developed. During the programme 2000-2006 the practice was optimised. Different elements were elaborated: more efficient routes or pathways to integration, a better match between different parts of a trajectory, more accurate monitoring and guidance of each unemployed in a trajectory.

In principle a trajectory consists of following parts:
- intake and orientation;
- training and vocational training;
- guidance and placement on the labour market;
- job consolidation.

A distinction was made between services for short term unemployed (axe 1) and long term unemployed people (axe 2). The long term unemployed needed more intensive guidance towards the labour market than the short term unemployed:

Axe 1: preventive approach:
• young unemployed, less than 25 years old and less than 6 months of unemployment;
• unemployed older than 25 years old, less than 12 months of unemployment;

Axe 2: curative approach: long term unemployed

Monitoring and evaluation system

In the EPD3 2000-2006 of Flanders, monitoring and evaluation arrangements were linked to each other. The mid term and final evaluation consist of three major elements:

• Recommendations for the monitoring system including the definition of relevant indicators;
• Process evaluation on programme and project level;
• Effect evaluation on programme level, project level and client level.

The main evaluation questions for the different axes of the programme were:

• Axes 1 & 2: effectiveness of the programme on client level; innovation and added value of the programme;
• Axe 3: effectiveness of initiatives for potential starters and sustainable entrepreneurship;
• Axe 4: reach and effectiveness of training activities for employees;
• Axe 5 and 6: effects of experiments, pilots and studies.

The effect evaluation on programme level refers to following elements:

• The effectiveness and efficiency of the programme: realisation of the quantitative and qualitative aims of the programme in relation to the budget;
• The macroeconomic effects of the implementation of the programme:
  o Job creation (f.i. jobs created by work experience, social economy, job-sharing, local initiatives, ...);
  o Collateral effects on employment (f.i. additional employment for counsellors, teachers, advisors, ...);
• The structural effects of the programme on the Flemish labour market policy:
  o Strategies of active labour market policy;
  o Generalisation of innovative methods and instruments;
  o Strategies and partnerships for sensitisation.

The effect evaluation on project level refers to following elements:

• Factors determining success and failure for effective and efficient implementation and process of the programme;
• The effect of innovative and successful implementation of measures, institutional arrangements and cooperation between actors;

The effect evaluation on client level refers to the direct beneficiaries (f.i. clients, participants, organisations, companies, ...). The type of effects differs, depending on the profile of the target group and the nature of the measure. Elements of evaluation are:

• Client satisfaction;
• Unemployed: net-labour market effects, ...;
• Persons threatened by exclusion: increase of employability in terms of individuals possibilities to integration into the labour market (qualifications, training, ...);
• Adaptability of individuals: mobility of employees, type of employment contracts, ...;
• Entrepreneurship: qualifications of entrepreneurs, income, innovation, ...
• Women: position of women on the labour market in terms of equal opportunities, level of responsibility, ...

**Good practice**

**Evaluation and impact assessment Axes 1&2**

The central question in the evaluation and impact assessment of the axes 1&2 was: which target groups follow which trajectories to the labour market, against what kind of effects in relation to the budget?

The two main elements in the effect evaluation of the axes 1&2 were:

- Innovation and added value of the programme
- Effectiveness of the programme (micro level)

As shown above, the ESF axes 1&2 have a large added value to the Flemish labour market policy, in financial terms, in terms of volume as well as in terms of innovation in methods and organisation of route counselling for unemployed persons. In the single programme of Flanders, the Flemish employment policy objectives and the performance indicators correspond to the European Employment Strategy. During the ongoing evaluation ESF Objective 3, the integrated employment strategy is reflected in the minutes of the monitoring committees of the axes of the ESF Objective 3 programme in Flanders and in the employment strategy of the public employment service ‘EMPLOYMENT SERVICE’. Interviews on policy level during the ongoing evaluation confirm these findings.

**Innovation**

During the programme period 1994-1999 the most important innovation was the method of route counselling for unemployed. In the programme period 2000-2006 the innovation and added value related to the monitoring system and the steering capacity of the ESF-related stakeholders.

**Monitoring**

The client follow-up system is developed by the Public Employment Service and makes it possible to monitor the activities and measures for unemployed an accurate way. The system allows the Public Employment Service, and in extension the ESF-policy stakeholders, to check and steer the reach of the unemployed on a permanent basis. Moreover the system makes it possible to measure the labour market results of the trajectories of pathways to integration for the unemployed (gross effects). Both elements were crucial in order to steer the programme during the programme period 2000-2006.

**Result driven**

The programme was partly result driven. A part of the budget of ESF-trajectories depended on the results in terms of reach and placement of the target groups. Possible side effect of this result driven approach was the creaming off of the target group in order to get better labour market results.

**Partnerships between actors on the labour market**

A trajectory sometimes exists of a sequence of diverse activities, organised by different actors on the labour market. To make these kinds of trajectories successful, cooperation and partnerships between actors on the labour market were essential.
**Added value**

Different forms of added value can be indicated.

**Reinforcement of the volume of the target group**

A very important added value of the ESF-budget is the reinforcement of the volume of route counselling activities for unemployed. Without the ESF-budget, a considerable amount of trajectories would not have been implemented.

**Increased efforts for certain target groups**

The overrepresentation of specific target groups like long term unemployed in the ESF-programme can be interpreted as an added value of the programme. ESF-trajectories represent more specific target groups with difficulties on the labour market than non-ESF trajectories.

**European Employment Strategy**

The measures in the axes 1&2 fit perfectly into the philosophy and the principles of the guidelines of the European Employment Strategy during the programme period 2000-2006.

**Effect evaluation on client level**

The client follow-up system for the axes 1 & 2 is a tool for guidance and follow-up of the unemployed in Flanders. During the programme period 2000-2006 the client follow-up system is used as a tool for monitoring and evaluation as well.

The method and organisation of route counselling for unemployed persons is developed during the former ESF Objective 3 period (1994-1990) and is supported by the Flemish policy makers and the public employment service during both periods (1994-1999 and 2000-2006). External ongoing evaluation has always been important to monitor and adjust the Flemish employment strategy. The quality of the methodological approach of the evaluation of the axes 1&2 can be explained by (1) the strength of administrative capacity, monitoring and statistical databases, (2) the institutionalisation of evaluation and (3) the need for acquiring highly credible knowledge. Because of the importance of the data sources of the public employment for the evaluation of the axes 1&2, the evaluation methodology could not be implemented in the other axes of the Objective 3 programme. In axe 4 for instance, the impact has been measured by the use of a survey.

**Labour market status**

In the final evaluation of the programme, special attention is given to the effects of the measures of the axes 1&2.

The follow-up of the labour market results of trajectories, subsidised by ESF, was important in order to management the ESF-budget and was crucial to steer the actions in the programme. The ESF-policy stakeholders have published the labour market results on a regular basis. The labour market status was measured at the end of a trajectory (client employed or unemployed).

**Impact**

Two evaluation approaches were followed:

In a first approach, each participant who has been unemployed for x months at the end of the trajectory (target group), is compared with a non-participant with a similar profile, who has been unemployed for the same x months, at the end of the trajectory (control group). The impact estimates were positive.
for the target group in ESF-trajectories (80% had a faster trajectory than the control group).

In a second approach, a comparison was made at the start of the trajectory. Each participant who has been unemployed for x months at the start of the trajectory (target group) is compared with a non participant with a similar profile, who has been unemployed for the same x months at the start of the trajectory. In this approach, the impact is significantly lower than in the first approach (only 40% had a faster trajectory than the control group).

Moreover, a preventive approach (axe 1) with limited trajectories in time and in type of guidance (formal training and dual learning in terms of training and work experience) seemed to be less effective than a curative approach (axe 2) with more intensive trajectories for long term unemployed.

The results cannot be interpreted as a negative evaluation of the method of route counselling. The evaluation has shown that the method and instruments for route counselling and custom-made guidance need to be implemented in a more selective way for specific target groups who are long term unemployed. Evaluation has pointed out that the method of counselling has a significant positive net-effect in the curative approach.

These results have been shown in the midterm evaluation and final evaluation reports during the ongoing evaluation of ESF Objective 3 axes 1&2. Besides, several scientific studies on the same data sources, have confirmed the results of the external evaluation. For instance, the Flemish employment policy in relation to the European Employment Strategy has been evaluated by order of the federal government and the European Commission (2004-2005).

**Evaluation results as an opportunity for learning**

Interviewees express convergent views that the findings of the evaluation studies have been intensively used by policy makers in Flanders (government, administration, public employment service, etc.).

**Dissemination activities** towards guidance and training organisation were limited to classical communication channels like website, newsletters and seminars.

Several **scientific studies** on the same data sources, have confirmed the results of the external evaluation. For instance, the Flemish employment policy in relation to the European Employment Strategy has been evaluated by order of the federal government and the European Commission (2004-2005).

The **direction of the employment policy** in Flanders has not been changed since the evaluation of the ESF Objective 3 programme axes 1&2. On the contrary, the policy direction has been continued during the new programme period 2007-2013. Moreover, the guidance and training organisation were obliged to work according the methods of route counselling for specific target groups through tendering procedures organised by the public employment service.

**Information sources:**

Annual reports ESF-Objective 3 EPD 2000-2006

Minutes of the meetings of the Flemish Monitoring Committee (Vlaams Monitoring Comité) of ESF-Objective 3 EPD 2000-2006

Vlaanderen, EPD 2000-2006. Evaluatie halverwege de looptijd. HIVA, Leuven, IDEA Consult, Brussel. (Mid-term evaluation)


Appendix 6 – Learning from impact evaluation: Hungarian case study

In Hungary, the intention to measure the impact of the labour market interventions has been present in the labour research community for more than a decade and to some extent it has also regularly been addressed by the labour market institutions and by the Ministry responsible for employment policy as well. In spite of the undoubted efforts, the learning paths of impact assessment have been fragmented and mainly localised to academic circles. The regular monitoring reports on the use of the domestic and/or EU investments financing active labour market policies (ALPs) were prepared by the labour administration and had been based on descriptive data provision which did not allow deeper analysis other than comparing subgroups of the data according to the available variables. Impact assessment based on comparison with counterfactual data has rarely been attempted. An important impact study was carried out at the end of the nineties. In the first phase of the study, the researchers used counterfactuals, while in the second turn the same data was used to explore the role of different impact factors as age, education and location. The results did not justify certain large scale interventions especially not the expected impact of public work. As far as the effectiveness of other interventions as training and entrepreneurial support was concerned, the results were somewhat more positive.

The HRDOP of the 2004-2006 National Development Plan intended to develop the effectiveness of the Public Employment Service (Measure 1.2.) and its interventions (Measure 1.1.). The new service model with extended complexity of new interventions introduced personalised and compound assistance. In the same time it was step by step recognised that the increasing costs of these tailor made help should be based on solid evidence. A key element of the developments related to this recognition was to improve the sophistication and reliability of the monitoring system. As part of the Measure 1.2. a comprehensive proposal was put on the table. The title was: Recommendation on the evaluation and monitoring system of the labour market interventions operated by the Labour Centres. This paper indentified several areas for development as harmonised data collection, IT development for managing the data, the development of human capacities and methodology for regular monitoring exercises.

The document also addresses the importance of the net impact assessment method stating that this approach cannot be a part of the normal monitoring operation. It should remain targeted, occasional and the findings must be used in the possibly widest context.

The developed IT surface became functional gradually during the programme. The above referred study published in October 2006 found that under the HRDOP Measure 1.1. treating large number of people many labour centres used individually developed paper based analytical registration techniques and Excel tables. The centrally developed so called "Frame software" gradually replaced the manual registration but not without complains of technical difficulties, errors and missing functionalities especially in the area of aggregating data and reporting. These facts show the difficulties of this evolving system, in which the conceptual advancement had to be accomplished with the necessary technical (IT) conditions, and the preparation of the staff.

In October 2007 the Managing Authority of the Human Resource Development Programmes launched an open public procurement procedure for the "mid-term" evaluation of the selected measures of the HRDOP. As a result of this
procedure, MEGAKOM Development Consultancy was chosen and contracted in January 2008 to implement the evaluation. The project activities were delivered between February and June, 2008. The evaluation covered 7 specific measures of the OP, among them the Measure 1.1.: Preventing and tackling unemployment. The evaluation used the net impact assessment method with counterfactuals. Although the exercise was probably not perfect from scientific point of view, it was a perfect learning opportunity.

**Selected impact study**

The selected impact study is very similar to that of Section 4.1.

**One impact evaluation among others**

There are currently other on-going impact evaluations that could also have been chosen, but the time constraint does not allow us to wait for their first results. The main goal of this study is the analysis and evaluation of the impact of cohesion policy interventions on the level and quality of employment in Visegrad countries (Czech Republic, Hungary, Poland and Slovakia).

The impact of chosen types of interventions (*bottom-up approach*):

- Support for enterprises,
- Support in the field of education and training (ESF)

The approach will be complemented with a top down macro model logic.

The analysis concerning employment will be carried out and presented in gross and net form. They plan to provide for the deadweight factor existence, displacement, substitution and multiplication effects. An impact analysis is going to be carried out using the counterfactual method.

**Description of the evaluation**

It is not easy to justify the validity of the data after two years. The evaluators faced difficulties in case of accessing reliable data regarding other measures of the OP as well. In case of Measure 1.1. the evaluators reported exceptional fluent data provision by the National Employment Office which resulted in the feasibility of the planned analysis. The other option might have been to work with the results of the decade old study referred above. Both the less probable availability of the involved experts and the relevance of the related information and lessons learned seemed to be problematic.

We can say that there is a significant improvement in data collection and of the understanding how reliable data can feed evaluations, even if a new software and IT provider resulted in some transition deficit. There were also serious commitments to institutionalise the regular monitoring process based on quantified facts. Parallel with the active labour market measures under the same OP priority ongoing systemic conceptual work had been done (Modernisation Handbook of the Hungarian PES). There is no evidence that the professional and institutional links between monitoring and evaluation have been explored. There is also a difference between the monitoring approach of the Ministry of Employment mainly concentrating on the impact on the placement opportunities of the assisted unemployed while the Evaluation Unit of the National Development Agency has more integrated macro-view putting more emphasis on net national impact.
**Reasons why the impact study was launched**

There were three main reasons for selecting this study. The first is the different assessment of the effectiveness of the more or less uniform intervention in the same time, even if the innovative nature of the tailor-made services made this uniformity somewhat diversified. The second reason was that there had been an increasing scepticism regarding the effectiveness of the labour market interventions. It can still use new evidences. The traditional tools of the labour market institutions seemed simply ineffective in helping the labour market reintegration of various groups of the registered unemployed. Measure 1.1.1 of HRDOP was the first decentralised programme with the objective of supporting the return to the labour market of these people through the use of innovative solutions (services and support) financed from the European Social Fund. Furthermore, the measure has also complemented the limited resources available at the labour centres for such purposes from domestic sources, and it has successfully introduced and disseminated the various practices currently constituting the foundations of the EU’s labour market policy and tools. The third is the availability of the needed information. A related recent complain of a civil servant of the ministry: there is serious evaluation burden on the Public Employment service due to the fact that the data and indicators of the impact of other ESF interventions in most cases are less measurable so less proper for analysis.

Even if the availability of the data seems to have been a key explanation, the impact evaluation put great emphases on the core issue of 'the success of innovative services and institutional arrangements such as decentralisation ', and that this issue was high on the agenda of the policy-makers when the evaluation was launched. According to a ministry official, although the evaluation was initiated and managed by the MA, later it became more important for the Ministry arguing in favour of a the ESF share for the follow up of this type of operation in the period 2007-2013.

**Impact study and the agenda of decision/policy-makers**

The analyses was organised in the right moment. When the public tender was launched the preparation of a similar measure for (2007-2013) had already been started. It was duly expected that the findings could be well used in designing the evolving ALP services.

The exercise meant novelty as it was not part of the usual practice. The existing methodological knowledge was accessible by the public authorities, because the high quality labour economist community often referred to the need of the evidence base of the Hungarian practice of the ALPs, however the ministry responsible for employment felt easier to faithfully follow the EU guidelines, and simply monitor the outcomes of the measures instead of taking the risk of not finding evidence of the existing practice. The evaluation was initiated by the MA, but on the contrary according to the interviews and the available documents the Ministry of Employment was motivated in using the (supportive) finding.

**Stakeholders involved in the impact study if there were any**

- The members of the Monitoring Committee
- The data providers of the National Employment Office
- The participants of the closing session of the project

**Overview of major findings**

On the control day (May 2008) in case of the treatment group 58,23%, while in case of the control group 47,56% had registered employment. It means
that the treatment group had better employment situation compared to the control group. The broad overall conclusion that the impact on the probability to find a job is +11%

The findings were not too detailed. The evaluators said that the technicalities take away the most of the available time, and short time remained for the analysis.

**Selected finding(s)**

This finding was simple, and confirmed the expected positive impact on of the largest interventions of the OP. The conclusion is about the success of the new service model with personalised and compound assistance. This finding was crucial for the policy makers of the Ministry of Employment who regularly had to meet with the scepticism of the MA.

**Dissemination of the findings**

According to the available documents and the interviews the quality of dissemination of the evaluation was poor. The lack of institutionalisation and experience of how the evaluation results can be translated into actions hampers the dissemination process. Nevertheless the easy access to the study might good opportunity for wide range of stakeholders to get acquainted with it.

The final report introduced the methodology and some details of the findings more or less in detail.

The documentation of the evaluation is available on the Internet ([https://hepih.nfu.hu/page.php?PageID=238&OpenClose=Időközi%20értékelés%202007/](https://hepih.nfu.hu/page.php?PageID=238&OpenClose=Időközi%20értékelés%202007/)). The final report is the most informative part of these publicly available documents.

The main dissemination event was the closing session of the evaluation project. The consultants presented their results, and they expressed their wish to think about the follow up of the evaluation. There is no indication or explicit text indicating measures derived from the whole complex ESF evaluation report.

The workshop took place on 26.08.08. The Summary Report and the presentations of the whole evaluation exercise of the 7 analysed measures gave the basis for the discussion. The challenges were clear: how to address those who should use the findings of the study. The generalities came from the nature and complexity of the evaluated measures.

The presentation was relatively shy presenting the findings of the impact study compared to the extent of the final report. The impact study results had no real chance to compete with other more detailed, colourfully presented evaluation results (e.g. CATI survey, questionnaire)

The composition of the participants did not allow to deliver messages for the decision makers (for example one of the key institutions delegated a brand new colleague, for whom the subject of the evaluation was the same new than the evaluation itself).

**Learning opportunities/Use of findings**

The Ministry of Employment used the selected finding for defending the financial share and further developing the professional content and extension of the studied measure.
The Ministry of Employment used several sources for knowledge accumulation. Qualitative studies, reports, among them the selected impact assessment as well, but they did not administer the practiced learning opportunity(ies).

The Monitoring Committee formally put on the Agenda the Summary report, but I could not get access to the minutes of the meeting that would prove that the discussion actually happened.

According to one of the interviewed ministry officials the evaluation results were internally analysed and were taken into account in planning the 2007-2008 Action-plan. 4 documents have been provided by the ministry to demonstrate the use of the findings. From these documents, only one refers directly to the evaluation results. The proposed text for the closing report of the operational programme has detailed reference to the methodology of the impact evaluation. However the quoted findings were rather referring to the whole evaluation. E.g.: according to the analysed indicator the measure 1.1 is one of the most successful interventions because of the involved participants exceeded the plan. The proposed changes targeted rather structural issues (stronger monitoring role of the ministry, better linkages between the employment and social institutions and services). The overall positive results let the launched innovations further extending to system level.

**Information sources:**

Ms. Lilla Jutkus, monitoring assistant to the director general of the Managing Authority

Mr Béla Kézy, Head of the Megakom Consulting, the company which delivered the impact analysis in 2008

Mr. Adam Baric, former project manager of evaluations of the MA responsible for HRD programmes

Mr. Kerekes Ildiko (HBF – subcontractor of Megakom)

Ms. Reanáta Tóth – head of division responsible for planning
Appendix 7 – Learning from impact evaluation: Italian case study

During the 2000-2006 programming cycle, Italy adopted a common information system for monitoring implementation of the structural funds at the national level (called Monit 2000 and later Monitweb). The information systems progressively developed in the Central-Northern regions, while significant delays were experienced in the South. By 2003, all the Northern regions had started their own placement surveys, while for Objective 1 regions the Ministry decided to delegate monitoring and data collection to the ESF national evaluation unit, situated inside ISFOL, the research institute on human resources development which supports the Italian Ministry of Welfare. Therefore, the country experienced two diverging trends in the data collection on employment outcomes of intervention: decentralisation at the regional level for Obj. 3 regions and centralisation on Obj. 1. In Obj.3 each region (under the coordination of Isfol) realised its own survey while in Obj.1 the Ministry of Labour decided to give charge to Isfol to realise a centralised survey.

In 2005, the Placement Working Group in the ESF National Evaluation Unit was extended to cover both Obj. 1 and Obj. 3 regions, in order to define a common framework for retrospective impact studies on ESF recipients. The Group is composed of representatives from all regions and autonomous provinces, central administrations and research institutes. It has progressively developed a series of tools, such as:

- a standard application form for interventions,
- a follow-up questionnaire for surveying the employment status of participants,
- a questionnaire for evaluating employment outcomes of continuing education (hereafter ‘placement surveys’),
- a common methodological approach to estimating employment impacts.

This activity resulted in the production of:

- For the cycle 2000-2006: standard application forms; questionnaires for evaluating the employment outcomes; Guidelines for the evaluation of ESF employment effects (2001).

80 Only exception being Sardinia and a few attempts in Sicily, but limited at the local level.
81 In the 2000-2006 ESF programming period Isfol was also the independent evaluator of CFS Obj.3.
82 The autonomous provinces of Bolzano and Trento.
83 The investigations on employment outcomes of interventions cofinanced by the ESF, hereafter “surveys placement”, are intended by ISFOL as “retrospective surveys, by sample or census, built specifically for the evaluation of effectiveness”. The follow-up surveys, instead, are merely forms or questionnaires administered to participants directly from training agencies shortly after the end of the course, as required by the national legislation on accreditation. In the past programming cycle, however, follow-up data has been improperly exploited by some regions (case in point being Emilia Romagna, as we shall see later on). Conference Proceedings Firenze 22 aprile 2008.
84 “Linee guida per la valutazione degli effetti occupazionali del Fondo sociale europeo 2000-2006” ISFOL
DG EMPL - Study on the Return on ESF Investment in Human Capital

• For the cycle 2007-2013: extended application forms\textsuperscript{85}; questionnaires for evaluating employment outcomes; discussion on different approaches to estimate employment impact.

ISFOL encountered a few obstacles in retrieving personal data records on the project beneficiaries, mainly because of:

• failure of project holders to transmit detailed data set to the regions
• lack of an adequate IT system (especially in Obj. 1 regions)

However difficult, it has to be underlined that in the final phase of the ESF evaluation process, ISFOL succeeded in the creation of a common “Placement database” which represents a rare example of integration among administrative archives and retrospective investigations carried out at the regional level\textsuperscript{86}.

**Impact evaluations**

Two reports have been identified which are relevant in terms of impact evaluations over the 2000-2006 period: the mid-term and final evaluations of Objective 3 CSF and the report “training and employment in Southern Italy”\textsuperscript{87}. Both these evaluations include estimates of the impacts on unemployed participants in terms of finding jobs. However, the two evaluations differ for their geographical scope and methodological approach. As we shall see, ISFOL only partially succeeded in developing a fully fledged impact analysis.

In the southern Objective 1 regions, the survey on placement returns was carried out on interventions ended between 2003 and 2004. In contrast with what was foreseen in the inception phase, an internal control group could not be identified, due to the gaps in regional data archives. By the same token, ISFOL could not develop an external control group, due to delays in the data collection and the changes which intervened in the labour work force surveys conducted by the Italian statistics institute (ISTAT).

Due to the bad state of the regional administrative archives on beneficiaries, the data set lacked information on the control group defined in the project, so Isfol abandoned the idea to realise impact evaluation. Because of the limited insight in terms of impact analysis, the Obj. 1 report is not retained for further investigation within the framework of this document.

**The impact evaluation under study**

**Evaluation process**

As regards the Objective 3 regions\textsuperscript{88}, the interim\textsuperscript{89} and final evaluations\textsuperscript{90} were mandated to the National ESF Evaluation Unit (ISFOL) by the Italian

\textsuperscript{85} Strumenti di rilevazione all’ingresso degli interventi.

\textsuperscript{86} Impact evaluation was not the main aim of the project financed by the Ministry of Labour. The principal aim was to measure employment outcomes of interventions. Impact analysis was an additional and experimental objective.

\textsuperscript{87} Formazione e lavoro nel Mezzogiorno - Indagine sugli esiti occupazionali degli interventi finalizzati all’occupabilità cofinanziati dal Fondo Sociale Europeo 2000-2006 nelle Regioni Obiettivo 1 – Rapporto finale, ISFOL, Febbraio 2008

\textsuperscript{88} Piemonte, Valle d’Aosta, Lombardia, Emilia Romagna, Toscana, Abruzzo, Autonomo Province of Trento and Bolzano, Lazio, Umbria, Marche, Veneto, Liguria, Friuli Venezia Giulia and Emilia Romagna.
Ministry of Labour and Social Policies and the European Commission. The final report, published as a book in July 2007, is based on interviews carried out between 2003 and 2005 by the regional authorities. Because of its structure, this report has also been defined as the central evaluation of a decentralised system.

**Placement surveys**

The impact analysis focuses on the placement results of training. Therefore, only measures to improve employability were considered, i.e. projects in formal education and lifelong learning, not on-the-job training. The updated intermediate evaluation of CSF Obj.3 used a sample of 35,174 individuals to analyze the employment results of intervention (interventions concluded in 2002).

For the impact evaluation, ISFOL considered more years (interventions concluded in 2001, 2002 and 2003) than in the employment results analysis (only interventions concluded in 2002). The final sample for the impacts analysis was thus composed of 47,454 individuals.

The effectiveness of training was first measured in terms of changes in employment status: 64% of participants interviewed via the regional placement surveys declared to be employed after one year. The probability of entering a formal education path was also considered. The total success rate of training activities, defined as the integration of recipients in the job market and/or in the education system, thus rises up to 76%. From the gender perspective, the most significant finding is that the low income job category is represented ten percentage points more by women (36%) than men.

**Impact analysis**

An external control group of 8,049 non-trained people was based on the ISTAT micro-data on the labour force over 2006-2004. The sample considered all the people concluding training programs, except for those individuals who find a job before completing the training programme.

---

91 “The evaluation of the European Social Fund in Italy: a central evaluation of a decentralised system” by Bulgarelli and Vergani at the 4th EES Conference, Lausanne 12-14 October 2000
92 The analysis was thus limited to the following formal off-the-job types of training: post graduate training, scholarships, training for re-entering employment, training during compulsory education, training after compulsory education and high school diploma, professional and technical training, apprenticeship and lifelong training. On the job training instead indicates all training offered to employees on the company’s premises.
93 This is because ISFOL had already published in other previous reports employment results analysis omitted in the update of intermediate evaluation of CSF Obj.3
94 Individuals in the control group were selected if unemployed at time t and had not undergone any sort of training until t+12. The relatively small size of the control group could admittedly represent a weakness in the impact analysis carried out by ISFOL.
95 Regions Lazio and Abruzzo were excluded from the dataset since their surveys could not be carried out within the deadlines for the final evaluation report. The 2001 survey was missing for: Valle d’Aosta, Lombardia, Liguria, Umbria, Marche, Autonomous Province of Trento, while the year 2003 was missing for Emilia Romagna and Toscana. The
The comparison on the employment status of the two groups confirms that training generally enhances the chances of finding a job. No evidence was found questioning the credibility of these results, probably due to the strong reputation of ISFOL and the solid evaluation methodology built in close coordination with the Managing Authorities.

A comparison of the two groups shows that participants have a 32% higher probability to be employed after 12 months (see next table, first line). The authors have also applied three different econometric approaches in order to deal with the selection bias (see next table, bottom lines)\(^96\). The most valid approach (endogenous switching model) shows that the impact is 26% and not 32%.

<table>
<thead>
<tr>
<th>Increased probability to be employed after 12 months</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple comparison</td>
<td>32</td>
</tr>
<tr>
<td><strong>Econometric approaches</strong></td>
<td></td>
</tr>
<tr>
<td>Probit</td>
<td>27</td>
</tr>
<tr>
<td>Biprobit</td>
<td>31</td>
</tr>
<tr>
<td>Endogenous switching model</td>
<td>26</td>
</tr>
</tbody>
</table>

Interviews at ISFOL revealed that, from the standpoint of policy-makers, the most relevant new pieces of knowledge were:

- Training has a positive impact on participants;
- The target groups identified benefit from interventions;
- Impact is stronger for women;
- Training is particularly effective for the most vulnerable categories, e.g. persons at risk of losing job, long term unemployed, women in general and particularly women from 35 to 45 years old, etc.

In addition, the study showed that the impact decreases with age, increases with the level of schooling, and is positively correlated to the local labour market conditions\(^97\).

If the findings were fairly expected and thus do not constitute a real breakthrough, they were extremely valued by all the interviewees\(^98\) as the first piece of evidence confirming the positive impact of training. However, trainings frequently include an internship period and often the project holder has close ties with local actors in the labour market. These very important explanatory factors were not included in the analysis and could admittedly represent a weakness.

**Dissemination**

The ESF national evaluation unit inside ISFOL played a leading role, through the Placement Working Group, in disseminating the findings of the study. IS-
FOL communicated on the methodological guidelines and findings at three levels: regional, national, and European.

At national level, the findings were tabled by the Monitoring Committee of Community Support Framework, but no significant evidence of these discussions could be retrieved. For this reason, the effectiveness of the dissemination processes is considered quite poor and the use of the findings at national level is not retained for further investigation within the framework of this document.

**Use of the findings at regional level**

*Mix of national and regional studies*

The impact evaluation at national level pooled regional data sets originating from harmonised placement surveys. Locally, these data sets were also used for carrying out regional evaluations.

ISFOL’s technical and methodological assistance played a key role in promoting such regional impact evaluations. ISFOL also acted as a platform for the exchange of best regional practices, which allowed for further spontaneous learning in terms of methodology.

*A good potential for learning*

According to the interviews at ISFOL, in Italy there had been a strong debate on the effectiveness of training, especially training courses, at the time when the reports were published. Many advocated that on-the-job training is preferable. Therefore, the robust finding that other forms of training improve employability was a potentially important input in the debates.

Generally in Italy there is a strong concentration of ESF interventions on particular target groups: persons at risk of losing job, long term unemployed, women from 35 to 45 years old, etc. Therefore, the finding that training benefited these groups particularly was a potentially important argument for legitimising and strengthening this option, in spite of the fact that such publics are more difficult to attract and to keep in training activities.

99 Among others, representatives of the National ESF Evaluation Unit inside ISFOL held meetings in Emilia Romagna (2004), Trentino, Toscany.

100 For example, ISFOL organised in 2001 a conference on "Training and employment: evaluating the employment impacts in the Italian Regions". The meeting was held at the National Council for Economics and Labour and saw the participation of Regional authorities, the European Commission, Ministry of Labour, various researchers and experts in the area. Particularly noteworthy the presentation "Measuring the impact of training measures on employment" by Berliri e Pappalardo.

101 Such as the presentation "Employment Impact of ESF Training Interventions on Particular Target Groups" Berliri at the Fifth European Conference on Evaluation of the Structural Funds (Budapest 2003).

102 At the national level, a common Monitoring Committee is set up by the Community Support Framework to cover all the Structural Funds.

103 It should also be noted that the great majority of the interviews, both at regional and central level, qualified the added value of ESF impact studies mostly in terms of methodological learning and spreading the evaluation culture within Italian public administration. For further details on the process of methodological learning, refer to Appendix 3.
Uneven interest in the findings

Managing authorities at regional level have paid uneven attention to impact evaluations. The involvement of Italian regions in ESF monitoring and evaluation system has varied greatly, depending on their understanding of these activities as a mere formal obligation or an actual learning opportunity.

Emilia Romagna was among the few forerunning regions that were strongly involved in the design, implementation and use of impact evaluations. This is why the remaining of this document focuses on it.

The case of Emilia Romagna

A case of good practice

The region has the following characteristics:

- Good monitoring system and availability of good professionals (both internal and external to the administration) in the field of evaluation;
- Traditionally a forerunner in the process of data collection and statistical treatment;
- Long standing commitment to lifelong learning and adult training;
- Strong link between technical and administrative officers in charge of data collection and treatment with the regional policy maker, who generally show a good level of attention to the results of evaluations.\(^{104}\)
- Emilia Romagna was thus chosen as a best practice, but the observed good methodological practices cannot be applied to all ESF funded activities in the country.

Several conclusive evaluations

In the first years of the Millennium, Emilia Romagna, through its regional employment agency\(^{105}\), decided to mandate an external consulting firm\(^{106}\) with the task of producing an annual evaluation report on the effectiveness of educational training in terms of job finding. These evaluations were carried out to answer to a legal obligation and do not stem from a specific decision linked to the national political agenda\(^ {107}\). Impact evaluation therefore resulted from a technical supply rather than a political demand. Three reports were published on the employment impacts of training courses delivered throughout 2000-2002\(^ {108}\). Trainees were phone-interviewed twelve months after the end of their participation in the courses. Overall, around 45,000 participants were interviewed on approximately 1,600 courses, totalling a 10 million hours of training activities.

\(^{104}\) As confirmed by interviews at ISFOL and by the intervention of the Director General of the Regional Labour and Training department at the roundtable organised by ISFOL in 2001.

\(^{105}\) AERL (Agenzia Emilia Romagna Lavoro)

\(^{106}\) The assignment was to be carried out under the scientific supervision of CAPP, the Public Policies Analysis Centre at the University of Modena and Reggio Emilia.

\(^{107}\) The impact evaluations were required in the ToR issued by the regional Managing Authority, with explicit reference to the EC regulation. However, according to the external evaluator ‘the list of services required reminded rather synthetic’.

\(^{108}\) To these 3 reports should be added an evaluation on the effectiveness of training activities on handicapped recipients.
The second evaluation report, published in 2004 on projects ended in 2001, included a quasi-experimental analysis, which, at least in Italy, had never been carried out on a regional scale before\textsuperscript{109}. For two years, the external consultants carried out an impact evaluation with a control group drawn from data collected by ISTAT. The studies took into account the initial characteristics of participants and non-participants, while the selection bias was controlled through a propensity score matching method\textsuperscript{110}.

The analysis revealed that interventions designed to favour reintegration of the labour market trigger a higher impact in general and for women in particular, whereas the continuous higher education interventions seem to favour male recipients. For both types of interventions, employment impact turned out to be stronger for women\textsuperscript{111}.

The findings of the study were as follows:

<table>
<thead>
<tr>
<th></th>
<th>2001</th>
<th>2002</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Men</td>
</tr>
<tr>
<td>Increased probability to be employed after 12 months (%)</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>7</td>
</tr>
</tbody>
</table>

These regional findings are much less positive than the national ones which were published two years later. This is due to the different methodologies used in the Emilia Romagna region and the national research. However, the analysis by gender ends into the same lesson: impact is higher for women.

The gender differential suggests that vocational training policies can constitute an effective form of intervention in order to dampen gendered differences in employability in the regional context of Emilia Romagna.

**A leading practice which was discontinued**

The placement surveys and the impact evaluations carried out in Emilia Romagna served as "pilot tests" for the development of the general guidelines within the Placement working group. The external consultants and the Programming and Evaluation Department inside the region worked in close cooperation with ISFOL to develop the methodology, sampling plan and questionnaire.

Although the regional administration has a long tradition in evaluation and analysis, not only for ESF, the local policy-makers are said\textsuperscript{112} to have taken this initiative to respond to a pressure from the national and European levels.


\textsuperscript{110} The methodology adopted compared the transition matrices of employment status one year after completion of courses with similar matrices of the overall regional population drawn from the quarterly labour force survey.

\textsuperscript{111} The 2000 evaluation revealed that 81% of the interviewed recipients of interventions aimed at re-entering the job market resulted employed after 12 months, while only 77% of the recipients of further higher education projects were employed. Still, the same results were consistently higher for women, which were employed at a rate of 83% and 74%.

\textsuperscript{112} As gathered from interviews with the Regional Managing Authority and the independent external evaluators.
and to have understood their evaluations as administrative and methodological exercises. Their level of ownership thus appears rather superficial and no clear cut evidence could be found to support the contrary.

Half way through the OP implementation, Emilia Romagna decided to change the survey methodology, because the cost and the sustainability of the external technical assistance was deemed excessive. Also, interviews\textsuperscript{113} revealed that the interest of annual surveys was perceived as limited:

- results confirmed the high level of employment in the region,
- they showed no significant change on an annual basis, and thus
- the costs were not justified.

The region opted instead for use of the follow up data coming from the training agencies\textsuperscript{114}, although ISFOL, in their role of technical assistance, advised against, due to the inherent conflict of interest. A regional directive\textsuperscript{115} was issued detailing the questionnaire to be used by the training centres and defining the role of the region in terms of data verification and treatment. The present monitoring system, as drawn in the 2007-13 regional evaluation plan\textsuperscript{116}, provides for annual control surveys by the region - a small sample and light questionnaire, to check the reliability of data while keeping costs low. Thus, the regional monitoring system now presents a two layered structure:

- Knowing the labour market status of past participants through an online database of certified training centres;
- Checking these data through an independent survey on a small sample of participants, and controlling biases through this survey.

The 2007-13 ESF regional OP underlines the Managing Authority’s commitment to carry out placement sample surveys, although none has taken place yet. The information gathered will feed into the results indicators for the specific objectives on active labour market policies for social inclusion and women’s integration in the labour market.

Lessons that have not been used

Potentially interesting lessons

According to the regional representatives interviewed, the use of a counterfactual allowed to confirm the positive impact of training in a context where doubts were raised because of the high employment level\textsuperscript{117}.

Because the impact evaluations were not mandated on an ad hoc basis, but rather to abide to an administrative requirement, they addressed broad questions which were not specifically tailored to the need of policy makers at the time.

\textsuperscript{113} Ibid.

\textsuperscript{114} In Italy there is a national certification system for training centres, whereby they are required to carry out annual placement surveys on their beneficiaries.

\textsuperscript{115} Determina 631 2008 del Responsabile Del Servizio Formazione Professionale Prot. n. (SSF/08/25623)

\textsuperscript{116} PIANO DI VALUTAZIONE UNITARIO 2007 - 2013

\textsuperscript{117} Over the period concerned, the region was characterised by high and increasing employment rates. The unemployment rate among the active population floated slightly above 3% in 2002 and 2003.
Although the impact evaluation produced potentially useful findings, the regional Monitoring Committee and the public authorities focused their attention on the placement surveys only, i.e. employment status of former participants instead of difference with the control group.

The reports also proved that training is much more effective on women in terms of employability. In absolute terms, gross impacts show that men tend to be more successful in finding a job. However, by comparing the population of treated to the counterfactual, the differential is more significant for women. In general, for the most disadvantaged target groups, training proved to be an important means for social and professional inclusion.

The main lessons as reported by interviewees within the regional Managing Authority would thus be that “training plays a key role in finding a job for disadvantaged individuals” and that “the programme should not reduce their focus on the weakest segments of the population”.

**Channelling lessons to policy-makers**

Regional representatives and external consultants actively participated to the works of the National Workgroup on Placement, such as the round table on "Public policy evaluation and decision making". The regional general director of the Culture, Training and Labour department attended the roundtable together with other regional authorities, representatives of the national ministries and social partners.

Results of the placement surveys were reportedly presented at the Monitoring Committee meetings at the time\(^{118}\). Still, regional policy-makers had a traditional interest in evidence. As a consequence, the preconditions for use were excellent.

The only opportunity for learning was identified by the interviewees in the drafting of the following regional OP.

**No short-term use**

No evidence could be found of an immediate and direct use of the lesson learned in Emilia Romagna.

On the contrary, it is reported that regional policy-makers were more interested in issues such as quality of jobs, and consistency and duration of job search.

**Delayed and indirect use**

At the date of writing this document, the regional policy-makers seem to have integrated one of the main lessons arising from the impact evaluation, i.e. that women and the weaker target groups greatly benefit from training. According to the interviewees, this is proven by the strong focus which was preserved in the regional ESF programme on the social inclusion axis and on projects addressing gender equality, despite the significant decrease in financial resources for 2007-13\(^{119}\). No other example of a specific decision quoting this lesson could be retrieved.

The administrative staff interviewed within the Managing Authority also stressed the link between the impact of training in terms of finding a job and

\(^{118}\) The evaluation team could not get a hold of the minutes of the Monitoring Committee meetings.

\(^{119}\) Reportedly, the present cycle saw a reduction of about 40% in national and regional funding.
the criteria for accreditation of training centres. This already existed in the
previous cycle, but since 2008 the Managing Authority introduced a minimum
threshold for the effectiveness of training in terms of finding a job\textsuperscript{120}.

However, this learning process is not just the consequence of the impact
evaluations reviewed. In fact, the lesson was reinforced by the EQUAL evalua-
tions, which concluded that the impact of training will be maximised by the
existence of a network linking the actors on the field and guiding recipients
from the educational activities to the labour market. The EQUAL evaluations
were published from 2005 to 2007\textsuperscript{121}, while the placement impact surveys
stopped in 2003. Because the EQUAL evaluations arrived at the beginning of
the design phase of the following programme, their findings had a much
stronger impact on the internal discussions. As an interviewee concluded: “Po-
litical interest tends to be higher when evaluations can influence the new
programming exercise”.

In another area the interest for the impact evaluation could also be revived
soon. In the current context of crisis which has somewhat deteriorated eco-
nomical situation, regional officers will be more focused on the capability to
keep and reintegrate people in the work force\textsuperscript{122}. In the present period, Emilia
Romagna experienced a reorientation towards continuous and lifelong training
activities targeting adults at risk of losing their jobs. This shift will be reflected
in their demands for evaluation, from assessing the advancement in human
capital qualifications to the improvements in adaptability and maintaining
people in the labour market.

**Information sources**

**Interviews**

M Paolo Severati, National ESF Evaluation Unit ISFOL (08.02.10)

Ms Cristina Berliri, State Treasury and previous National ESF Evaluation Unit
(01.03.10)

Ms Paola Stocco, National ESF Evaluation Unit ISFOL (08.02.10)

M Enrico Toti, National ESF Evaluation Unit ISFOL(08.02.10)

Ms Francesca Bergamini, Director at the Programming and Evaluation de-
partment, Emilia Romagna (23.03.10)

Ms Paola Zaniboni, Continuous and professional training, in charge of the
area: certification and accreditation of professional training centres, Emilia
Romagna (23.03.10)

M Mario Demurtas, Director of Poleis, (17.03.10)

\textsuperscript{120} With the new regional regulation, in order for each training centre to maintain their
accreditation, at least 70% of recipients must be employed within 12 months and at
least 50% must have found a job coherent with the training received. See: Determina
631 del 25/01/2008

\textsuperscript{121} La Valutazione di Impatto dei Progetti Equal I Fase Sulle Politiche Sociali in Emilia-
Romagna “ Poleis, Maggio 2005

“Equal II Fase l’azione 2 in Emilia-Romagna” CRAS, Novembre 2007

\textsuperscript{122} As was confirmed by interviews at the regional level Bergamini, Director at the Pro-
gramming and Evaluation department and Zaniboni, Continuous and professional
training, in charge of the area: certification and accreditation of professional training
centres (23.03.10) and at the national level: Severati ISFOL (8\textsuperscript{th} February 2010)
Documents

National level


L’analisi degli esiti occupazionali del FSE nel periodo di programmazione 2000-2006 e le prospettive per il nuovo settennio 2007-2013, Paolo Severati Firenze 22 Aprile 2008

Formazione e lavoro nel Mezzogiorno - Indagine sugli esiti occupazionali degli interventi finalizzati all’occupabilità cofinanziati dal Fondo Sociale Europeo 2000-2006 nelle Regioni Obiettivo 1 - Rapporto finale, ISFOL Febbraio 2008

Indagini regionali FSE 2000-2006 Note sulla metodologia da utilizzare nella rilevazione degli esiti occupazionali, Gruppo di lavoro sul placement ISFOL

Rapporto di monitoraggio delle politiche occupazionali e del lavoro, Ministero del Lavoro, della Salute e delle Politiche Sociali settembre 2008


L’analisi degli esiti occupazionali del FSE nel periodo di programmazione 2000-2006 e le prospettive per il prossimo, Paolo Severati ISFOL 22 aprile 2008, Grand Hotel Baglioni, Firenze, “Gli esiti occupazionali delle attività formative finanziate dal FSE”


Linee guida per la Valutazione degli effetti occupazionali del Fondo sociale europeo 2000-2006, ISFOL - SL (SV) - 1/03

The Evaluation of the European Social Fund in Italy: a central evaluation of a decentralised system, Bulgarelli and Vergani ISFOL 4th EES Conference, Lausanne 12-14 Oct 2000


Osservatorio "Formazione Orientamento Occupazione Nuove Tecnologie Professionalità” Anno XXII n 6 ISFOL Novembre-Dicembre 2001

Regional level


Rapporto Finale Di Valutazione Di Efficacia Delle Azioni Rivolte Alle Persone (Anno 2000), POLEIS Marzo 2003


Questionario Per La Rilevazione Degli Esiti Occupazionali Per Gli Enti Di Formazione Accreditati, POLEIS Luglio 2007

Appendix 8 – Learning from impact evaluation: Polish case study

Description of the selected impact study

In the context of ESF impact analysis carried out in the framework of the previous financial perspective “The Study on the ultimate beneficiaries of the Sectoral Operational Programme Human Resources Development 2004-2006”, presenting the data on the employed and the unemployed was identified and further constitutes a subject of this case study, primarily as it provides for in-depth insights in terms of impact analysis and learning opportunity. The study further investigated in this document was commissioned in 2006 as one of the first attempts of the Polish authorities to evaluate the effects and the impact of ESF intervention in Poland being a continuation of another impact study identified but not retained for further analysis. The study was included in the Evaluation Plan for SOP HRD 2004-2006, following the requirements of the EC “Guidelines for system of monitoring and evaluation of EFS assistance in the period 2000-2006”. The results of this and other evaluations foreseen in the Evaluation Plan were also to be used for the 2007-2013 programming period as well as for other strategic documents such as the National Action Plan for Employment or the National Action Plan for Social Inclusion.

In Poland the system has been devised to integrate evaluation processes during the period 2007-2013 for National Development Plan 2004-2006 and for the National Strategic Reference Framework 2007-2013 (NSRF). In April 2004 the National Evaluation Unit (NEU) was established as the main element of the system of implementation of the evaluation process. This function is performed by the Department of Structural Policy Coordination of the Ministry of Regional Development. NEU performs the entrusted tasks in cooperation with the European Commission and the NSRF Coordination Institution, as well as the Institution Coordinating the Regional Operational Programmes (ROP), and also the institutions at the level of management and implementation of all operational programmes. In the period of the years 2007-2013 evaluation studies are conducted at the level of NSRF/NDP and cover the horizontal and transversal issues concerning more than one operational programme. Moreover, evaluation studies are also commissioned in the framework of all operational programmes. The evaluation work conducted at both levels is coordinated by the National Evaluation Unit. Furthermore, the Unit is responsible for the development of evaluation standards, so as to improve the

123 the summarized results of the analysis are presented in two Final Reports, one presenting the data on the employed and the second one on the unemployed participants, both published in May 2008, (Raport Końcowy z ‘Badania beneficjentów ostatecznych Sektorowego Programu Operacyjnego Rozwój Zasobów Ludzkich 2004-2006 przedstawiający dane dotyczące osób pracujących’, Raport Końcowy z ‘Badania beneficjentów ostatecznych Sektorowego Programu Operacyjnego Rozwój Zasobów Ludzkich 2004-2006 przedstawiający dane dotyczące osób pracujących’, ABR Opinia, 2008.

124 The previous impact analysis carried out in 2005 as a pilot study, covered a lower number of participants surveyed (1145) with no control group identified („Badanie beneficjentów ostatecznych Pomocy udzielonej w ramach Sektorowego programu operacyjnego Rozwój zasobów ludzkich 2004-2006”, Anna Chrościcka, 2005).

125 The respective details are determined by Guidelines No 6 concerning the evaluation of operational programmes approved by the Minister of Regional Development on 30 May 2007.
quality of the conducted analytical studies. The Managing Institutions are accountable for the evaluation of the respective Operational Programmes. The evaluation plans prepared for SOP HRD and later for OP HC defined the framework of the evaluation process in Poland covering the ESF support until 2015.

The impact study analyzed, as foreseen in the SOP HRD Evaluation Plan for 2004-2006, was to provide the quarterly monitoring data which subsequently were used in fulfilling the monitoring function of the ESF Unit dedicated - at the time when the study was designed - to monitoring, programming and evaluation. The data collected during the quarterly surveys of final beneficiaries on ESF assistance were subsequently fed into the reporting on ESF implementation, thus the analysis had to be planned and implemented in such a way that the needs of the information suppliers could have been matched and the respective roles and tasks of the Unit met. The credibility of the study was confirmed by the persons in charge of the impact study in the Managing Authority. The findings of the study did not contradict with the main findings of the previous impact study conducted on a smaller scale.

The representatives of the ESF Management Department in the Ministry of Regional Development acknowledge the unprecedented scale of the impact analysis in terms of the number of the ultimate beneficiaries surveyed. The analysis implemented during the period 2006-2008 included the samples of the ultimate beneficiaries of ESF assistance under nine Actions of the SOP HRD subjected to five successive waves of questionnaire surveys and the control group samples subjected to two waves of surveys. The beneficiaries surveyed benefited from ESF support under the following Actions under Priority 1 (Active labour market and professional and social inclusion policy) and 2 (Development of knowledge-based society) of OP HRD 2004-2006:

- 1.1 Development and modernisation of labour market instruments and institutions (employed)
- 1.2 Perspectives for youth (unemployed)
- 1.3 Counteracting and combating long-term unemployment (unemployed)
- 1.4 Vocational and social integration of the disabled (unemployed)
- 1.5 Promotion of active social policy in support of high risk groups (unemployed)
- 1.6 Vocational integration and re-integration of women (unemployed, employed)
- 2.2 Improving education quality and relevance to the labour market needs (employed)
- 2.3 Development of personnel of modern economy (employed)
- 2.4 Strengthening administrative capacity (employed)

The majority of the ultimate beneficiaries (for the employed participants 3464 out of 3646 and all 6841 unemployed) were surveyed using computer-assisted telephone interviews in a period of six to nine months after the completion of their participation in projects. The samples were selected through a layered randomised process (layers referred to action, gender, age, region)

---

126 The study was carried out in 5 sessions of the surveys in 2007. In February and March 2007, the study was conducted with beneficiaries after their participation in projects in the second quarter of 2006; in April and May of 2007 with the ultimate beneficiaries who have participated in projects in the third quarter of 2006 etc. In December 2007 the last session was carried out for with the ultimate beneficiaries who have completed participation in projects in the second quarter of 2007.
from the lists of ultimate beneficiaries whose details were submitted by implementing institutions. For Action 1.2 and 1.3 addressed to the unemployed the control samples of participants were identified which reflected the structure of a sample of final beneficiaries covered in the fourth and fifth session of survey127.

The main objective of the study was to assess the impact of SOP HRD projects and determine the effectiveness of the projects. The results of the analysis were to improve the implementation of individual Activities. More specifically the study was to seek the opinions of the final beneficiaries on the quality of projects in which they participated, determine the extent to which beneficiaries have used qualifications acquired or increased as a result of participation in projects, estimate to what extent the improvement of employment situation of the beneficiaries was the result of the support obtained, determine whether ESF contributed to a change of the unemployed situation on the labour market or provide learning about the potential for increasing the usefulness of aid. The study focussed on the issues highly interesting in the context of the ESF support implementation and resulting from Poland’s obligations related to benefiting from the EU assistance128. In line with the ToR of the study, the data collected were to be used for detailed evaluation of the SOP HRD and to contribute to a better allocation of resources between different groups requiring support in future programming periods.

The results provided by the impact study were learnt through application of the credible evaluation method, with the evaluation techniques applied subject to adjustments over the study implementation period as necessary (eg. postal questionnaires used to collect the data were abandoned given the low response rate from beneficiaries; consequently the CATI was solely used as the method of gathering responses for the ultimate beneficiaries). The extended evaluation method has been adopted for the similar study of ESF impact on ultimate beneficiaries under OP HC 2007-2013 commissioned by the OP HC Managing Authority in the first quarter of 2010129.

The study did not provide for clear recommendations to be followed but its strong cognitive character has been confirmed. The interviews conducted with the Managing Authority representatives revealed that the results of the study did allow for valuable observations of tendencies and factors which could have affected the effectiveness of ESF support and which subsequently led to the decisions on conducting the in-depth qualitative analyses of some selected results. The study was designed as an on-going analysis hence the findings from the consecutive waves of surveys of the ultimate beneficiaries were systematically used in planning of thematic qualitative studies in the subsequent period.

**Selected findings**

The study was characterised by the novelty in providing the findings on the net effect of the ESF intervention for Poland as it was the first analysis measuring the ESF net impact. In this respect the study provided evidence that the

---

127 The control group for the employed was not possible to be identified due to obstacles in retrieving personal data records on the project beneficiaries.


129 Including the measurement of the net effect of ESF intervention.
ESF support does improve employment and job opportunities in Poland, as confirmed through the comparison of situation of the ESF project participants with those who did not receive any support (control groups). The net effect was estimated at 11% for the beneficiaries of Action 1.2 and at 16% for the ultimate beneficiaries of Action 1.3. This finding has not been analysed further particularly as it did not convey into proposals or clear recommendations as confirmed by the MA.

Similarly, the finding reporting on the level of effectiveness of the respective types of assistance has been assessed as not robust enough for further investigation in the framework of this case study. The most effective type of support provided to the ultimate beneficiaries proved to be grants to start their own business activity. In the final group of the unemployed beneficiaries to whom the grant was awarded, the percentage of those people performing regular work was far higher than among other (respectively 97% and 61%). The Managing Authorities reported that although the study was analysing the situation of beneficiaries after six months of completion of the participation in projects, the lack of information on the performance of the beneficiaries in the long time perspective could not sufficiently justify transferring such or similar findings into credible recommendations (the analysis currently underway in the framework of the 2007-2013 programming period under the "Study on the values of indicators achieved for regional component of OP Human Capital 2007-2013" foresees the evaluation of sustainability of support in the form of two rounds of panel study at the turn of 2010/2011 and 2011/2012).

The results of the impact study revealed that after six months 42% of the unemployed recipients of the assistance was employed, 16% started their own business, 14% continued learning and 28% did not continue education or remained unemployed. By reaching these targets all the monitoring indicators, as foreseen for the ultimate beneficiaries covered by the ESF support under SOP HRD, were exceeded. However, the likelihood of taking a permanent job differed depending (inter alia) on the gender and the situation of unemployed women has improved less than men's situation: after about six months after the completion of participation in the projects 51% of women had a permanent job compared to two-thirds (68%) of men. The study did allow for observation of lower effectiveness of ESF support towards the unemployed women when compared to men. This finding was reported by the authors of the study on the level of the majority of Actions directed to both the representatives of both gender:

<table>
<thead>
<tr>
<th>Percentage (%) of men and women employed after 6 months by Measure/Action</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>67</td>
<td>53</td>
</tr>
<tr>
<td>1.3</td>
<td>73</td>
<td>56</td>
</tr>
<tr>
<td>1.4</td>
<td>19</td>
<td>21</td>
</tr>
<tr>
<td>1.5</td>
<td>40</td>
<td>25</td>
</tr>
</tbody>
</table>

---

130 Excluding Action 1.6 under SOP HRD which was addressed to the women with low and outdated skills and to women running business or intending to start a business.
Selected learning opportunities

Respecting the principles of gender equality is one of the rules of horizontal policies of the European Union in line with the provisions of the Treaty of Amsterdam art. 2 and 3\(^{131}\). This commitment of the Member States is respected also through the implementation of programmes and projects co-financed by the Structural Funds. The ESF co-financed projects should provide for actions aimed at developing solutions to support equal opportunities, reconciliation of private and professional life, mobility and retraining, as well as promoting equal opportunities for women and men in the labour market and combat stereotypical perceptions of male and female roles in working life and social life.

In the case of the study investigated the final and the quarterly study reports provided the evidence that the principle of equal opportunities for men and woman in the SOP HRD projects might have been treated relatively superficial.

The authors of the study reported that women were very close in being successful in finding an employment (41 % of woman compared to 45%) yet the men did start their own business two times more often then women (when surveyed 10% of women and 23% of men declared having their own business). Measuring the net effect in the impact analysis concerned provided another evidence that men performed systematically better than women. Whereas the difference was not significant for Action 1.2 (11% for women and 13% for men), the net effect in the second group (Action 1.3) was nearly three times higher (23%) for men than women (8%).

The Ministry of Economy and Labour, Department for European Social Funds Implementation acting as the Managing Authority, was responsible for the implementation of SOP HRD 2004 -2006 with the respective implementing authorities responsible for individual Programme Measures. In late 2004 the MA established the Steering Group for SOP HRD Evaluation to coordinate the evaluation process and in particular provide support to the process of contracting evaluations or identification of the areas which could further be subjected to deeper analysis. The respective quarterly reports from the impact analysis have been undergoing a series of internal (within the ESF Implementation Department) and external consultations (within the Evaluation Steering Group). The findings of the consecutive reports were also discussed at the SOP HRD Monitoring Committee and further investigated through the implementation of “The Study on a perspective of equal gender opportunities within SOP 2004-2006”\(^{132}\). Another evaluation “The OP HC pilot projects study”\(^{133}\) also confirmed the lesson learnt of weaknesses in respecting the principle of gender equality in case of projects implemented under the new OP HC 2007-2013.

“The Study on a perspective of equal gender opportunities within SOP 2004-2006” reported that although no cases on violation of the principle of equal opportunities for women and men were identified, the activities implemented in the ESF projects were rarely accompanied by in-depth reflection on their

\(^{131}\) Treaty of Amsterdam amending the Treaty on European Union, the Treaties establishing the European Communities and related acts, Official Journal C 340, 10 November 1997


\(^{133}\) Badanie projektów pilotażowych PO KL, Pozarządowa Agencja Ewaluacji przy Stowarzyszeniu BORIS, 2008.
impact on the situation of women and men. This was a consequence of low awareness of the perspective of gender mainstreaming, and the lack of knowledge on how to take concrete action arising from the rule. The study showed that a gender equality perspective, understood as ensuring that all activities and initiatives in a full and active way at all stages of the implementation take into account the impact on women and men, was not fully applicable to the SOP HRD. This applied to both institutional as well as a project level. The diagnosis of the situation conducted under the study also clearly indicated the need for particular training at all levels of the programme implementation. The authors of the "OP HC pilot projects study" reported that the description of the horizontal policy principle of gender equality should not only be limited to a declaration of application of the rule in the projects implemented - this being often the case and understood mainly in terms of an equal number of representatives of both gender in the recruitment project activities. It was also stressed that the aspect of equal opportunities for women and men should be emphasised through equal access to goods, rights, opportunities, resources and benefits, responsibilities in both private and public life.

The findings of the reports have been consequently disseminated at the national and regional level to raise the low awareness of the public in the area of equal gender opportunities in planning for ESF co-financed projects. Both the staff of the institutions responsible for implementation and project applicants did not have adequate knowledge necessary to conduct a thorough analysis of the projects in accordance with the principle of equal opportunities. This finding was converted into several operational recommendations.

The Monitoring Committee for SOP HRD appointed by the Managing Authority ensured the correct monitoring, evaluation and effective implementation of the Operational Programme. The SOP HRD Monitoring Committee by resolution 27 dated 17th March 2008 adopted a number of recommendations to ensure the implementation of the principle assuring equal access to opportunities of participation in project activities for both men and women. The effective implementation of the principle of equal opportunities for women and men under the OP HC was also discussed by the Monitoring Committee for SOP HRD and OP Human Capital in June 2008.

As a consequence a range of actions to strengthen compliance with the principle of equal opportunities between women and men were implemented including inter alia:

- In March 2008 the Managing Authority established the Working Group on Principles of Equal Opportunities for Women and Men, which included representatives of the Intermediate Bodies, 2nd Level Intermediate Bodies (IB2 and the OP HC Managing Authority). The main tasks of the Group is initiation of any actions to promote gender equality within the framework of OP HC, exchange of experiences and information on the implementation of the principle of equal opportunities. The Working Group members and staff of the MA were covered by the training on the principles of equal opportunities for women and men.
• In 2009, the "minimum standard" of implementation of the principle of equal opportunities for women and men was introduced to the system of implementation of OP HC. Within the Substantive Assessment Sheet for HC OP calls for proposals project co-financing application checklist checking the "minimum standard" was introduced. This list will allow members of the Project Evaluation Committee to ascertain whether the grant application meets the minimum requirements for compliance with the principle of equality opportunities for women and men.

• The MA prepared the ‘Agenda for equality opportunities for women and men within the Human Capital’ which provides for a series of activities, to be conducted in 2009-2013, to raise awareness about equal opportunities for men and women among the staff of the institutions involved in implementation of Human Capital and the potential project applicants (including training, conferences).

• From the second half of 2008 MA of OP HC requires that the Contractors of any evaluation studies examine the compliance with the principle of equal opportunities. This solution has also been recommended for IB and IB2.

• The guide on “The principle of equal opportunities for women and men in the OP HC projects” was prepared and addressed primarily to the project applicants, with major focus placed on process of preparation of the application to guide preparation of an analysis of the situation of women and men and a gender impact assessment, required under the OP HC in each application for funding. This assessment makes it possible to verify whether the project actually accounts for the needs of women and men, for example by offering different, individualised forms of support corresponding to the different needs of participants of the project.

The impact study analysed is considered as sufficiently relevant and its usefulness has been confirmed. The findings of the study were used indirectly, for designing the subsequent analysis which as a result strengthened the robustness of the initial findings of the impact study. The study was planned in a way to meet the needs of the ESF Implementation Department in their evaluation, programming and monitoring functions, mainly through its ongoing character and periodic reporting as required by the information suppliers. The gaps in respecting the principle of gender equality has been conveyed and converted into feasible recommendations. These are continuously disseminated to the national, regional and local level actors involved in the ESF projects implementation in line with the ‘Agenda for equality opportunities for women and men within the Human Capital’.

**Information sources**

**Interviews**


---

134 Principles of Selecting Projects under the Human Capital Operational Programme, April 2009.

135 ‘Zasada równości szans kobiet mężczyzn w projektach PO KL”, Ministerstwo Rozwoju regionalnego, Departament Zarządzania EFS.
Marta Rawluszko – Gender Mainstreaming Expert (30.04.2010)
Appendix 9 – Return on investments

Reverse calculation
The research synthesis did not reveal any clear picture concerning the return on investment (see pages 11 and 25), an issue which is assessed as a knowledge gap.

Acknowledging that problem, the European Commission suggested that the study team provide their own estimates of the range of potential rates of returns on the basis of reasonable assumptions.

In this spirit the study team carried out a series of reverse calculations showing the conditions that need to be fulfilled for a training programme to be profitable enough as seen from the society’s point of view.

As regards public investments, it is normal procedure to prepare a cost-benefit analysis including both individual and social dimensions, and to consider that the rate of return should reach a reasonable threshold. Discussion within the study team ended in a bracket of 6-8% for this threshold. A search into the EC publications suggests that a 4% threshold should be preferred136.

In order to undertake the calculation, the study team has designed three series of scenarios including: (1) type of training, (2) cost of the training, (3) impact, and (4) individual and social benefits in case of successful impact.

Considering the coverage of ESF supported activities, three types of investments would deserve to be studied in this section:

- Training of long term unemployed with a return-to-work purpose
- Training of low skilled employees in at-risk enterprises in order to prevent unemployment
- Support to training systems at sector level in order to compensate for underinvestment at the level of individual firms.

In fact, only the first type of training is considered in this study because:

- It is relatively well known in terms of benefits, and it is covered by the four country studies of this report
- Both other types of training correspond to knowledge gaps

First series of scenarios

Types of training
Training programmes targeted at the unemployed may be combined with other types of activities such as counselling or job search assistance. The literature review indicates that such a mix of instruments is the most effective

136 European Commission, Directorate-General Regional Policy, Evaluation and additionality (August 2006) Guidance on the methodology for carrying out cost-benefit analysis (Programming period 2007-2013). “The Commission proposes the following indicative benchmarks for the social discount rate: 5,5% for the Cohesion countries and 3,5% for the others. Member States may wish to justify different values reflecting specific socio-economic conditions. For instance, the Commissariat Général du Plan, France, recently lowered its reference to 4%, while the UK Treasury consistently applies a 3,5% social discount rate for public sector investents”.
approach. In line with this lesson, the following scenarios are generally based on a 4 month training\(^{137}\) combined with a 9 month individual coaching.

The effectiveness of the training depends on the targeted people and their specific employability problem. This is something the existing studies take into account, but relatively little is known about the differences between social groups, except that training of young job seekers tends to be less effective than average. The study team’s scenarios are therefore based on a training programme targeted at “average” long term unemployed, i.e. adult low skilled people.

It must be expected that there are considerable differences between the effectiveness of training, depending on the subject, length and whether the training is focusing on a specific type of job or is more general training, and more generally the quality of training. Little is known about that, but research seems to indicate that shorter training programmes are often more effective than longer ones. In this respect, a 4 month training duration is a reasonable assumption. As regards ESF, it can be considered as a medium duration (see Table 14 and Table 27).

**Costs of training**

The costs of a labour market training programme are primarily the expenditure on planning and designing the programme, selecting participants, renting premises, and paying the training staff, plus administrative fees and overheads. In the case of the study team’s scenarios, the cost of a four month training and coaching programme for 17 participants is estimated in four steps:

- Direct cost is first estimated as the gross salary of 9 person-month (designer, trainer, coach);
- Overheads are assumed to be 125% of the above, including premises and functioning of the training institution;
- An additional 20% is added in order to include the cost of running the training system.
- For the purpose of the study, the cost of a person-month is estimated at € 2000 (monthly median salary on the labour market, all social contributions included);

Building on these assumptions, the overall cost of the training is:

- Total cost: € 47 000.00
- Cost per participant: € 2 750.00
- Cost per hour\(^{138}\) : € 5.80

The time of the unemployed participants is counted as zero, as they are unemployed. It is also assumed that all participants are living in the local area, and do not need accommodation.

**Individual and social benefits**

First, it is assumed that that only 85% of participants take part in the entire programme. Some find a job in the meantime, and others just drop out. They may not have needed the training, or they may have learned enough during

\(^{137}\) Training courses are assumed to gather 17 participants in average.

the first part of the programme. No positive or negative impact is counted in any case. Drop-outs do not reduce the cost of the training.

Second, it is assumed that there is a 10% lock-in effect during the training. This means that 10% of those finalising the training would have found a job during this period if they had not participated in the programme. The corresponding negative impact is valued at the same price as that of the benefits (see later).

The success rate is the proportion of participants who find a job within one year after the finalisation of their training, and who would have remained unemployed without the training. The literature review and the country studies suggest that this proportion varies in a range of 5 to 15%. The scenarios do not select any specific point within this range. Three values are used in successive calculations {5%; 10%; 15%}.

The individual benefit is assumed to be the salary. Because participants are poorly employable people, they may not find high quality and full time jobs. For this reason, the benefit is estimated at 60% of the average monthly salary. An equivalent option would be to include (1) wage minus welfare benefits as a profit for the successful participant, and (2) reduction in welfare benefits as a profit for the society.

On the top of individual benefits, it is worth adding a certain proportion of social benefits, i.e. satisfaction of having a job (see page 25), plus the multiple benefits for the whole society\(^{139}\) arising from the fact someone exits unemployment, e.g. improvement of family and social relationships, reduction in informal economy and anti-social behaviour. In the various scenarios, social benefits represent 10 to 150% of individual benefits. This wide range of assumptions reflect the almost total lack of knowledge in this area.

Another key assumption is the duration of the benefits, which may be related to ageing, skill obsolescence, and business cycles. The Belgian country study shows that the duration of benefits is in a range of 1 to 2 years (see Figure 11 and Figure 12), but there is a doubt about the validity of the findings (see 4.5.1). In a large and recent study carried out in the USA\(^{140}\) the annual benefits were shown to grow constantly from year one to year 4. Referring to this lesson, the study team considers that the duration of benefits may vary in a range of 3 to 10 years.

Last but not least, the analysis needs to pay attention to the very important issue of substitution, i.e. the fact that someone else becomes or remains unemployed for the very reason that a participant found a job. Substitution is assumed to be limited if the training correspond to some skill mismatch on the local labour market. Knowledge is however completely missing on this issue. This is why the study team did not make any assumption in this respect. The proportion of substitution among successful participants may range from 0 to 100%.

Cost effectiveness

The return on a given investment is indeed strongly affected by the relationship between cost and impact. Highly effective types of training would achieve high success rates and low substitution effects at low cost. Once again, this issue is assessed as a being a knowledge gap. For this reason, some scenarios

\(^{139}\) Externalities  
\(^{140}\) Reported in Heinrich C. (2010) Contribution to the Conference ‘Shaping the Future of the ESF’
build upon the assumption that the same impacts may be achieved by a standard training (i.e. 4 month duration costing € 2 750 by participant), and by a very efficient training (i.e. 2 month duration costing € 1375 by participant).
Appendix 10– Focus group

Minutes of the focus group meeting 'Learning from impact evaluation' held in Brussels on 10/06/2010

Presentation of the study and the meeting

Attendance

Participants in the meeting:
- Paulo SEVERATI, ISFOL, Italy
- Misa LABARILE, European Commission, DG EMPL
- Sigried CASPAR, European Commission, DG EMPL
- Mathieu LEFEVRE, European Commission, DG EMPL
- Benedict WAUTERS, Managing Authority, Flanders

Participants in the follow-up telephone conference:
- Simon MARLOW, Managing Authority, England
- Malgorzata JABLECKA-KILUK, Managing Authority, Poland

Purpose of meeting

The focus addressed the issue of learning from impact evaluations. Four questions were discussed in this respect:
- Where should we learn?
- How to learn valid lessons?
- Should we make learning mandatory?
- Can we learn across Member States?

The discussion was based on a working document summarising the available knowledge on the impact of human capital investments, and the findings of four case studies investigating the actual utilisation of impact evaluations that were carried out in respectively Poland, Hungary, Italy and Belgium.

Knowledge gaps: where is learning urgent?

Presentation

The case studies and the review of available knowledge conclude that there are more knowledge gaps than lessons learned as regards the impact of ESF-like investments in Human Capital.

A reasonable amount of knowledge is available on the effect of training employees with secured jobs, and assisting job seekers through training and assistance.

Yet, there are knowledge gaps such as:
- Workers at risk, whether this risk applies to the economic sector, the firm or the individual himself (e.g. obsolete skills);
- Employable inactive people (especially elders);
- Poorly employable people.

These gaps are mainly due to the few number of studies dealing with these issues, and the difficulty to draw conclusions from existing studies.

The study team identified six areas where learning is a priority:
- Relating costs and effects ... since knowledge can be found about the costs or the effects but the connection between the two is not enough known; from our literature review, this seems to be also the case in studies outside Europe.
- ‘At-risk’ sectors and firms;
- Skill obsolescence;
- Effects of specific investments ...
  e.g. specific training approaches or specific publics (women, younger, elder, vulnerable); knowledge is not stabilised in this area, for instance, studies show contradictory results about the efficiency of training for women compare to men. The conclusion that training is less effective for young job seekers seems to be the only consensus in this area;
- Effects over the business cycle;
- Effects beyond trainees;
- Effects of a macroeconomic nature ...
  such as substitution effect in relation to skill mismatch.

**Discussion**

The participants pointed out that the effect of the **quality of training** has not been covered by the study up to now. There is indeed a knowledge gap in this area. Some studies specify whether the training was certified or not, but no connection is made with the subsequent impact on trainees. There is evidence that off-the-job training is more effective than on-the-job training, and that labour oriented training achieves better results. As regards the unemployed, training needs to be associated with complementary assistance in order to be effective (e.g. counselling, job search assistance, incentives to take a job).

The participants stress the importance of the **quality of the training system**. Italian regions require that training institutes be certified in order to be funded, but this certification process focuses on regulatory compliance rather than factors securing good results. In some regions there are thousands of training agencies which run their business with limited connection with the labour market in a context of strong inertia in the training policy. Call for tenders tend to remain unchanged and vulnerable target groups or poorly employable people are often left aside.

As regards quality, a determining factor is the **quality of teachers/trainers**, as well as the number of trainees per class/group. In the area of education, a study based on PISA data reached the conclusion that the quality of teachers and their level of professionalism accounts for more than any other factor in the positive outcome of training.

Participants do not recommend that quality be assessed through client satisfaction surveys. Trainers should also be consulted as they may have an accurate view of the training process. More important would be to undertake post training studies relating quality and outcomes.

Another success factor is a close connection between training and human resource management. For instance, the Flanders’ Managing Authority requires that companies develop their training plan, and even their global human resource strategy, to a minimum level before funding their training programme. This approach builds upon a lesson learned, i.e. training as such is not converted into sustainable skill improvements without fertile conditions in the working environment. The working organisation plays a crucial role in this process, since there would be no benefit from the training without some professional evolution of the trainee.

Finally, the issue of **future skill needs** is an important one, and especially that of who should assess such needs. Some Managing Authorities rely on
companies’ own studies on this matter. In Flanders, the Managing Authority carries out studies on “bottleneck jobs”, i.e. obsolete skills and skill mismatch, and public resources are targeted at these skill gaps in priority.

**Methods: how to learn valid lessons?**

*Presentation*

The study recommends not establishing an absolute hierarchy of methods. Methodological choices always depend on the context in which the evaluation takes place. The study team presented their provisional views about a kind of decision tree enabling evaluation managers to adapt to the constraints, by answering successive questions such as:

- Does a counterfactual make sense?
- Can we access individual data from monitoring / statistical databases over a long time period?
- Is it possible to carry out surveys over a long enough time period?
- Do we have sufficient resources?

First, a counterfactual makes sense if, and only if, the impact under study occurs at individual level (trainee, training group, enterprise), there are limited interferences with wider systems, and macro-level phenomena can be neglected. In the frequent cases where such conditions are not matched, theory-based evaluation methods are to be chosen, e.g. contribution analysis.

If a counterfactual makes sense, then the evaluation may proceed through an attribution analysis. The next question is that of duration. Both the literature review and the investigations in the four countries show that the return on training investment is a matter of years. For instance, the Polish study was not assessed as credible enough for policy-making because impacts were analysed only six months after the training. Conversely, in Flanders, the study covered a 2-3 year period and demonstrated that short term impacts tend to be negative, while longer term impacts become positive. Therefore, the second question is that of timing. Is it possible to wait long enough, for learning? Or to learn from old enough programmes, which is certainly a better option? Is it possible to collect data over a long enough time period. Repeating surveys on the same sample over years is knowledgeably difficult and a preferable option is to access adequate databases. If this is feasible, then the best approach (at least in the context of human capital investments) is to carry out an econometric analysis applying to a long chronological dataset including a quasi-experimental control group.

In case it makes sense to analyse impacts over a shorter time period, then a the evaluation should proceed through a mix of monitoring data and post training surveys (up to one or two years after training). The next question is that of the feasibility and acceptability of a randomised control trial (RCT). This option deserves to be preferred at this stage if it is feasible. Otherwise, a relevant alternative is an econometric analysis of two separate random samples (participants and non-experimental control group), as was done in the case of Poland, but over a longer time period.

Considering the above recommendations, there are three basic ways of improving the quality of impact evaluation:

- Accessing and connecting databases, from monitoring systems, public employment service and statistical offices;
- Learn from long term observation, through methodological approaches that capture long term effects such as duration models and longitudinal surveys; however, the issue of the lengthening of the learning process should be managed;
• Learn from rich datasets by combining monitoring data, pre- and post-training surveys, and internal control groups.

Discussion

The participants stressed the need to focus on relevant evaluation questions. In particular, there is a need to specify which type of training is to be evaluated, and what are the success criteria. Programmes and their evaluations also tend to be weak in terms of eliciting the causal chains of what happens after the training, or even what is expected to happen. Such assumptions should be a basis for both monitoring and evaluation.

The issue of data collection was also highlighted. The quality of impact analyses may be particularly weak if it relies on datasets that results from pooling data from many sources without sufficient quality checks. Another important issue is to plan evaluations long enough in advance in connection with policy design and implementation. Usually, data collection problems arise if the evaluation is too distant from the policy.

Accessing databases may raise privacy issues. In the case of England, a recent scandal has entailed severe limitations in the access to databases in order to prevent accidental leakages of individual information. Analysing datasets internally is the current solution. In Italy, institutional research institutes cannot be refused an access to databases, but this privilege does not extend to academic research. Another option could be to certify some organisations which would be given a fully secured access to personal data, which would then be made public in an anonymous form. In general, participants considered that privacy issues should be handled more professionally, but should not serve as an excuse for restricting access to databases. They suggested that the EC disseminate good practices of data accessibility which could inspire national authorities.

The participants agreed that there is no absolute hierarchy of methods. They stressed the interest of analysing the impact of training over a long enough time period, e.g. three years. It was also recommended that counterfactual approaches be combined with theory-based approaches eliciting and testing causal chains. They considered that the main difficulties applying RCTs are related to political and ethical acceptability, and not to time, money, or methodology.

The participants suggested that impact evaluations apply less on large scale routine programmes and more on pilots with an aim to learn transferable lessons.

Finally, the participants insisted on the need to develop the practice of peer reviews in the area of impact evaluations.

Regulations: should learning become mandatory?

Presentation

Through its investigation into the actual use of four impact evaluations, the study team found that policy-making decisions were taken in line with the findings of the impact analyses. However, an in-depth contribution analysis found that the causes of this seemingly good news had almost nothing to do with what was assumed to be a successful learning process.

Four reasons prevented the learning process to work:

• Limited political ownership;
- Excessive focus on methodology ... for instance dissemination efforts devoted much room to methodology at the expense of substantial information that would have interested policy makers;
- Lack of explanatory studies ... and insufficient focus on causal chains and explanations of the findings;
- Lack of selectivity ... possibly due the willingness to evaluate the programme as a whole; this prevented sound knowledge to be generated.

Considering that the four cases studies applied to some of the best European practices in terms of impact evaluation in the area of the ESF, and that almost no learning could be observed, there is a challenging problem.

In order to address this problem, the preliminary recommendations of the study team were:

- Rolling evaluation plans ... such multi-annual plans should be endorsed by policy-makers;
- Better and more selective impact evaluation ... it should be recognised that it is simply not possible to measure and analyse the impacts of everything; through their evaluation plans, Managing Authorities should decide to learn at least one lesson in at least one specific area, at least every five years; such lessons should be accumulated in a kind of shared knowledge tank; by this means, the amount of knowledge would pile up and all programmes could benefit from the lessons learned by others;
- Quality assessment of the lessons learned... quality of evaluations should not be assessed globally; on the contrary, each individual lesson should be assessed in terms of credibility and transferability;
- Annual follow up report ... many institutions that have good evaluation practice already do that; it consists of a summary of lessons and recommendations arising from previous evaluations, complemented by a monitoring of which actions were taken, or not taken subsequently; this report should be discussed by policy makers;
- Knowledge synthesis within evaluations ... each evaluation should contain a chapter on what is already known from former researches or studies;
- Knowledge synthesis before decision... any decision involving significant financial consequences should refer to what has been learned in the past, in line with EC’s practice of impact assessment.

Altogether, the recommendations are assumed to progressively create a culture of learning at the level of policy makers. This assumption is made credible if compared with the development of an evaluation culture in the public administration across the EU over the last 20 years, as a consequence of the fact that evaluation was mandatory in the context of structural funds.

Discussion

The participants agree with the diagnosis of politicians’ lack of learning culture. A show-case is highlighted in Flanders. Policy-makers are just starting to require that any company applying to a public support should demonstrate that their training programme is part of a broader human resource strategy. This policy reform is fully due to a group of university professors that have done years of sustained lobbying for bringing their knowledge into the political
sphere, however, their knowledge resulted from a series of research works carried out 40 years earlier or more. This story exemplifies the fact that the learning curve is extremely slow if knowledge is pushed by suppliers and not pulled by users. In the case of England, the culture of ‘evidence-based policy’ is said to have led to significant cultural changes.

Participants welcomed the study team’s recommendations and acknowledged that this was the first attempt to address the issue of learning. In particular, it was said that evaluation plans are already recommended under the convergence objective but not under the regional competitiveness objective. The issue is often discussed and there are good practices. The recommendation was considered as feasible.

Participants also reflected on the issue of learning in a context of multi-level governance. More can be learned from specific impact evaluations (specific kind of training, specific public), but evaluation is also needed at the level of programmes, Member States, and the EU. At present, it is almost impossible to concluding on impacts at upper levels. Progress could be made by combining evaluative information at these levels with lessons learned from lower level impact evaluations. In this respect, knowledge generated by impact evaluations throughout the ESF system should be considered as a public good.

For a participant, the recommendation of requiring evidence before financing ESF projects would be perceived as another administrative burden for beneficiaries who already encounter difficulties to cope with the existing ones. The opposite opinion is expressed by another participant who feels that all draft recommendations of the study team could easily be implemented in the form of pure administrative compliance without any change at the level of policy-makers. Experience suggests that politicians will never learn if there is not a real risk of loosing money in case of poor learning.

Incentives: how to learn together?

Presentation

For both methodological and financial reasons, it is impossible to learn from impact evaluation everywhere at any time. Yet, an extended learning could be achieved through, and only through, a lively knowledge community. In its draft recommendations, the study team suggests that the EU create incentives for developing such communities, especially where there are knowledge gaps.

Learning together is assessed as being:

- A necessity … since good impact evaluations are long, costly, and specific;
- An opportunity … to strengthen the validity of the lessons learned and to assess their transferability;
- A feasible approach … as the study team has been able to conduct this approach for this study and is also aware of other cross-country studies (on-going Visegrad study) or cross-region studies (as in Northern Italy).

The sharing of knowledge should also extend to ESF-like investments inside and outside Europe.
Discussion

The participants confirmed that learning is very slow in the areas that are identified as knowledge gaps (see section 1). For this reason, there is a rationale for the ESF to create incentives for impact evaluations, studies, or action oriented research projects in these areas. Moreover, incentives should apply to knowledge communities where scientific and political languages could merge. The participants did not trust that the EU Research Framework Programme would be a relevant instrument for building such knowledge communities.

The participants were of the opinion that learning could be hampered by the diversity of impact evaluation practice across Europe. If there are doubts about the reliability of data, or the validity of analyses, the transfer of knowledge will be discouraged. For this reason, the participants suggested that precise methodological guidance be disseminated and that incentives be conditional to peer reviews.

As regards the incentives themselves, the participants suggested that financial resources be earmarked for sound and relevant evaluation, i.e. peer-reviewed impact evaluations focusing on knowledge gaps. Such resources should be detached from technical assistance, and even detached from the programme itself, something which means that no evaluation money would go to Managing Authorities in case of poor evaluation practice.
## Appendix 11 – Vocabulary of the study

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attrition bias</td>
<td>Bias stemming from the fact that (1) some members of the treatment group drop out from the training or (2) some members of the treatment or control group cannot be reached by the follow-up survey</td>
<td></td>
</tr>
<tr>
<td>Change</td>
<td>Evolution of a variable / indicator which is assumed to be affected by the programme, e.g. change in the labour market status of surveyed people</td>
<td>Do not use ‘gross effect’ or ‘outcome’</td>
</tr>
<tr>
<td>Conclusion</td>
<td>Any kind of value judgement building upon the evaluation findings</td>
<td></td>
</tr>
<tr>
<td>Effectiveness</td>
<td>The extent to which actual effects correspond to the objectives</td>
<td>Do not use ‘net effectiveness’</td>
</tr>
<tr>
<td>Experiment, experimental design</td>
<td>Approach to impact evaluation based on a randomised control group</td>
<td>Do not use ‘social experiment’</td>
</tr>
<tr>
<td>External control group</td>
<td>Control group drawn from statistical sources that are external to the study, e.g. statistical database</td>
<td></td>
</tr>
<tr>
<td>External validity</td>
<td>The extent to which the findings derived from the analysis of a given data set (e.g. samples) may also apply to the whole population under study</td>
<td></td>
</tr>
<tr>
<td>Finding</td>
<td>Any kind of knowledge arising from data analysis</td>
<td>Synonymous of ‘analysis result’, a term which should be avoided since there is a risk of confusion with the results of the programme</td>
</tr>
<tr>
<td>Impact</td>
<td>Part of the observed change which is attributable to the programme</td>
<td>Considered as synonymous of ‘effect’, ‘results’, and ‘outcome’ in the framework of this study</td>
</tr>
<tr>
<td>Individual data</td>
<td>Data applying to individual people or companies, e.g. participation, labour market status, socio-economic characteristics</td>
<td>Synonymous of ‘micro data’</td>
</tr>
<tr>
<td>Internal control group</td>
<td>Control group built in the study through a special survey</td>
<td></td>
</tr>
<tr>
<td>Internal validity</td>
<td>The extent to which findings derive from the data set under study without any bias</td>
<td></td>
</tr>
<tr>
<td><strong>Term</strong></td>
<td><strong>Definition</strong></td>
<td><strong>Comments</strong></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Lesson</td>
<td>Conclusion which is potentially transferable</td>
<td>Lessons might be useful for designing the next programme, or other similar programmes</td>
</tr>
<tr>
<td>Quasi-experiment, quasi-experimental design</td>
<td>Impact evaluation with a control group which is not randomised</td>
<td>Do not use 'social experiment' since this term also applies to pilot-testing innovative policies</td>
</tr>
<tr>
<td>Scalability</td>
<td>The extent to which the conclusions applying to a local pilot programme may also apply to a future routine programme in a wider context</td>
<td></td>
</tr>
<tr>
<td>Selection bias</td>
<td>Bias stemming from the fact that individuals in a non-randomised control group might not have applied to or been selected by the programme, because of some individual characteristics, knowing that these characteristics may also explain a part of their subsequent labour market status.</td>
<td></td>
</tr>
<tr>
<td>Transferability</td>
<td>The extent to which the lessons learnt through one or several evaluations may also apply to other programmes in other regions, countries, or policy areas</td>
<td></td>
</tr>
<tr>
<td>Treatment group</td>
<td>Individual who are randomly enrolled in the programme in the case of an experimental design</td>
<td>To be used in the case of experiments only. A term such as ‘surveyed participants’ could be used in the case of quasi-experiments</td>
</tr>
</tbody>
</table>