



Department
for Work &
Pensions



European Social Fund impact evaluation: research design and scoping study

A report of research carried out by Ecorys on
behalf of the Department for Work and Pensions

July 2018

Research Report 962

A report of research carried out by Ecorys on behalf of the
Department for Work and Pensions

Crown copyright 2018.

You may re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. To view this licence, visit <http://www.nationalarchives.gov.uk/doc/open-government-licence/> or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gsi.gov.uk.

This document/publication is also available on our website at:
<https://www.gov.uk/government/organisations/department-for-work-pensions/about/research#research-publications>

If you would like to know more about DWP research, please email:
Socialresearch@dwp.gsi.gov.uk

July 2018 .

ISBN 978-1-5286-0529-8

Views expressed in this report are not necessarily those of the Department for Work and Pensions or any other Government Department.

Executive summary

The European Social Fund (ESF) is the European Union's (EU) main instrument for supporting citizens to access and progress in employment, and to promote social inclusion. Guided by an Operational Programme (OP), the ESF in England is being implemented in the current 2014-2020 ESF programming period through a range of provision that aims to address employment and social inclusion issues, along with supporting skills development. To examine the extent to which these aims are met, the European Commission (EC) requires that EU Member States evaluate the impact of the ESF Programme. This report aims to provide the groundwork for this through considering how an impact evaluation of the ESF could best be undertaken to meet the Commission's requirement.

This study was part-funded by ESF Technical Assistance.

Contents

- Acknowledgements.....i
- Authors’ credits ii
- List of abbreviations iii
- Glossary of termsv
- Summary vi
- 1 Introduction..... 16
 - 1.1 Background to the ESF and the design study 16
 - 1.1.1 Overview of the ESF in England 16
 - 1.1.2 The requirement for an ESF evaluation..... 17
 - 1.2 Study aims and objectives..... 18
 - 1.2.1 Study aims and objectives..... 18
 - 1.2.2 Key research questions..... 18
 - 1.3 Methodology..... 19
 - 1.4 Report structure 20
- 2 Review of the ESF programme and developing a theory of change.....21
 - 2.1 Programme architecture: key features and considerations for the evaluation21
 - 2.2 Issues arising from the stakeholder consultations 24
 - 2.3 Intervention logics for the programme 28
 - 2.3.1 Intervention logic for PA 1 29
 - 2.3.2 Intervention logic for PA 2 32
 - 2.3.3 Additional insights from the theory of change workshop 35
 - 2.3.4 The role of the intervention logic in the proposed impact evaluation.....37
- 3 Key design considerations..... 38
 - 3.1 Scope and focus of the evaluation 38
 - 3.1.1 Coverage..... 38
 - 3.1.2 Focus of the impact evaluation..... 39

ESF impact evaluation: research design and scoping study

- 3.2 Determining the level of analysis 40
- 3.3 Evaluation criteria: effectiveness and impact.....42
- 3.4 Considering value for money..... 43
- 4 Potential counterfactual impact evaluation approaches.....47
 - 4.1 CIE – initial feasibility assessment47
 - 4.1.1 Defining CIE in the context of ESF impact evaluation47
 - 4.1.2 The potential for assessing ESF impacts through CIE 49
 - 4.1.3 The potential for assessing other ESF results through CIE..... 53
 - 4.2 Review of potential CIE designs 54
 - 4.3 Assignment onto provision 59
 - 4.4 Data availability and potential comparison groups..... 65
 - 4.4.1 Data availability for the treatment group..... 65
 - 4.4.2 Potential comparison groups and data availability..... 66
 - 4.5 Sample size considerations.....74
- 5 Recommended design for the impact evaluation 78
 - 5.1 Recommended approach and rationale 78
 - 5.1.1 Recommended approach to the CIE..... 78
 - 5.1.2 Recommended approach to assessing other impacts..... 79
 - 5.1.3 Potential approach to incorporating Priority Axis 3.....81
 - 5.1.4 Recommended additional data sources81
 - 5.1.5 Recommended approach to assessing value for money.....81
 - 5.1.6 Learning from related evaluations 82
 - 5.1.7 Producing early evidence 82
 - 5.2 Proposed research questions for the impact evaluation 83
 - 5.3 Challenges, risks and limitations 84
- 6 Annex one: additional technical detail..... 86
 - 6.1 Intervention logics at the IP level..... 86
 - 6.2 Activity groupings 94
- 7 Annex two: Research tool for the stakeholder consultations 97

List of Tables

Table 2.1 Overview of the programme by Thematic Objective, Priority Axis and Investment Priority 22

Table 2.2 Contextual factors that could affect effectiveness and results..... 26

Table 4.1 High level examination of CIE feasibility against programme impacts 50

Table 4.2 Summary overview of potential CIE designs 55

Table 4.3 Participant eligibility criteria by CFO, PA and IP61

Table 4.4 Principal target groups of provision by IP 67

Table 4.5 Overview of comparison groups and potential data sources..... 69

Table 4.6 Target participant numbers.....76

List of Figures

Figure 2.1 Intervention logic for Priority Axis 1 30

Figure 2.2 Intervention logic for Priority Axis 2 33

Figure 2.3 Activity groupings across ESF IPs and their causal links with the intended results and impacts of the programme..... 36

Acknowledgements

The study team wish to thank all of the individuals that contributed to the completion of this research. In particular, we wish to acknowledge the helpful advice, support and guidance of Department for Work and Pensions staff involved with the study, notably Leanne Dew, Anna Taylor, William Farbrother and Nicholas Campbell. We would also like to thank a range of other DWP colleagues who gave valuable advice and input to the issues examined through the research. Finally, we would also like to express gratitude to all of the other individuals interviewed as part of the study fieldwork, and to those who contributed to the theory of change and impact evaluation workshops held to inform the issues under consideration in this report.

Authors' credits

Ian Atkinson is an Associate Director at Ecorys. He specialises in employment and labour market research and has a particular interest in the design and implementation of employment programme evaluations. As part of this Ian has contributed to a range of studies for the European Commission and UK Government Departments around this theme.

Anja Meierkord was previously at Ecorys, where she conducted research on employment, education, training and skills policies, and now works for the OECD. She is an econometrician who works in the labour market and skills fields.

Rachel Wooldridge is a researcher at Ecorys whose background is in researching and evaluating programmes designed to support vulnerable groups in the areas of employment, skills and social inclusion.

List of abbreviations

APS – Annual Population Survey
ATT – Average effect of Treatment on the Treated
BAME – Black, Asian and Minority Ethnic
BEIS – Business, Energy and Industrial Strategy
CBA – Cost-Benefit Analysis
CEA – Cost-Effectiveness Analysis
CEM – Coarsened Exact Matching
CIA – Conditional Independence Assumption
CIE – Counterfactual Impact Evaluation
CIS – Customer Information System
DFE – Department for Education
DiD – Difference-in-differences
DWP – Department for Work and Pensions
EC – European Commission
ELSA – English Longitudinal Study of Ageing
ERDF – European Regional Development Fund
ESF – European Social Fund
ESFD – European Social Fund Division
ESIF – European Structural and Investment Funds
ESOL – English for Speakers of Other Languages
EU – European Union
GDP – Gross Domestic Product
GVA – Gross Value Added
HESA – Higher Education Statistics Agency
HMRC – Her Majesty’s Revenue and Customs
IAG – Information, advice and guidance
IMD – Index of Multiple Deprivation
ILR – Individualised Learner Records
IP – Investment Priority
ITT – Invitation to Tender
IVA – Instrumental Variables Approach
LA – Local Authority
LEO – Longitudinal Education Outcomes
LEP – Local Enterprise Partnership
LSYPE – Longitudinal Survey of Young People in England
LTU – Long-term unemployed
MA – Managing Authority
MI – Management Information
MDE – Minimum Detectable Effect
NBD – National Benefits Database
NPD – National Pupil Database
NEET – Not in Employment, Education or Training
OECD – Organisation for Economic Cooperation and Development

ESF impact evaluation: research design and scoping study

OP – Operational Programme

PSM – Propensity Score Matching

RDD – Regression Discontinuity Design

RTI – Real Time Information

SME – Small and Medium Enterprise

STEM – Science, Technology, Engineering and Mathematics

ToC – Theory of Change

UC – Universal Credit

VfM – Value for Money

YEI – Youth Employment Initiative

Glossary of terms

Counterfactual impact evaluation (CIE) – CIE is a type of impact evaluation using a counterfactual analysis approach. Counterfactual analysis compares the real observed outcomes of an intervention with the outcomes that would have been achieved had the intervention not been in place (the counterfactual).

Difference-in-differences – Difference-in-differences is a statistical technique used to estimate the impact of an intervention on a set of specified outcomes. It mimics an experimental research design by comparing the average change on these outcomes experienced by a treatment group with that experienced by a comparison group over time. The technique is therefore used to estimate post-programme differences in outcomes between a treatment and comparison group, with the approach including an assessment of pre-programme differences.

European Social Fund (ESF) – The ESF is the European Union's (EU's) main financial instrument for supporting jobs, helping people get better jobs and ensuring fairer job opportunities for EU citizens. The European Commission works with countries to set the ESF's priorities and determine how it spends its resources.

(ESF) Managing Authority (MA) - The Department for Work and Pensions (DWP) ESF Division is the ESF Managing Authority for England. It has overall responsibility for administering and managing the ESF and reporting to the European Commission.

(ESF) Operational Programme (OP) – Operational Programmes describe the priorities for ESF activities and their objectives at national or regional levels within the European Union.

Local Enterprise Partnerships (LEPs) – LEPs are voluntary partnerships of local authorities and businesses with responsibility for deciding on general economic priorities at the local level.

Propensity score matching (PSM) – PSM is a statistical technique used to estimate the impact of an intervention on a set of specific outcomes. It mimics an experimental research design by comparing outcomes for a treatment group and a statistically generated comparison group, which is similar to the treatment group in its composition.

Theory-based evaluation – An evaluation approach which involves the explicit development of hypotheses to test through the gathering of relevant evidence, for example concerning whether and in what ways the objectives of an intervention have been achieved.

Theory of change – Theory of change is an evaluation methodology drawing on work developed in the United States to evaluate community and social programmes. The approach involves identifying the logic behind an intervention in terms of its rationale and aim, key objectives, inputs, activities and short, medium and long term outcomes and testing this 'intervention logic' through a range of evaluative methods. In this study, theory of change has been used to describe how different ESF activities are intended to lead to particular outcomes (i.e. the intervention logic) as a basis to develop evaluative methods to test if, how and why this has happened.

Summary

1. Introduction

This summary presents the key findings of a study undertaken to inform the design of a planned impact evaluation of the 2014-2020 European Social Fund (ESF) programme in England. The study was undertaken by Ecorys between July 2017 and February 2018. It is intended to help fulfil the European Commission's (EC) requirements for monitoring and evaluation in respect of the ESF, specifically the need to undertake an impact evaluation of the fund.

2. Methodology

The study sought to inform the design of a robust and cost effective impact evaluation for the ESF programme through several tasks. These involved:

1. A desk-based review of ESF programme and project related documentation
2. Stakeholder consultations to inform the development of an intervention logic for the programme, including with representatives of the ESF Managing Authority (MA) (x7), national co-financing organisations (CFOs) (x5), local CFOs (x1), and local European Structural and Investment Fund (ESIF) sub-committees (x6). Consultation also engaged two other stakeholders involved with the European Regional Development Fund (ERDF) programme and evaluation
3. Consultations with six further stakeholders to discuss possible counterfactual impact evaluation (CIE)¹ approaches, including representatives from the Department for Work and Pensions (DWP), the Department for Education (DFE) and the Welsh Government
4. A workshop to explore the ESF intervention logic further, considering the causal links between ESF activities and the programme's intended results and impacts, along with thinking through the contextual factors that might affect these
5. An impact evaluation workshop, bringing stakeholders with insights into CIE approaches and available datasets together to explore the potential scope and focus of a CIE
6. Further development and appraisal of potential approaches to evaluating the impact of the ESF, the main findings of which are presented below

3. Key findings

3.1 Overview of the recommended approach

A mixed-method evaluation of the ESF is recommended in order to assess the totality of potential results and impacts stemming from the programme, whilst also examining the effectiveness of ESF provision in generating these. CIE approaches to assessing impact are likely to be feasible, but only in respect of some results and impacts

¹ A counterfactual impact evaluation compares the outcomes of an intervention with the outcomes that would have been achieved had the intervention not been in place (the 'counterfactual'), typically by comparing outcomes between a treatment group and comparison group.

the programme seeks to generate. The suggested approach would thus combine CIE techniques, specifically propensity score matching (PSM) and difference-in-differences (DiD) with theory-based evaluation. The latter approach is recommended specifically to assess potential results and impacts that are unable to be feasibly estimated through a CIE, along with helping to explore the reasons behind the programme's achievement or otherwise of its objectives.

3.2 Recommended approach to CIE

Based on a review of CIE designs and their likely feasibility, it is recommended that propensity score matching (PSM), drawing on ESF Management Information (MI) and administrative datasets, forms the core of a CIE approach within the broader evaluation design advocated. The review indicated several advantages to a PSM approach over others, specifically in terms of its potential to be successfully implemented in order to estimate the key intended employment and skills impacts of the ESF at the level of participating individuals. A review of administrative datasets suggests that the requirements for operationalising such an approach can be met, conditional on access being granted to the data and the ability to link individual level data across different data sources.

PSM also has the potential to be used for subgroup analysis, for example in respect of particular ESF target groups, though such an approach would need to be carefully designed given the likely overlaps between such groups. While the assessment does not preclude the use of CIE to estimate the impact of particular types of provision, for example at the level of ESF Investment Priority (IP), there is a need to be realistic over the likely limitations to any 'sub-treatment' analysis within the PSM model proposed. The main challenge is being able to adequately define and discretely analyse effects, particularly at the level of specific activities such as volunteering or advice to promote employability. There are two main concerns in respect of this: that ESF provision typically combines a range of support within single projects, lessening the potential for effects of particular activities to be disaggregated and discretely analysed, along with the extent to which ESF MI could reliably be used to identify that individuals have participated in particular activities.

Where the approach is feasible and applicable, it is recommended that PSM be complemented with DiD analysis when estimating impacts. The rationale for this relates to the need to help address potential selection bias remaining from the application of a PSM approach. While PSM can be used to effectively model a comparison group with similar characteristics to a treatment group (i.e. ESF participants), some unobservable characteristics such as motivation can cause bias to emerge. DiD is thus suggested as a way of correcting for this.

Examining the potential of the approach in light of the ESF's intended outcomes indicates that PSM-DiD analysis should principally be used to estimate the effect of ESF provision on entry to employment, (re-) engagement with education and training, and skills outcomes in respect of accredited levels and qualifications. These represent well defined outcomes at the level of participating individuals with a clear change in state that can be measured and estimated through a CIE. Broader, more structural, and less clearly defined impacts and results would, it is contended, need to be assessed through alternative approaches.

3.3 Recommended approach to assessing other impacts

To cover a broader set of impacts than those able to be estimated through CIE techniques, the study findings indicate that a theory-based evaluation design should complement the PSM-DiD approach suggested. This would develop hypotheses concerning the presumed results of particular activities or types of support within ESF provision, with these being tested through the collation of an appropriate evidence base. It is suggested that this encompasses qualitative data gathered from in-depth interviews with ESF providers, participants, and other stakeholders, potentially on a case study basis, allied to analysis of ESF MI and insights from Leavers Survey data. The latter would be used to complement, contextualise, and compare against qualitative fieldwork findings to build up a detailed understanding of the role of ESF in generating its intended results and impacts.

Within a theory-based approach, the study recommends that effectiveness considerations form part of the focus. This is primarily to help understand which activities lead to which results and impacts, along with how and why, rather than focusing on results and impacts in isolation. This should help ensure that an ESF impact evaluation offers lessons for related programmes and provision. However, our assessment suggests that an impact evaluation should only consider process aspects where likely effects on the ESF intervention logic are direct and significant. From this perspective, examining effectiveness should focus solely on explaining why impacts occurred or otherwise in respect of the provision and support offered.

3.4 Recommended approach to assessing value for money

As well as considering effectiveness and impact, the extent to which ESF offers value for money will be a central concern when evaluating the programme – not least in terms of helping to inform decisions over the design of future provision. Examining the options available for such an assessment highlighted that the DWP's in-house cost-benefit analysis (CBA) model could be used as a basis for this where it is possible to monetise both costs and benefits. Detailed data on costs and spend is available from ESF MI to feed into this model. Equally, the model can be used to estimate many of the likely fiscal and social benefits of the programme as they relate to employment, including those concerning, for instance, changes in income, benefits payments and tax receipts. Using the Department's in-house model as the basis for a CBA is also recommended in terms of its potential to facilitate comparison with other employment programmes in the context of assessing the value for money of the ESF.

The CBA frameworks or approaches developed by other government departments can be used to complement this model, where appropriate. While this will require further exploration in designing the precise parameters of a CBA framework in the ESF context, it is anticipated that such models should facilitate assessment of the returns resulting from skills development or from activities to reduce re-offending. However, caution should be applied, and appropriate caveats used, to account for differences in how these approaches used by other departments vary from that used by DWP. Where other results of the programme are considered significant to incorporate into a value for money (VfM) assessment, but are difficult to monetise such as those relating to wellbeing, it is recommended that additional approaches such as cost-effectiveness analysis (CEA) be incorporated as appropriate into the assessment. The results of the CIE should, moreover, be fed into any models developed and used, where feasible, to enhance the accuracy of estimates made.

While combining CBA and other approaches suited to assessing value where benefits are unable to be reliably monetised is recommended as an approach for the evaluation, such combined approaches within an overall VfM assessment will need to be approached with caution. In particular, while the potential CBA approach outlined can be used to derive a VfM ratio or ratios, this would not be able to be directly combined with ratios derived from (potential) CEA approaches for effects unable to be monetised. With careful development, however, different techniques should be able to be used in combination to provide a detailed and rounded understanding of the value offered by the programme.

3.5 Challenges, risks and potential limitations to the recommended approach

While the design study gives confidence that a robust impact evaluation of the ESF is feasible, and that CIE techniques can be effectively deployed within this, it is important to recognise that there are some challenges and risks to be worked through when finalising the evaluation approach. As is common practice, it is thus recommended that study timescales facilitate a further scoping and refinement phase as part of the implementation of an impact evaluation. At this point, a key challenge relates to the need for further consideration around how to maximise the accuracy and robustness of any subgroup or sub-treatment analysis, particularly in relation to applying CIE techniques. Similar care and attention will also be needed to overcome some of the challenges inherent in mixed method evaluations, particularly in terms of how different evidence sources and findings are weighted and combined.

Key risks largely relate to the potential for some of the datasets that are central to the recommended approach being unavailable for unforeseen reasons, and/or delays in accessing these. Changes in programme delivery arrangements, timescales, or lifetime due to the United Kingdom's exit from the European Union (EU) also represent a risk in this sense, as does the potential for changed interpretations of the requirements for an impact evaluation on the part of the EC.

In terms of limitations, it should be clear that the scope of a CIE as part of the recommended approach can only assess some of the totality of the ESF's likely results and impacts. The need to apply such techniques only where feasible, where they can produce robust estimates of impact, and where they are methodologically defensible, also needs to be kept in mind when implementing and evaluation. Likewise, in respect of the theory-based component of the recommended approach, limitations in terms of precisely determining and proving causality will have to be acknowledged to the degree they are inherent in such approaches.

1 Introduction

This report was commissioned by the Department for Work and Pensions (DWP) to inform the design of a planned impact evaluation of the 2014-2020 European Social Fund (ESF) programme in England. Work to inform the report was undertaken by Ecorys between July 2017 and February 2018. This introductory chapter sets out the context for the design and scoping study, details its aims and objectives, and presents the methodology used to undertake it. It also outlines the structure of the remainder of the report.

1.1 Background to the ESF and the design study

1.1.1 Overview of the ESF in England

As the European Union's (EU) main instrument for supporting citizens to access and progress in employment, and to promote social inclusion, the ESF involves the European Commission (EC) and EU countries working in partnership to set the priorities of the fund. Forming part of the 'EU Structural and Investment Funds (ESIF) Growth Programme', alongside the European Regional Development Fund (ERDF) and part of the European Agricultural Fund for Rural Development (EAFRD),² implementation of the ESF in England is guided by the ESF Operational Programme (OP) 2014-2020.³ The OP sets the objectives and priorities for the delivery of the ESF over the 2014-2020 programming period. Delivery is structured around three priority axes:

- Priority Axis 1, Inclusive Labour Markets, combines activities to address employment and social inclusion issues
- Priority Axis 2, Skills for Growth, supports the development of skills
- Priority Axis 3, Technical Assistance, supports the management, monitoring and evaluation of the fund.

² In England the ESIF include the ESF, ERDF, EAFRD, and the European Maritime and Fisheries Fund (EMFF) which is managed by the Marine Management Organisation (MMO), a non-departmental public body sponsored by the Department for Environment, Food & Rural Affairs (DEFRA).

³ *European Social Fund Operational Programme 2014-2020*, Department for Work and Pensions, 2015

As the Managing Authority (MA) for the ESF programme, the Department for Work and Pensions (DWP) is responsible for producing the OP and managing delivery of the programme.⁴ As detailed further below, DWP also acts as a co-financing organisation (CFO)⁵ for the ESF programme. Furthermore, through a team functionally independent from both DWP's role as the MA for the fund and its CFO role, DWP is also responsible for evaluation in respect of the programme.

At the national level, DWP's role as MA for the programme is overseen by the Growth Programme Board (GPB), the Programme Monitoring Committee (PMC) for the fund. The Programme Board is responsible for monitoring the ESF, advising the MA, and helping align the ESF with the other ESIF funds where possible. At the sub-national level, local ESIF Committees operate in each Local Enterprise Partnership (LEP) area. These committees complement the functions of the MA and GPB by providing advice on local priorities, for example through ensuring that the specifications developed to procure ESF provision locally reflect local priorities.

ESF provision in England is delivered through two main routes: through co-financing organisations (CFOs), or through 'direct bids' to the MA. There are four national CFOs operating as part of the 2014-2020 ESF Programme – DWP, the Education and Skills Funding Agency (ESFA), Her Majesty's Prison and Probation Service (HMPPS) and the Big Lottery Fund (BLF). These national CFOs run invitations to tender or calls for proposals to select providers to deliver provision, with input from the local ESIF sub-committees for particular areas. Direct bids involve potential ESF providers responding to open calls for proposals run by the MA. The aim of this process is to enable the MA to develop calls (with the support of local ESIF committees) to achieve OP priorities and meet local needs.

A smaller proportion of ESF provision is commissioned through local co-financing routes. In particular, in London the Greater London Authority (GLA) acts as an Intermediary Body (IB). In this role the GLA oversees and manages the design and commissioning of ESF provision. Oversight of this role is provided by the MA. The GLA also has CFO status, granted in March 2016. In addition, Trafford Council, on behalf of Greater Manchester Combined Authority (GMCA), was granted CFO status in March 2016, enabling it to design, procure and manage ESF provision locally whilst still, as with all CFOs, being overseen by DWP acting as the MA. As well as this local co-financing, a small proportion of the ESF is used to support Community Led Local Development (CLLD), with the small community based projects funded under CLLD being commissioned by a 'direct bid' mechanism to the ESF MA.

1.1.2 The requirement for an ESF evaluation

The ESF evaluation team within DWP leads on monitoring and evaluating the OP for the 2014-2020 programming period. As noted, this team is functionally independent from DWP activities in respect of its role as the MA for the ESF and its CFO role in respect of the fund. The requirement for monitoring and evaluation is stipulated by

⁴ In the context of this report references to DWP are understood to refer to the Department's specific role as MA for the ESF, rather than DWP's broader remit and functions (i.e. as a CFO and, through a functionally independent team, its responsibility for evaluation of the programme), unless specified otherwise.

⁵ CFOs are public bodies that bring together ESF and domestic funding for employment and skills so that ESF complements national programmes.

the EC, with the relevant guidance outlining that MAs should ensure that evaluations, including evaluations to assess effectiveness, efficiency and impact, are carried out for each ESF programme.⁶ To ensure that evaluation activity has an adequate focus on impact, rather than concentrating solely on implementation issues, the Commission's guidance makes assessing the extent to which the objectives of each priority axis have been achieved compulsory.⁷

1.2 Study aims and objectives

1.2.1 Study aims and objectives

As the EC guidance for ESF evaluation also outlines, assessing impact comes with a number of challenges. These include, for example, determining a suitable methodology and assessing the availability of data to inform estimations of impact. This design and scoping study is intended to examine and address such challenges, a number of which are inherent in conducting impact evaluations. More broadly, the study was commissioned to enable the ESF evaluation team within DWP to design a robust and cost-effective impact evaluation for the ESF programme. In doing so, the study is likewise intended to help ensure that the ESF evaluation requirements outlined above are met.

Within these broad aims, a number of specific objectives were developed by DWP for the design and scoping study. These were to:

1. Develop and refine an intervention logic model(s)⁸ for the ESF 2014-20 Programme (covering both priority axis 1 and priority axis 2)
2. Provide advice on the feasibility of impact evaluation design options, including: the possible scope of the evaluation, the evaluation questions it should seek to answer, and possible outcome measures
3. Outline the methodological approaches which could feasibly be used to conduct an impact evaluation, and the pros, cons, requirements, and risks of each method
4. Provide an outline of the contractor's recommended research approach.

1.2.2 Key research questions

A number of questions were devised by DWP for the research design and scoping study to address as follows:

⁶ *Programming period 2014-2020: Monitoring and Evaluation of European Cohesion Policy, European Social Fund Guidance document*, European Commission, 2015

⁷ Ibid.

⁸ 'Intervention logic models' are associated with evaluation techniques seeking to define, and then test, the presumed outcomes and impacts from an intervention. While they can vary in terms of presentation, they typically seek to capture the rationale for an intervention, the inputs supporting its development and delivery, the activities used to implement it, and the presumed outcomes and impacts intended to flow from these activities.

Intervention logic modelling

1. What is the logical framework for the intervention(s) under ESF programme for priority axis 1 (inclusive labour markets) and priority axis 2 (skills for growth)?
2. What are the intended outcomes and impacts of the ESF 14-20 programme?
3. What are the mechanisms through which the interventions aim to deliver the intended outcomes of moving people into employment (priority axis 1) and improving the skills of the workforce (priority axis 2)?

Research design and scoping

4. Based on the logical framework for the intervention, what measures of success (outcomes) could be assessed through an impact evaluation?
5. What should the scope of an ESF programme impact evaluation be?
6. What evaluation questions should it aim to answer?
7. What variables could be used to measure cost-benefit and value for money of the ESF programme/ESF provision?
8. To what extent is it possible to conduct a counterfactual impact evaluation? What are the alternative approaches to gathering robust evidence on programme impacts?
9. What approaches and methodology could be used to conduct the impact evaluation? What are the pros and cons of each approach?
10. What data is available for use in an impact evaluation and what does this mean in terms of possibilities for constructing a suitable control group or counterfactual?
11. Given the complexities of the ESF programme, what is the best way to measure impact, given data robustness requirements, contextual and policy factors, and value for money?

1.3 Methodology

The methodology developed to meet the above study aims and objectives, and the address the specified research questions, involved the following principal stages:

1. Desk-based review of ESF programme and project related documentation.
2. Consultations with ESF stakeholders to inform the development of intervention logics and an accompanying theory of change⁹ for the programme, comprising interviews with representatives of the MA (x7), national CFOs (x5), local CFOs (x1), and local ESIF sub-committees (x6), accompanied by two further interviews with other stakeholders, including those involved with the ERDF programme and evaluation.

⁹ In this context, 'theory of change' refers to a narrative account of how the ESF programme is intended to lead to the intended outputs, outcomes and impacts the programme seeks to occasion. It therefore complements the intervention logic models developed for the study specifying the programme's rationale, inputs, activities, outcomes and impacts.

3. Consultations with six further stakeholders to discuss possible counterfactual impact evaluation (CIE)¹⁰ approaches, including representatives from DWP, the Department for Education (DFE) and the Welsh Government.
4. A theory of change workshop, bringing a range of stakeholders connected with the programme together to consider the presumed causal links between ESF activities and the programme's intended results and impacts, along with thinking through the contextual factors that might affect the extent to which the theory of change is evident in practice.
5. An impact evaluation workshop, bringing stakeholders with insights into CIE approaches and available datasets together to explore the potential scope and focus of a CIE, possible comparison groups, and any related work using similar methods that has been, or is being, undertaken.
6. Further development and appraisal of potential approaches to evaluating the impact of the ESF, the results of which are presented in chapters two to five below alongside insights gained from the preceding tasks.

1.4 Report structure

The remainder of the report is structured as follows:

- **Chapter two** identifies key features of the ESF programme requiring consideration, details the results of the stakeholder consultations, and presents suggested intervention logics and a theory of change to guide the planned impact evaluation.
- **Chapter three** examines a series of 'key design considerations' for the planned evaluation, including evaluation scope and focus, evaluation criteria, and how VfM might be assessed.
- **Chapter four** assesses the feasibility of undertaking a CIE of the ESF programme, examining likely design options and considering some key issues such as data availability, potential comparison groups, and required sample sizes.
- **Chapter five** presents a recommended design for the ESF impact evaluation based on the findings of the preceding analysis.
 - o **Annex one** includes additional technical detail to that presented in the main report
 - o **Annex two** presents the research tool used to inform the main group of stakeholder consultations.

¹⁰ A counterfactual impact evaluation compares the outcomes of an intervention with the outcomes that would have been achieved had the intervention not been in place, typically by comparing outcomes for those receiving an intervention with a 'comparison group' of similar individuals not subject to that intervention.

2 Review of the ESF programme and developing a theory of change

Through examining the ESF programme and gathering stakeholder inputs, this chapter aims to provide the required underpinning to develop an approach to the proposed ESF impact evaluation. It first identifies the key features of the ESF programme architecture and related considerations for the evaluation, before summarising relevant insights emerging from the interviews undertaken with programme stakeholders. Building on this, a theory of change for the programme is then outlined, including intervention logics for the programme as a whole and for the key Priority Axes likely to be the focus of the impact evaluation.

2.1 Programme architecture: key features and considerations for the evaluation

Taken as a whole, the 2014-2020 ESF Programme in England is relatively wide-ranging and complex. Its aims encompass supporting individuals to move closer to and into work, promoting social inclusion, addressing youth unemployment, and developing the skills of individuals both in and outside of the labour market to support productivity and the adaptability of the workforce. The programme is designed around three priority axes – inclusive labour markets (Priority Axis (PA) 1), skills for growth (PA 2), and technical assistance (PA3) – within which there are a number of Investment Priorities (IPs). The IPs within the programme, along with the Thematic Objectives (TOs) it seeks to address, form the building blocks for each PA and are drawn from the ESF regulations set by the European Commission. Drawing upon the 2014-2020 ESF Operational Programme (OP), Table 2.1 below provides a summary overview of the TOs, PAs and IPs that make up the programme.

Table 2.1 Overview of the programme by Thematic Objective, Priority Axis and Investment Priority

Priority Axis (PA)	Thematic objective (TO)	Investment Priority (IP)
PA1: Inclusive Labour Markets Aims to increase participation in the labour market and thereby improve social inclusion and mobility.	TO8: Employment Set at the EU level to promote sustainable and quality employment and support labour mobility.	1.1 Access to employment for jobseekers and inactive people – to help those who are disadvantaged but still relatively close to the labour market to tackle their barriers to work, and enter and sustain employment. 1.2 Sustainable integration of young people – to focus on helping young people, particularly those who are not in employment, education or training (NEET), or at risk of being NEET, to participate in the labour market and learning. 1.3 Youth Employment Initiative (YEI) – to focus on helping young people, who are NEET, to participate in the labour market and learning in areas eligible for the YEI. ¹¹
	TO9: Social Exclusion and Poverty Set at the EU level to promote social inclusion and to combat poverty and discrimination.	1.4 Active inclusion – to help people who are more distant from the labour market and may face multiple disadvantages to tackle their multiple, complex and profound barriers to work and to move towards or into employment, or to sustain employment. 1.5 Community Led Local Development (CLLD) – to support activities initiated by local action groups.
PA2: Skills for Growth¹² Aims to support growth in the economy through supporting skills development amongst individuals.	TO10: Skills Set at the EU level to support investment in education, training and vocational training for skills and lifelong learning.	2.1 Enhancing equal access to lifelong learning – focuses on improving the skills of individuals to meet their goals and the needs of the local economy, primarily training, advising or supporting individuals, including those in work but at risk due to skills deficiencies or facing redundancy. 2.2 Improving the labour market relevance of education and training systems – focuses on improving employer participation and engagement in learning so that it is responsive to the needs of the local economy and more individuals progress into or within learning. This is primarily about improving partnerships and systems.
PA3: Technical Assistance Supports the effective and efficient delivery of the ESF.	PA3 supports delivery of the ESF programme as a whole, thereby covering all TOs.	PA3 supports delivery of the ESF programme as a whole, therefore covering all the above IPs.

¹¹ The YEI is geographically targeted, covering regions of the EU where the youth unemployment rate in 2012 was higher than 25 per cent, or where youth unemployment was more than 20 per cent but had increased by more than 30 per cent in 2012. Only parts of England are eligible for YEI funding.

¹² While PA2 relates primarily to TO10: Skills, it is also intended to contribute indirectly to thematic objectives 8 and 9 by training people in skills that sustain and enhance their employment and reduce their risk of social exclusion.

As well as the thematic dimension of TOs, PAs, and IPs, the ESF Programme also has a geographical dimension, covering three categories of region (less developed, transition and more developed). Regions are categorised based on regional Gross Domestic Product (GDP) per head compared to the European Union (EU) average. Less developed regions have a GDP per head below 75 per cent of the EU average, transition regions between 75 and 90 per cent of that average, and developed regions above 90 per cent of the EU average. Importantly, however, the Operational Programme (OP) makes the point that, in the English context, the extent of variation **within** regions, in terms of the need for ESF support, is often greater than that **between** regions. As the OP outlines:

“...the highest rates of unemployment are in areas that include major cities, either within transition or more developed regions; and pockets of deprivation are found in rural areas that are otherwise relatively affluent.”¹³

A further important feature of the ESF is the requirement it places for ‘match-funding’ to be provided at the level of European Union Member States in order to complement the monies provided from the European level through the fund. For co-financed provision, initially anticipated in the OP to account for 70 per cent of all provision, match funding is provided by the public bodies acting as co-financing organisations (CFOs) from mainstream budgets and programmes. For ‘direct bid’ provision, match funding must be secured by the organisation making the bid. In the context of designing an approach to evaluate the ESF programme, the match funding aspect requires consideration as part of determining the scope and focus of the planned study.

As outlined in the report introduction, the delivery architecture for the programme, encompassing CFO and direct bid provision, adds further complexity and challenges for undertaking a programme level evaluation able to assess the impact of the ESF as a whole, as well as its constituent parts. In particular, the programme design gives rise to a number of issues and challenges that any potential evaluation will need to address. These issues and challenges are examined in detail in chapter three, and can be summarised as:

- Determining the level or levels at which evaluative activity should focus – for example, at the ‘whole-programme’ level; at the TO, PA or IP levels; at the level of particular co-financed provision, such as that run by the Department for Work and Pensions (DWP) or Education and Skills Funding Agency (ESFA); or some combination of these or other options.
- Whether, and if so how, the geographical dimension of the programme should be reflected in an evaluation design; that is, the extent of focus on the less developed, transition and more developed regional aspect to the programme.
- How the value for money (VfM) of the provision to be evaluated can be assessed, given its heterogeneous nature and the range of different activities and target groups involved, including both those in work and those relatively far removed from the labour market for example.
- The extent to which the effectiveness of delivery, as opposed to just its results, requires consideration in assessing the ESF’s impact, and how this might be used to explain impacts if necessary, given the range of different delivery models apparent within the fund.

¹³ *European Social Fund Operational Programme 2014-2020*, Department for Work and Pensions, 2015, p.6

Desk research undertaken to inform the study also provided the opportunity to review a sample of ESF project applications and contracts approved at the delivery level. Covering projects within PA 1 and PA 2, and including projects within the national CFO provision and direct delivery projects, this review provided several further insights of note when considering the potential for an ESF impact evaluation. These included:

- The fact that, at the delivery level, there was a wide degree of variance in the approach to project delivery in terms of delivery models and approaches, along with the target groups provision focuses on.
- At the same time, there was more consistency in the primary objectives and intended results of provision, with these closely reflecting the key ESF OP aims around supporting individuals to move closer or into work and to develop their skills, thereby promoting social inclusion.
- The objectives of the provision commissioned strongly reflect the intended results and impacts of the IP within which they sit, as detailed above in table 2.1.
- Reflecting the intention of the OP that ESF provision should complement, and not duplicate, existing mainstream programmes, there is an evident focus across projects around complementing, and filling gaps in, such programmes.

2.2 Issues arising from the stakeholder consultations

As outlined in the report introduction, a series of consultations with stakeholders were used to test and refine initial intervention logics for the ESF programme, covering PA 1, PA 2 and the IPs within them. The draft intervention logics were produced through a review of key programme documentation, notably the OP. Further details of this process and the resulting intervention logics and theory of change are presented in section 2.3 below. The interviews also explored a number of additional issues of relevance to a potential design for evaluating the ESF's impact. These covered views on the aims of the programme and the intended results an evaluation should test, contextual factors that might affect intended results, how VfM might be assessed, and any other views on what the focus and scope of an evaluation should be. The consultations also provided an opportunity to explore any additional data collected around ESF provision, and the extent of any other evaluative activity.

Discussions around the purpose and aim of the ESF tended to closely reflect the articulation of this in the OP, namely supporting individuals to move closer to and into work, promoting social inclusion, addressing youth unemployment, and developing the skills of individuals. While there were some differences in emphasis, reflecting the place of the stakeholder concerned in relation to the programme, the fund was not seen as having any major additional aims outside of these that would influence any impact evaluation.

In line with this, views on the key outcomes an impact evaluation should seek to examine related to ESF results around entry to employment and skills development. The role of the ESF in respect of employed individuals was also noted, specifically in terms helping them to remain and progress in employment. From the social inclusion perspective, a number of stakeholders also felt that there should be a specific focus on examining the results of activity aimed at moving people closer to the labour

ESF impact evaluation: research design and scoping study

market, in terms of 'distance travelled', rather than just focusing on job-entries and sustainability. While this was seen as more challenging, given that such 'softer' outcomes may be harder to quantify, this was nonetheless seen as important in capturing the full impacts of the programme.

Stakeholders generally felt that the draft logic models produced through the desk review, covering PA 1, PA 2, and each IP within them, accurately captured the rationale, activities, intended results and impacts of the programme. Suggested refinements, which influenced the final intervention logics outlined in section 2.3, included:

- Taking note of the OP review underway at the time of this study, particularly in terms of reflecting the renewed clarity this seeks to give that PA 2 is able to develop the skills of unemployed people as well as those in work.
- The importance of assessing the aspects of the programme intended to support the localism agenda, hence explicitly reflecting the role of PA 2 in meeting local employer and skills needs and testing this.
- Ensuring that the assumed results of IP 1.4 (active inclusion) adequately reflect the focus on tackling barriers to move people closer to the labour market, as well as results relating to entering employment.
- Taking account of the varied definitions of sustained employment outcomes within the programme, depending on the CFO or provision concerned.
- Reflecting the interest there may be in progression towards employment in terms of distance travelled, the view being that this should explicitly be captured in the programme's intended results (particularly in IP 1.4).

In respect of contextual factors, perspectives on what may affect the presumed operation and results of the programme were gathered as part of the theory of change workshop held during the study, as well as from the individual stakeholder interviews. Table 2.2 below summarises a list of relevant factors from this exercise that may, when the impact evaluation is implemented, help to explain why aspects of the ESF intervention logic are or are not evident. It should be understood that the table simply reflects the stakeholder views offered, rather than endorsing their relevance or utility for an ESF impact evaluation. The contextual factors raised by stakeholders can be broadly categorised as internal to the ESF programme itself, or relating to 'external' issues such as labour market conditions. The most common external issues raised related to changes in economic conditions, potential effects of the UK's exit from the EU, and internal factors relating to delivery of the programme. The latter included the potential for the clarity of ESF requirements and guidance, or misunderstanding of these on the part of providers, to impact on how effectively the programme is delivered.

Table 2.2 Contextual factors that could affect effectiveness and results

Internal factors	
Factors identified	Effects on presumed intervention logic
Co-ordination of provision	Potential for duplication, reducing scale of results achieved
Role of LEPs	Ensuring local needs addressed
Clarity in ESF regulations	Effects on provider understanding and delivery
Evidence requirements and programme guidance	Positively: providing detail on outputs and results and helping inform effective delivery Negatively: Administrative burden affecting delivery / financial impact on delivery partners
Review of OP	May be positive in better reflecting labour market needs, local needs, and policy changes in the second half of the programme
Funding levels between different 'categories of region'	Potential for overlooked areas of deprivation within 'transition' and 'more developed' regions that are under-funded relative to 'less developed' regions.
Co-financing / direct bid context	Impact on second half of the programme – possibly greater resource for direct bids but also risk that too much direct bid provision leads to over-provision and duplication.
Funding and programme design	Impacts on provider behaviour influencing success or otherwise
Individuals moving between providers and restrictions on this	Confusion over outcomes claiming and potential for double counting and/or individuals not being claimed for (though audit / compliance checks should avoid this)
Amount of time providers spend with individuals, including considering this by different ESF priority groups	Effects on the extent to which presumed results and impacts can be observed in respect of the different ESF priority groups
Degree to which provision is oriented around class-based support for groups or individual-based activities	The relative effectiveness of different approaches and their relative prevalence may affect (scale of) results and impacts
Provider competition for outputs	Lack of information sharing and reduced effectiveness of provision overall
Lack of facility to record softer outcomes	Inability to fully capture impacts of the programme in some areas (e.g. effects of debt management support)
Length of intervention/ support	Could affect degree to which positive results are achieved and/or sustained
How far projects are designed to encourage ease of access and engagement for participants	Could affect levels of engagement and the extent to which different ESF priority groups are able to access support

ESF impact evaluation: research design and scoping study

Evaluation timing given cyclical project lifetime and cyclical customer journey	Potential to affect judgements over outcomes and impacts depending on, for example, when fieldwork is conducted or quantitative analysis of impacts undertaken
Effectiveness of provider outreach activity	Could affect levels of recruitment onto provision and hence the numbers supported and the results/impacts achieved
Difficulty of employer engagement	Reduced effectiveness in terms of anticipated numbers of participants and outcomes

External factors

Factors identified	Effects on presumed intervention logic
Local labour market conditions and range of jobs available	Effects on number of potential participants, depending on levels of local unemployment and activity, as well as the potential for individuals to access work through ESF support
Welfare reform in general and introduction of Universal Credit in particular	Numbers of potential participants and issues around eligibility; rent caps affecting activity of Housing Associations to focus more on employment provision
UK leaving the EU	Uncertainty – potentially affecting provider recruitment and retention of staff
Tighter labour market	Certain groups harder to find than envisaged
Geography	Potentially harder to support participants with next steps if in isolated areas and reluctant to travel
Wider LEP activity in local areas	Can complement but also compete with ESF provision, hence influencing outcomes
LEP approaches	At times may have a different understanding of what is needed to the provision on offer affecting outcomes; also can have a lack of understanding of the OP and procurement rules impacting on effectiveness and efficiency of provision
Wider economic conditions	Affecting engagement and success rates positively or negatively
Uncertainty of devolution	Affecting the design and development of provision locally

In terms of how an impact evaluation might address VfM considerations, several considerations were raised. Comparing unit costs of provision against results achieved was viewed as a possible basis for assessing value, in particular the extent to which expected unit costs for anticipated results were achieved in practice. Commonly, stakeholders also noted that VfM should be examined with reference to the context of the ESF provision. In particular, it was noted that any approach should consider the target group being supported, in that some groups may require more support and be more expensive, as well as potentially geographical context, in that provision may be more expensive in areas of greater population dispersal.

The intended results of ESF provision were also seen as important to capture as broadly as possible, so as to inform assessments of the benefits of the programme. For example, stakeholders cited potential benefits, or cost savings, around reducing

re-offending within Her Majesty's Prison and Probation Service (HMPPS) provision, along with improved health and wellbeing, as well as more obvious savings to the Treasury from reduced benefits claims due to ESF participants entering work. The need to disaggregate different activity, and assess costs on this basis, was also raised.

Additional suggestions concerning the evaluation as a whole included the importance of being able to inform the delivery of any successor programme to the ESF, or related programmes, particularly in terms of exploring what types of provision achieve the greatest impacts for the lowest cost. It was also noted that it would be helpful to be able to explain why impacts occurred in addition to simply estimating their scale. Alongside this, the importance of capturing the new role of LEPs in guiding the programme at the sub-regional level, and the relative success or otherwise of this in generating the intended results of the programme, was stressed by several stakeholders. Finally, the importance of robustly estimating the effect of the programme in relation to its key objective of supporting individuals was advanced as being a key requirement for the evaluation.

2.3 Intervention logics for the programme

As outlined in the report introduction, a key aim of the design and scoping study was to develop an intervention logic or logics for the programme. This enables a theory of change behind the ESF to be articulated – specifically, by describing how inputs and activities are intended to lead to results. Developing an understanding of the ESF's intervention logic and theory of change has several purposes in light of the proposed impact evaluation. These include the need to identify the intended results of activity that the evaluation should focus on, the importance of understanding the mechanisms through which the activities of the programme intend to lead to these results, and providing a basis on which different evaluation methods can build to assess the impact of the programme.

While the research questions developed to inform the design and scoping study anticipated that intervention logics would be developed for PA 1 and PA 2, during the initial desk review it became apparent that only focusing on the PA level would pre-judge the level of analysis at which evaluative activity should be undertaken. As explored in section 3.2, in theory the programme could be assessed as a whole, at the PA level, at the level of individual IPs that sit within the PAs, through focusing on the programmes of national CFO provision along with direct provision, by grouping similar projects together across different PAs and IPs, or some combination of all the above elements. As a result, drawing on the initial desk research, intervention logics were developed at both PA and IP levels, with draft versions of these informing the stakeholder consultations and theory of change workshop held as part of the study.

While it is important to understand the intervention logic of the ESF at different levels, and from different perspectives, the core of the programme can be captured through reference to PA1 and PA 2. These axes are intended to meet the main aims of the ESF – supporting individuals to move closer to and into work, promoting social inclusion, addressing youth unemployment, and developing the skills of individuals and responding to local skills needs. While PA 3 – covering technical assistance for programme implementation – forms part of the OP, it is of less relevance for the type

of impact evaluation that this design and scoping study aims to inform. Specifically, PA 3 is concerned more with **processes** to ensure that results and impacts are achieved, rather than directly generating results and impacts.

While the intervention logic at IP levels may have a role in the proposed ESF impact evaluation, as explored in section three of the report, the outline of the intervention logic that follows mainly concentrates on the PA level. This is for two main reasons: firstly, this is sufficient to achieve the main purposes of defining the intervention logic of the programme as outlined above; and, secondly, to maintain clarity in articulating the main ways in which the programme seeks to achieve its results and impacts.

2.3.1 Intervention logic for PA 1

The diagram below presents the intervention logic developed for PA 1. The theory of change implicit in this logic begins with the rationale for intervention and its objectives. As outlined in the OP, despite an increase in employment levels following the recessionary period of the late 2000s, there are still concerns that particular groups of people face disadvantage in entering and sustaining employment. The intervention under PA 1 thus reflects the objective of supporting access to employment amongst the unemployed and inactive, especially people over 50, those from an ethnic minority background, those who are disabled, those who are lone parents, and the most marginalised and disadvantaged young people.

Alongside this, the OP also outlines a rationale for addressing youth unemployment, particularly amongst those not in employment, education or training (NEET). The need to support NEET young people is predicated on evidence that a lack of skills, particularly basic skills such as English, Maths and ICT, is a key barrier for many young people to entering work or re-engaging with education and training. As the OP also reflects, those from the “most marginalised and disadvantaged backgrounds (such as care leavers) require more intensive and specialised support” by way of supporting the objective of helping young people into work, education or training.¹⁴

The rationale for intervention also relates to the objective of promoting social inclusion, and tackling poverty, through supporting individuals to move closer to and into the labour market. As the OP outlines, in 2013 “23% of the UK population were considered to be at risk of poverty or social exclusion according to the EU definition.”¹⁵ Evidence suggests that poverty is typically a symptom of deeper and more complex disadvantages, such as family problems, worklessness, crime and debt, low educational attainment, and drug and alcohol dependency.¹⁶ The OP thus outlines a need to address poverty and social exclusion through helping some of the most marginalised groups, including those furthest from the labour market, to move closer to and eventually into work. The OP also recognises a need to address poverty and social exclusion amongst marginalised communities, hence supporting community-led local development (CLLD) strategies under PA 1. CLLD aims to mobilise local actors, assets and resources to stimulate local growth and provide individual pathways for people from disadvantaged groups into work.

¹⁴ *European Social Fund Operational Programme 2014-2020*, Department for Work and Pensions, 2015, p.25

¹⁵ DWP. 2015. *European Social Fund Operational Programme 2014-2020* (p15-16)

¹⁶ *Ibid* (p16)

ESF impact evaluation: research design and scoping study

Figure 2.1 Intervention logic for Priority Axis 1

	Rationale	Overarching objectives and inputs	Activities	Qualitative Results	Quantitative Results	Impacts
Employment	Particular groups of people face disadvantage in entering and sustaining employment, especially people over 50, those from an ethnic minority background, those who are disabled, those who are lone parents, and the most marginalised and disadvantaged young people. Support is needed for such groups to reduce labour market inequalities and enhance the ability of all to find and remain in work.	<p>Objective: Supporting access to employment for job-seekers and inactive people, including the long-term unemployed and people far from the labour market.</p> <p>Inputs: €1,008 million ESF to improve labour market participation channelled through the ESF infrastructure (CFOs and direct bids). Overseen by the MA benefitting from advisory inputs from ESIF sub-committees.</p>	<ul style="list-style-type: none"> Additional and innovative approaches to pre-employment training; Additional support for long-term unemployed people, including those who have left the Work Programme, and including new approaches to work experience and training; Improving the employability and transferable skills of unemployed, inactive and disadvantaged people; As part of wider support, using self-employment as a route out of worklessness; Activities to encourage the unemployed to start and grow businesses (including social enterprises). 	<ul style="list-style-type: none"> Improved employability of long term unemployed people Individuals with labour market disadvantages can compete effectively in the labour market Improved employability of inactive people Unemployed and inactive people see their basic skills needs satisfied Women at a disadvantage in the labour market receive support Employment outcomes are sustained 	<ul style="list-style-type: none"> 22% of unemployed participants into employment on leaving 33% of inactive participants into employment or job search on leaving 4% participants gaining basic skills 36% participants with childcare needs receiving childcare support 31% (less developed regions) or 34% (other regions) in employment 6 months after leaving 	<ul style="list-style-type: none"> Higher employment rate Fewer barriers to employment for individuals from disadvantaged groups Reduced gender gap
Youth Unemployment	Evidence indicates that a lack of skills (particularly basic skills such as English, Maths and ICT) is a key barrier for many young people preventing them from entering work or re-engaging with education and training. In particular, those from the most marginalised and disadvantaged backgrounds (such as care leavers) require more intensive and specialised support, indicating a need for provision to support these young people, along with others who are NEET, to help them enter work or re-engage with education and training.	<p>Objective: Supporting the sustainable integration into the labour market of young people, in particular those not in employment, education or training, including young people at risk of social exclusion and young people from marginalised communities.</p> <p>Inputs: €160m from the Youth Employment Initiative, plus an equivalent amount of ESF money, channelled through the ESF infrastructure. Overseen by the MA benefitting from advisory inputs from ESIF sub-committees.</p>	<ul style="list-style-type: none"> Improving the employability and transferable skills of unemployed, inactive and disadvantaged people; Training workless people and those facing redundancy, to upgrade their skills or learn new skills (including basic skills and English for speakers of other languages); Responding flexibly to employer demand in local labour markets where specific needs are identified; Funding additional basic skills provision, e.g. in literacy and numeracy; Helping disadvantaged young people into traineeships and apprenticeships and support them to stay there; Increasing the provision of careers advice and guidance, and strengthening engagement with employers 	<ul style="list-style-type: none"> Additional traineeships and apprenticeships are provided Marginalised 15-18 year olds re-engage with education or training NEETs complete more effectively in the labour market Unemployed 18-24 year olds are provided with additional work experience and pre-employment training opportunities Young lone parents participate more effectively in the labour market 	<ul style="list-style-type: none"> 4% participants gaining basic skills 55% participants (below 25 years of age) in employment, including self-employment, or education/training upon leaving in less developed regions (43% in other regions) 70% of unemployed participants complete the YEI supported intervention 48% of unemployed participants receive an offer of employment, continued education, apprenticeship or traineeship upon leaving 31% (less developed regions) or 34% (other regions) in employment 6 months after leaving YEI only: 60% long term unemployed participants complete the YEI supported intervention 38% long term unemployed participants receive an offer of employment, continued education, apprenticeship or traineeship upon leaving 60% inactive participants not in education or training complete the YEI supported intervention 32% inactive participants not in education or training receive an offer of employment, continued education, apprenticeship or traineeship upon leaving 32% inactive participants not in education or training are in education/training, gaining a qualification, or are in employment, including self-employment, upon leaving 15% participants in continued education, training programmes leading to a qualification, an apprenticeship or a traineeship 6 months after leaving 33% participants in employment 6 months after leaving 3% participants in self employment 6 months after leaving 	<ul style="list-style-type: none"> Increased number of young people in education, employment or training Reduced number of young people who are NEET or at risk of being NEET Contributing to a reduction in NEET levels and in youth unemployment
Social Inclusion	In 2013 23% of the UK population were considered to be at risk of poverty or social exclusion according to the EU definition. Evidence suggests that poverty is typically a symptom of deeper and more complex disadvantages, such as family problems, worklessness, crime and debt, low educational attainment and drug and alcohol dependency. Supporting individuals closer to and into work is a key route to addressing these issues. Social exclusion also has a geographical dimension in that marginalised communities are particularly affected, leading to a need to mobilise local actors, assets and resources to stimulate local growth and provide individual pathways for people from disadvantaged groups into work.	<p>Objective: Promote social inclusion and combat poverty and discrimination, supporting those more distant from the labour market to tackle their multiple, complex and profound barriers to work and to move towards or into employment, or to sustain employment. Mobilise local action to develop employment pathways within marginalised communities.</p> <p>Inputs: €703 million to support social inclusion, including €136 million for Her Majesty's Prison and Probation Service activity and €55 million for Community Led Local Development, channelled through the ESF infrastructure. Overseen by the MA benefitting from advisory inputs from ESIF sub-committees.</p>	<ul style="list-style-type: none"> Tackling barriers faced by disadvantaged people in rural communities, lacking access to services; Supporting community-based action including Community Grants, community learning and community led local development; Targeting disadvantaged communities and deprived areas; Tackling barriers faced by people in workless households and troubled families or with multiple disadvantages, so they move towards, enter and make progress at work; Helping inactive people with potential to enter the labour market but needing considerable support to enter work; Targeting people facing barriers - homelessness, debt, drug or alcohol problems, and mental health issues; Tackling discrimination and barriers faced by ethnic minorities and disabled people; Integrating offenders into the labour market; Supporting the development and growth of social enterprises 	<ul style="list-style-type: none"> People in particularly deprived areas move towards or into employment People with multiple and complex barriers to participation move closer to the labour market through having their barriers addressed Improved employability of offenders Marginalised individuals are supported to re-engage with education, training or employment 	<ul style="list-style-type: none"> 20% (less developed regions) or 19% (other regions) participants in education or training on leaving 17% (less developed regions) or 16% (other regions) unemployed participants in employment, including self-employment on leaving 29% of inactive participants into employment or job search on leaving 17% participants in education or training on leaving 20% (less developed regions) or 22% (other regions) participants in employment, including self-employment, 6 months after leaving 14% of unemployed participants into employment on leaving 27% of inactive participants into employment or job search on leaving 36% participants with childcare needs receiving childcare support 	<ul style="list-style-type: none"> Contributing to addressing the root causes of poverty that are barriers to work Individuals are moved closer to – or into – work Sustained bottom-up regeneration and economic development

ESF impact evaluation: research design and scoping study

As figure 2.1 indicates, to meet the objectives of PA 1, and reflect the rationale for intervention, a number of inputs support the development and delivery of activities. These inputs include the allocated funding from the ESF, alongside a range of inputs that can be broadly characterised as relating to the 'ESF delivery infrastructure'. Such inputs include, for example, the role of DWP as the MA for the fund, and the role of the Growth Programme Board in overseeing the programme. The co-financing mechanisms established to support delivery can also be viewed as inputs, as can strategic inputs into programme design on the part of Local Enterprise Partnerships (LEPs) and the role of the local ESIF sub-committees. At the level of ESF delivery, the expertise of providers offering support to participants would likewise form one of the inputs involved.

Collectively, these financial and other inputs are used to design, develop and implement a range of provision delivered under PA 1. Such provision, or activities, include: support with employability skills such as job-searching; work experience; skills development, including of basic skills; providing access to qualifications; supporting participants to access traineeships and apprenticeships; providing support to address barriers to work such as debt, housing or mental health issues; and supporting offenders to develop skills and re-engage with the labour market upon release. As this indicates, the ESF is able to support a wide range of activities under PA 1, with the aim of developing tailored support to move individuals and priority groups closer to and into work, along with helping individuals to develop skills and (re-) engage with education and training.

As the intervention logic diagram on the previous page shows, these activities, whether developed under IPs supporting employment, youth employment or social inclusion, are intended to lead to a series of results. This intervention logic, of a series of inputs, responding to a rationale and objectives, being used to develop activities, which in turn are intended to lead to a set of results and impacts, forms the theory of change behind PA 1 of the ESF programme.

A range of results are anticipated from PA 1 activity, including:

- improved employability amongst unemployed and inactive participants;
- increased labour market competitiveness amongst those supported;
- improved basic skills;
- reduced inequality in labour market access for particular groups, including women;
- participants gaining sustained employment;
- greater access to apprenticeships and traineeships;
- (re-) engagement with education and training; increased labour market competitiveness amongst young people through work experience and pre-employment training;
- improved labour market participation opportunities for young lone parents, individuals in deprived areas, and those distant from the labour market and/or facing multiple barriers to work moving closer to or into work;
- improved employability of groups facing significant barriers to work such as offenders; and
- improved pathways to work, education and training for marginalised individuals and those living in marginalised communities.

A number of the above results are quantifiable, as reflected in the OP's setting of outcome and result targets (see the 'quantitative results' column in figure 2.1). These targets focus on ESF priority groups, including inactive, unemployed, and long term unemployed individuals, NEET young people, and those with childcare needs. The targets cover employment outcomes, including self-employment, and sustained employment results; childcare support; the gaining of basic skills; access to education, training and apprenticeships; and the completion of specific provision, in the case of the Youth Employment Initiative (YEI). The measurement of these targets is undertaken through a combination of the Management Information (MI) collated by the MA and a separate piece of primary survey research, the 'ESF Leavers Survey'.

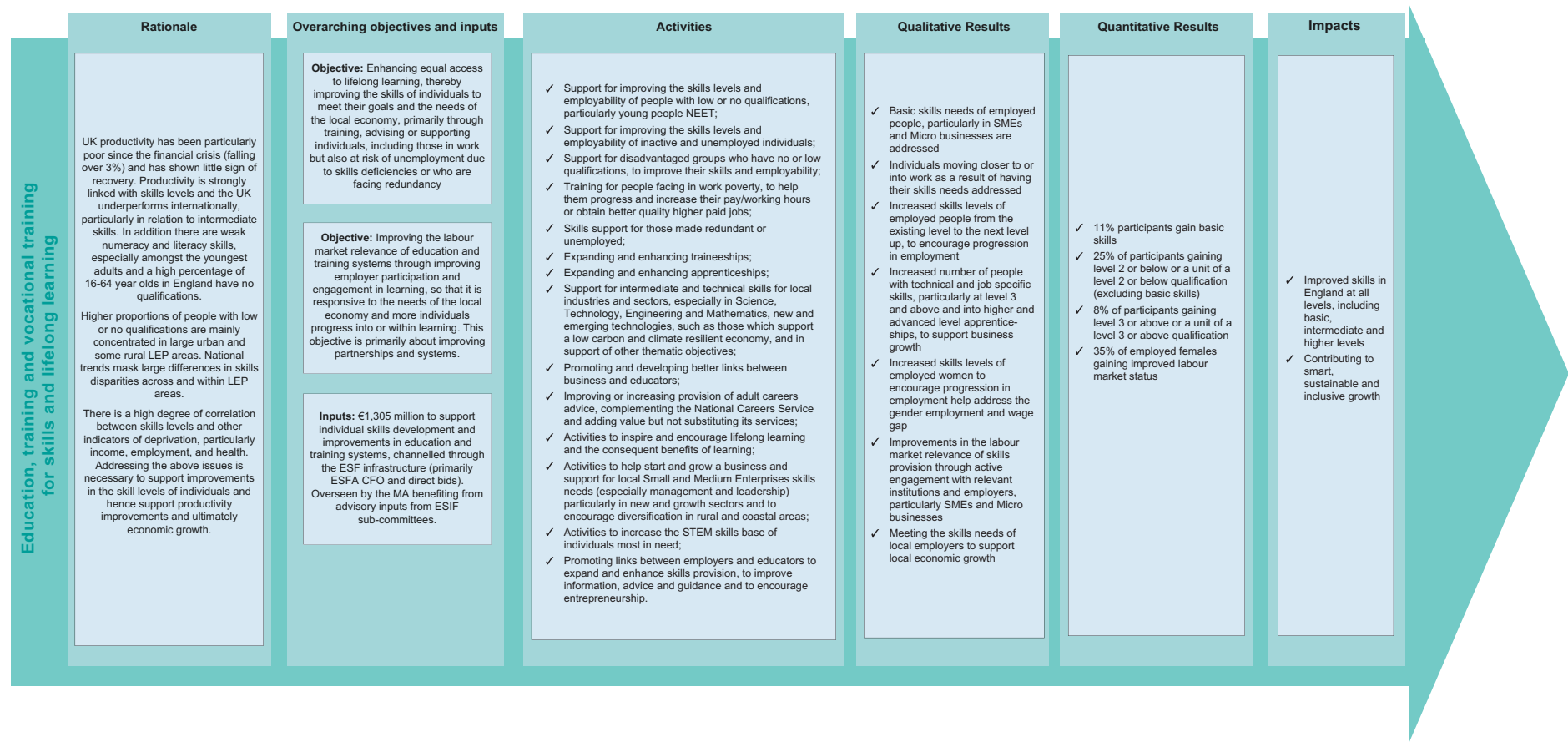
The final column of figure 2.1 details the impacts that can be expected to flow from the activities and results outlined. In respect of employment (covering IP 1.1 within the PA), impacts derived from the OP include a higher employment rate, fewer employment barriers for individuals from disadvantaged groups, and a reduction in the gender gap in terms of employment access. In terms of youth unemployment the intended impacts, from IPs 1.2 and 1.3, relate to an increase in young people in education, employment and training, and a concomitant reduction in those who are NEET, or at risk of being so, including a reduction in youth unemployment. Activities addressing social inclusion through IP 1.4 are intended to contribute to addressing the root causes of poverty that act as barriers to work and move individuals closer to – and into – work. IP 1.5, involving CLLD, activity aims to engender more 'place-based' impacts, in terms of helping promote sustained bottom-up regeneration and economic development in marginalised communities.

2.3.2 Intervention logic for PA 2

The rationale for intervention under PA 2 relates to a perceived need to invest in skills development to address the UK's poor comparative productivity levels since the recession of the late 2000s. As figure 2.2 below highlights, the assumption is that supporting the development of basic, intermediate and higher level skills amongst individuals will help boost productivity and, ultimately, economic growth. Connected to this is the recognition that poor skill levels are linked to lower income levels, poorer employment outcomes and other aspects of deprivation such as poor health. To address these issues, PA 2 intends to support 'skills for growth' through meeting its objectives focused on developing the skills of individuals and developing the wider 'skills infrastructure', including education and training systems.

In respect of individuals, the objective of PA 2 is to enhance equal access to lifelong learning, thereby improving skills so individuals can meet their goals, with this process also supporting growth in local economies. PA 2 also has the objective of supporting structural change by improving the labour market relevance of education and training systems, in part through improving the engagement and participation of employers in learning so that these systems are more responsive to local economic needs, and more individuals progress in learning. Importantly, as the OP reflects, as well as supporting skills objectives, PA 2 also aims to contribute to the objectives of PA 1 around employment and social inclusion.

Figure 2.2 Intervention logic for Priority Axis 2



ESF impact evaluation: research design and scoping study

As the intervention logic in figure 2.2 shows, several inputs support the development and delivery of activities to address the rationale and objectives outlined. These inputs include the allocated funding from the ESF, alongside the role of the MA in managing provision and the Growth Programme Board in overseeing the programme. Further inputs linked to the ESF delivery infrastructure include co-financing mechanisms established to support delivery, in respect of PA 2 primarily through the Education and Skills Funding Agency (ESFA) CFO, and strategic inputs into programme design through Local Enterprise Partnerships (LEPs) and the role of the local ESIF sub-committees. As with PA 1, the expertise of providers offering support to participants can also be viewed as an input.

In line with the intervention logic outlined in the above diagram, these inputs support the delivery of activities covering:

- support to improve individuals' skill levels and employability amongst the unemployed, inactive, NEET young people, those made redundant, and those from groups facing labour market disadvantage such as individuals with low skill levels;
- training to support in-work progression amongst those facing in-work poverty;
- activity to expand and enhance traineeships and apprenticeships;
- support for skills development in local industries and sectors, including in the science, technology, engineering and mathematics (STEM) fields;
- activities supporting the development of business – education links; improving or increasing careers advice;
- activities to promote lifelong learning; activities to support new business starts, support local small and medium enterprises (SMEs), and help diversify the economy in rural and coastal areas; and
- activities to support the STEM skills base.

This wide range of activities and provision is, in turn, intended to achieve a number of results articulated in the ESF OP. Intended results for individuals include: addressing their skills needs and, for the unemployed and inactive, moving them closer to or into employment by doing so; helping to encourage in-work progression through improving skill levels; and improving individuals' technical and job-specific skill levels, particularly higher level skills. This latter result in particular is also intended to support business growth through the development of the skills base and increasing access into higher and advanced level apprenticeships. A crossover between results for individuals and structural outcomes is also evident in the anticipated result concerning improved skill levels amongst employed women, helping to address the gender employment and wage gap. Further anticipated results of a more structural nature include improved labour market relevance in terms of available skills provision and meeting local employers' needs to support local economic growth.

As with PA 1, the above results articulated in qualitative terms are complemented by the intended achievement of a set of more directly measurable quantitative results. Reflecting the skills focus of PA 2, these cover the achievement of: basic skills, Level 2 or below qualifications¹⁷ and units of qualifications, Level 3 and above qualifications or units of qualifications, along with a result target for the percentage of females in work achieving improved labour market status.

¹⁷ In England, qualifications are assigned to levels under the Regulated Qualifications Framework (RGF), with each level having criteria indicating the required competencies within it.

As figure 2.2 shows, the qualitatively and quantitatively articulated results of PA 2 are intended to lead to impacts, namely improved skills in England at all levels, including basic, intermediate and higher levels, and contributing to smart, sustainable and inclusive growth. These impacts form the end of the intervention logic chain, or theory of change, reflecting the ultimate objectives of PA 2.

2.3.3 Additional insights from the theory of change workshop

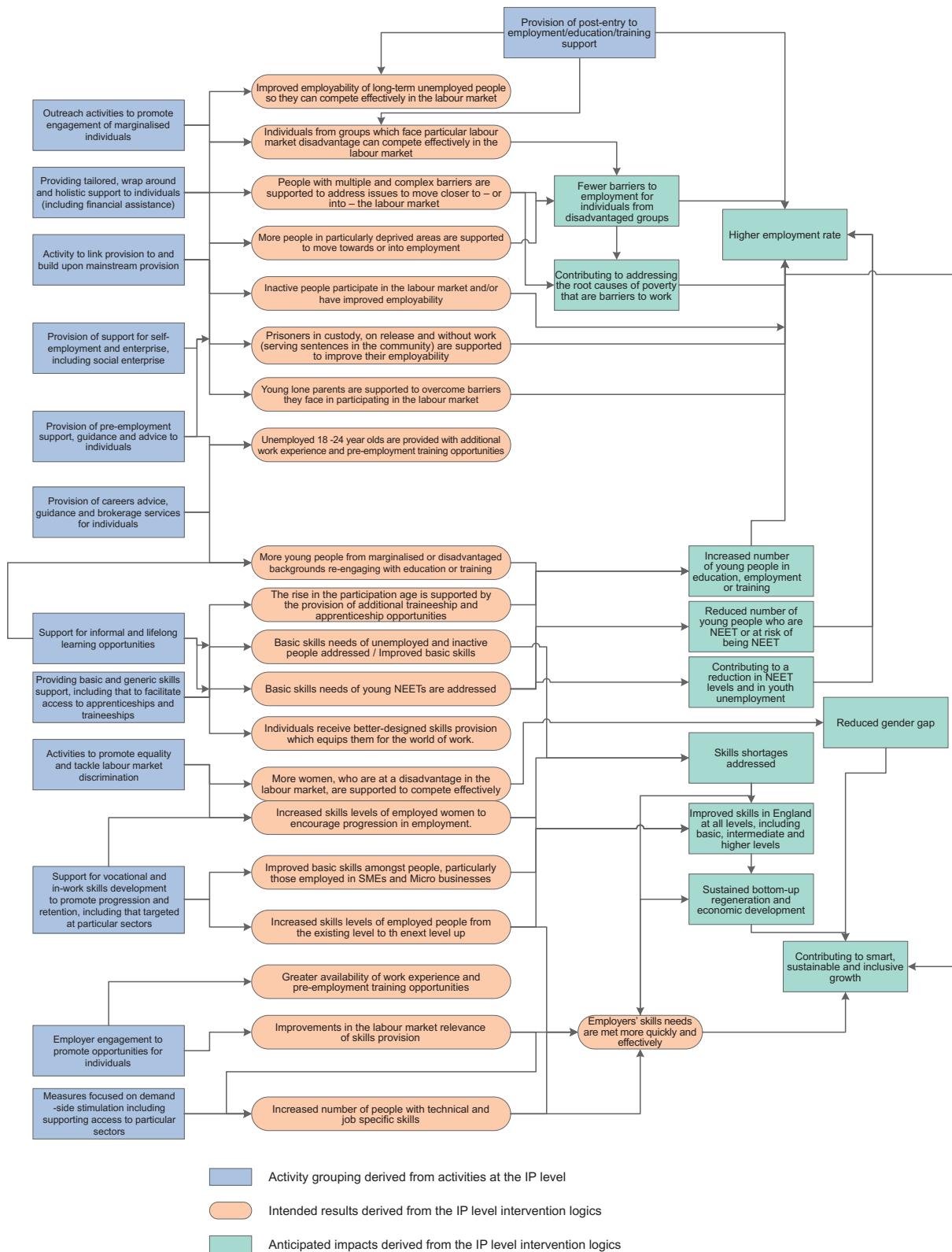
In addition to developing the intervention logic for the key components of the ESF programme, the study also involved bringing stakeholders together for a 'theory of change' workshop. This had two principal objectives. The first was to explore the likely contextual factors that would influence the extent the intervention logic, or theory of change, might be observable in practice. The results of this exercise were integrated with the other stakeholder inputs discussed in section 2.2 above. The second sought to inform the impact evaluation further through developing a more sophisticated understanding of the likely causal chains linking inputs to activities and on to results and impacts.

To facilitate the exercise involved in this second objective, activities drawn from the intervention logics developed at the PA and IP levels¹⁸ were grouped together into a series of 'activity groups'. Results and impacts drawn from these intervention logics were also provided so that, working in groups, participants could explore which activity groups or types have the potential to lead to which results and impacts. Figure 2.3 overleaf draws on the results of this exercise to present a presumed programme level causality linking activities, results and impacts. Both the IP level intervention logics and the component activities within each activity grouping are included in Annex one for reference. The IP level intervention logics also provide the source for the presumed results and impacts detailed in the diagram overleaf.

The approach of developing activity groupings and tracing causality has two main potential uses in informing the planned impact evaluation. Firstly, it offers a more sophisticated understanding of the intervention logic that the impact evaluation might seek to test. Secondly, it offers a different way of grouping activities, other than at the PA and IP level, which might facilitate the development of hypotheses concerning the ESF's presumed effects. Such hypotheses would be required for a theory-based evaluation approach, or could form a component of a wider impact evaluation approach. In forming part of such an approach, qualitative research – potentially in the form of case studies – could be used to explore linkages between the delivery and results of different forms of provision, and different activities undertaken by providers, in detail. The potential utility and role of this understanding of causality is picked up further in subsequent chapters in this report.

¹⁸ The intervention logics developed at the individual IP levels are included at Annex one for reference.

Figure 2.3 Activity groupings across ESF IPs and their causal links with the intended results and impacts of the programme



2.3.4 The role of the intervention logic in the proposed impact evaluation

The understanding of the ESF intervention logic at PA and IP levels presented above provides a key underpinning for the proposed impact evaluation. The process involved in developing these intervention logics is intended, in particular, to identify the qualitative and quantitative results of the programme against which to assess the impact of the ESF in England. In addition, as noted, the tracing of causality and the development of activity groupings opens up further potential methodological approaches that might be used in the proposed ESF evaluation. This includes the potential, for example, to adopt a theory-based evaluation approach to assess impacts where counterfactual impact evaluation (CIE) approaches are infeasible. In addition, such an approach might also be used to explain and/or confirm the results of a CIE.

As well as assessing impacts, in line with the EC ESF guidance,¹⁹ the evaluation will need to assess the extent to which the objectives of PA 1 and PA 2, as captured in the intervention logics described, have been met. The intervention logics and theory of change articulated therefore form an important basis for further developing the evaluation approach and subsequently implementing it. Chapters three and four of this report examine how this might be achieved, whether through a CIE and/or alternative approaches.

¹⁹ *Programming period 2014-2020: Monitoring and Evaluation of European Cohesion Policy, European Social Fund Guidance document*, European Commission, 2015

3 Key design considerations

This chapter examines a series of what can be termed ‘key design considerations’ for the proposed ESF impact evaluation. It first explores issues around the scope and focus of the evaluation. The level at which evaluation activity might be undertaken is then considered, for example at the whole programme level, Priority Axis (PA) level, or some other level or levels. The chapter then considers the evaluation criteria that an assessment of the ESF programme should include, in particular to address the study questions around whether effectiveness of processes should be considered alongside impact, and, if so, how. Possible approaches to assessing value for money (VfM) in the context of the evaluation are then examined.

3.1 Scope and focus of the evaluation

The key issues for determining the scope and focus of the proposed ESF impact evaluation relate to coverage, in the sense of which elements of the programme should be in scope for evaluation, and focus, in terms of which aspects of ‘impact’ the evaluation should focus on.

3.1.1 Coverage

As highlighted in the previous chapter, all of the activities and Investment Priorities (IPs) within PA 1 and PA 2 are intended to lead to results or impacts. As such, there is a strong a priori case that all aspects of these priority axes should be in scope to offer a comprehensive assessment of the intended results of the programme. This also reflects the requirements set out in the European Commission (EC) guidance on ESF impact evaluation, which call for an assessment of the extent to which the objectives of priority axes within Operational Programmes (OPs) have been met.²⁰ As outlined in chapter two, in summary these objectives cover:

- Supporting access and entry to employment, including supporting the sustainable entry of young people into the labour market (PA 1)
- Promoting social inclusion and combatting poverty by moving individuals closer and into work and mobilising action to support pathways to employment in marginalised communities (PA 1)

²⁰ *Programming period 2014-2020: Monitoring and Evaluation of European Cohesion Policy, European Social Fund Guidance document*, European Commission, 2015

ESF impact evaluation: research design and scoping study

- Enhancing equal access to lifelong learning, thereby improving the skills of individuals to meet their goals and the needs of the local economy, along with improving the labour market relevance of education and training systems through improving employer participation and engagement (PA 2).

Examining the extent to which these objectives have been met involves considering how far, and potentially in what ways, the results and impacts required to achieve these objectives are apparent from ESF investment and activity. The case for including all of PA 1 and PA 2 in the proposed impact evaluation is therefore clear.

However, as described in the report introduction, the OP also includes PA 3, aimed at providing technical assistance to achieve the objectives of PA 1 and PA 2 and the programme as a whole. As noted in section 2.3, PA 3 is theoretically less relevant for the type of impact evaluation that this design and scoping study aims to inform, given its concern with processes to ensure that results and impacts are achieved, rather than directly generating those results and impacts. Reflecting this, the specification for this design study did not anticipate that PA 3 would be in scope, or at least the main activities involved in the study, such as the development of intervention logics, should not be a focus for PA 3 as they were for PAs 1 and 2. While there is a case for including PA 3 in evaluative activity around the ESF programme, it should arguably be treated as part of any process evaluation activity therefore, rather than within an ESF impact evaluation.

Nonetheless, the EC guidance does imply a need to evaluate how far the objectives of each PA have been met. If, to meet this requirement to “assess to what extent the objectives under **each of the priority axes** for the ESF 2014-20 programme have been achieved” necessitates including PA 3 in the scope of evaluative work, there are two options. Firstly, a separate process evaluation could be commissioned, tightly focused on assessing the effectiveness of technical assistance activity. Secondly, within the overall planned ESF impact evaluation, a theory-based approach could be taken, a small part of which could be used to assess how PA 3 has influenced, or otherwise, the achievement of the objectives of PA 1 and PA 2. Such an approach is accepted as an alternative, or adjunct to, counterfactual impact evaluation (CIE) in the Commission’s guidance. In addition, as outlined in the sections that follow, there are reasons to assume that such a theory-based approach may be required for assessing aspects of the results and impacts of PA 1 and PA 2; hence, some assessment of PA 3 may be undertaken as part of this.

3.1.2 Focus of the impact evaluation

A further issue for the ESF impact evaluation concerns the results and impacts what it should focus on. As noted in discussing the evaluation scope, there is a strong ‘a priori’ case for examining all of the intended results and impacts of PA 1 and PA 2. However, as the previous chapters should make clear, there are a large number of results and impacts anticipated across the programme as a whole. This raises a potential need to prioritise resources to focus on certain impacts more than others, particularly in light of the importance of any evaluation being manageable and proportionate in its scale. Equally, considering the evaluation focus highlights the need to explore which results and impacts are amenable to a CIE and which may require alternative methods.

This latter issue is the focus of chapters four and five. However, it is worth noting a couple of key considerations concerning the relative significance of types of impact. Firstly, reviewing the intervention logics of the programme and examining a sample of the ESF projects commissioned illustrates that the most common, and arguably significant, results the programme aims to achieve are around employment and skills development at the level of participating individuals. Available data on expenditure similarly indicates the prevalence of projects aimed at supporting individuals into work and developing skills. This clearly indicates that employment and skills impacts on participating individuals should be a, if not the, primary focus of an ESF impact evaluation.

Secondly, some impacts are necessarily likely to be more feasible, or easier, to assess than others. In particular, more wide-ranging and multi-faceted structural impacts, such as on education and training systems, are commonly harder to robustly estimate or assess than more specific impacts on, for example, individuals. This does not mean the former should not be addressed in an impact evaluation but, rather, that expectations for the precision with which such impacts can be measured should be realistic. Similarly, such a recognition also has implications when examining which methods to use for particular types of impact. Both of the above considerations influence the further detailed consideration of which impacts should be assessed, and how, presented in chapters four and five.

3.2 Determining the level of analysis

Determining the level of analysis at which an ESF impact evaluation should be conducted is a further aspect to determining the evaluation focus. As reflected in the questions set in the initial study specification, a key issue is to determine what the unit of analysis should be. As described in section 2.1, this issue stems in part from the relatively complex design and delivery architecture of the ESF programme. As outlined in the preceding section, there is a strong case for impacts on individual participants being a significant focus for the evaluation in terms of a 'unit of analysis' where applicable (that is, where intended results are articulated in relation to programme participants as opposed to, for example, structural impacts on training systems). However, there is a further question of what level or levels such impacts should be estimated at, or aggregated to, in terms of the different levels of the ESF architecture outlined in section 2.1.

In particular, the evaluation could be focused at the 'whole-programme' level; at the Thematic Objective (TO), PA or IP levels; at the level of particular co-financed provision; such as those run by the Department for Work and Pensions (DWP) or Education and Skills Funding Agency (ESFA); or some combination of the above. In addition, the geographical aspect to the programme design, covering developed, transition and more developed regions, offers a further way of focusing the evaluation.

For simplicity, it may be tempting to focus activity at the level of particular co-financed provision, groups of similar projects, or at the IP level, and then aggregate results of stand-alone evaluation activity into a form of meta-evaluation of the programme. However, this is potentially problematic if not properly thought through. In particular, the EC guidance anticipates that judgements should be made at least at the level to which PAs have met their objectives or not. Therefore, any approach looking to

ESF impact evaluation: research design and scoping study

'build from the ground up' must consider how these judgements can be made when looking from 'the top down', or from the PA level down. They would also need a well-developed articulation of how individual evaluation components can be combined, and made sufficiently comparable, to make PA level judgements feasible.

In addition, there is a strong presumption in the EC guidance, reflected in the study specification, that the proposed impact evaluation should provide an assessment and, where possible, estimation of the totality of impacts generated. This implies a need to develop an analytical approach able to estimate and assess impacts for participants, but also potentially broader outcomes, at least at the PA level, and ideally across the whole programme. If we also consider the point made in the previous sub-section concerning the main impacts of the programme relating to employment and skills development for individuals, with both being relevant across the IPs and, to a large extent, the PAs as well, this again points towards the need to consider impacts at the whole programme level.

Importantly, however, while the focus of the evaluation should arguably be at this whole-programme level, this does not mean that analysis and findings are unable to be disaggregated at lower levels, in particular at the PA and IP levels or in respect of particular co-financed programmes such as the Big Lottery's Building Better Opportunities programme. In principle, it should be possible to develop an evaluation approach that can facilitate an analysis of impacts on participating individuals at different levels (e.g. PA, IP), while also assessing impact in respect of the ESF as a whole. While impacts on individual participants are likely to be a significant focus of any CIE, for example, estimations can be made in terms of a treatment effect on such individuals by disaggregating the results of such analysis at different levels – e.g. whole programme, PA, IP etc. We return to this in the subsequent chapter on the potential for a CIE as part of an ESF evaluation.

Equally, intended results beyond those for individuals, such as on training systems, could also be assessed through such an approach, though as later chapters discuss this is likely to be through methods other than a CIE (for example, through qualitative data collection within a theory-based evaluation approach, further details of which are provided in chapter five). Importantly, such an approach implies the development of a well thought through analytical framework able to guide the focus of data collection and analysis, both qualitative and quantitative, within the impact evaluation.

Such a framework might usefully build on the IP and PA level intervention logics outlined in this report and, in particular, on the cross PA/IP consideration of activity groupings and causality discussed in section 2.3.3 and summarised in figure 2.3. As noted, looking at the programme from this perspective offers the potential to develop a theory-based approach within which to frame the evaluation. Using qualitative insights gained from, for example, case study fieldwork, this approach could facilitate the development and testing of hypotheses around the degree to which particular types of activity lead to specific assumed results, while exploring how and why this was the case. Such an approach could complement a CIE focused on estimating the impact of the programme around its key intended results – as noted, those linked to employment and skills development for individuals. The potential shape of such a combined approach is outlined in more detail in chapter five.

Finally, in terms of the potential to examine the programme from a geographical perspective according to categories of region, there is a strong argument that such an approach would be problematic and unlikely to add much insight. The key reason

for this is the fact that, as the OP explicitly recognises, differences within these categories of region are as important as differences between them. Focusing analysis at this level would thus be of limited utility, particularly given the mixture of more and less deprived areas within the categories concerned. Using these broad regions to frame the analysis is therefore not recommended.

3.3 Evaluation criteria: effectiveness and impact

As reflected in the study specification, part of the design considerations for an ESF impact evaluation relate to the extent effectiveness should be considered alongside impacts as a key criterion for the proposed ESF evaluation to incorporate. In this context, drawing on definitions used by the EC in its evaluation work, effectiveness involves assessing "... the progress made towards achieving the objectives of the intervention, looking for evidence of **why, whether** or **how** these changes are linked to the ... intervention."²¹ This 'why, whether and how' aspect implies an exploration of the reasons that impacts, and evidence of objectives being met, are evident or otherwise, as well as the scale of those impacts. In turn, this raises questions of the degree to which the impact evaluation should consider the processes that lead to objectives being met and intended impacts achieved.

A strong argument for including effectiveness considerations in the impact evaluation concerns its anticipated role in informing any successor domestic programme to the ESF, or indeed related non-ESF interventions. Without some understanding of which activities lead to which results and impacts, along with how and why, the evidence base for future interventions would be significantly lessened. More generally, exploring these 'how' and 'why' issues is widely seen as good practice when conducting impact evaluations. For example, as the Organisation for Cooperation and Economic Development's (OECD) principles of impact evaluation state:

"A well-designed impact evaluation can also answer questions about program design: which bits work and which bits don't, and so provide policy-relevant information for redesign and the design of future programs. We want to know why and how a program works, not just if it does."²²

While effectiveness implies some consideration of processes in terms of exploring why and how impacts occurred, it does not necessarily require a full process evaluation. In the ESF context, such an evaluation would ordinarily involve full consideration of, for example, questions around the design, implementation and 'strategic fit' of the programme with domestic activity. A number of these questions have already been examined in, for instance, the ex-ante evaluation of the current ESF programme. In contrast, the expectation for the planned ESF evaluation is clearly that its primary focus should be on impact, as opposed to processes.

²¹ European Commission (2017), Better Regulation Guidelines, Chapter VI: Guidelines on evaluation (including fitness checks), p.59

²² OECD, Outline of Principles of Impact Evaluation, available at: <http://www.oecd.org/dac/evaluation/dcdndep/37671602.pdf> (accessed 30.11.17)

It can therefore be suggested that the evaluation should only consider process elements in the context of examining effectiveness where their likely effects on the intervention logics developed are significant. In other words, where gaining an understanding of the functioning of key delivery processes could play a significant explanatory role in explaining why impacts occurred or did not. This is precisely where the criteria of effectiveness is of relevance, leading as it does to a focus on identifying the degree to which key aspects of the implementation of the ESF were effective or otherwise.

As noted previously, the development of an understanding of presumed causality in respect of the programme, and the potential to develop and test hypotheses concerning which activities should lead to which results, is one likely mechanism through which effectiveness could be explored. The potential for a theory-based element to the impact evaluation is significant here, offering as it would a framework through which to explore how and why inputs and activities were effective and their role in contributing, or otherwise, to results and impacts.

Such an approach could also link the assessment of effectiveness to the estimations of impact gained through CIE methodologies. This would provide explanatory insights, building on the potential contextual influencing factors and causal links identified, to offer a more holistic understanding of the totality of the programme's impacts. It is thus recommended that effectiveness is a criterion for the proposed evaluation, but that its focus should be strongly linked to explanatory considerations for why results and impacts occurred or otherwise, rather than on examining broader process considerations.

3.4 Considering value for money

A further aspect to the design and scoping study involved considering how VfM might be examined in the context of the proposed impact evaluation. Examination of the potential role of VfM assessment needs to start by considering its purpose. Again, the role of the evaluation in potentially informing other provision is an important consideration, as is the wider importance of ensuring that the results of interventions using public monies are justified by their cost. Both considerations imply that V-f-M should play a significant role in the ESF evaluation, and that assessment of it should offer as detailed an understanding of the costs and benefits of the programme as possible.

The question then becomes what approaches and data are available, or could be gathered, to inform an assessment of VfM. Typically, such assessments adopt a cost-benefit analysis (CBA) approach in cases where the respective costs and benefits of an intervention can be (largely) quantified in monetary terms. Alternative approaches such as cost-effectiveness analysis (CEA) are available for assessing benefits that, for a number of possible reasons, are not amenable to, or fully suitable for, monetisation. CEA can be used either as a replacement for, or adjunct to, CBA.²³

²³ For a further discussion of the methodological basis, history and potential uses of CEA see Levin, H and McEwan, P (2000), *Cost-Effectiveness Analysis: Methods and Applications*, 2nd Ed., Sage

ESF impact evaluation: research design and scoping study

Drawing on the desk research and stakeholder consultations undertaken to inform the study, our understanding is that information on the direct costs²⁴ of the ESF intervention should be available, or can potentially be collated, at a number of levels. These include at the whole programme level, PA level, IP level, for particular national co-financed programmes such as those run by DWP and Big Lottery, and for individual projects. This suggests that data for the direct cost element of CBA or CEA approaches should be able to be accessed and analysed within an ESF impact evaluation.

While identifying the direct costs of the programme should be relatively straightforward, in the sense of these equating to the funding allocated, a full understanding of VfM should ideally take account of indirect or hidden costs. These may be in the form of, for example, cross-subsidy of provision by providers or 'in-kind' contributions. The extent that such costs can be systematically identified and consistently monetised across the programme is likely to be limited. Therefore, their potential use in CBA or CEA models is questionable. Nonetheless, any fieldwork could be used to identify the type, prevalence and broad scale of such indirect costs by way of contextualising any CBA or CEA undertaken.

On the benefits side of the equation, a key advantage to situating an assessment of VfM within an impact evaluation is the potential this offers for estimating the net, as opposed to gross, benefits of the intervention, at least in respect of some of its anticipated impacts. The results of any CIE approach taken within the evaluation, presuming this is possible and successfully implemented, could be used to identify the additionality of certain benefits that are evident, through its role in estimating benefits that can directly be attributed to the programme. Importantly, however, this would only be feasible in respect of impacts that are estimated through such approaches. As noted earlier in the report, the range of anticipated results and impacts from ESF activity is considerable, and not all may be amenable to a CIE, as explored further in chapters four and five.

The question of what potential benefits could be monetised to facilitate a CBA is also linked to this, in that certain impacts, for example entering employment or developing skills, can be given a monetary value. Our understanding is that the DWP CBA model should be able to be used as the starting point for a CBA in respect of the programme, and its component elements, where it is feasible to monetise both costs and benefits. Where available, we also understand that there may be potential to use CBA frameworks or approaches developed by other government departments where appropriate and/or required in the context of likely benefits accruing from the ESF programme. For example, this might include estimating the returns resulting from skills development, or from activities to reduce re-offending.

If there are existing standard frameworks or approaches in these areas, for example from related work undertaken by the Department for Education and Ministry of Justice/Her Majesty's Prison and Probation Service, this may provide a good starting point and enable calculations to be consistent with approaches developed by other government departments. However, it should be acknowledged that combining approaches with the DWP CBA model is likely to be challenging due to likely differences in what benefits are valued and how. As such, further work would be needed to establish the utility and feasibility of approaches used by

²⁴ That is, the funding allocated and/or used.

other departments, Equally, if feasible, caution would still need to be applied, and appropriate caveats used, to account for differences in how these frameworks and approaches used by other departments vary from those used by DWP.

In our understanding, using the DWP CBA framework as a basis for assessing the costs and benefits that can be ascribed to the ESF programme, and by extension the VfM it offers, should enable a range of relevant benefits to be quantified. In the context of the ESF, these include, for example, most of the fiscal and social impacts of the programme as they relate to employment, including those concerning, for instance, changes in income, benefits payments and tax receipts.²⁵ The approach to a CIE advocated in later chapters should offer the potential to gather data on, for example, participant characteristics such as benefit history so as to enhance the accuracy of the estimates concerning such benefits.

Taking the distinction made in the Department's approach between primary and secondary market impacts into account,²⁶ our understanding is that use of the DWP model also has the potential to enable estimates to be made concerning other relevant impacts, including those relating to health for example. As the Department's working paper relating to the model outlines,²⁷ should a relatively comprehensive CBA of the programme's effects be possible, it is also likely to be important to consider 'equilibrium effects' in the assessment (defined as effects on the general equilibrium of the economy as a whole). Thus, where possible, a full CBA should include an estimation of such effects, for instance in respect of substitution, displacement or general wage equilibrium effects.²⁸

Further consideration concerning the exact parameters of a CBA approach in this context will be required when refining the design and implementation of the planned ESF evaluation. However, the above suggests that CBA may well be feasible for considering VfM in respect of significant parts of the programme's likely costs and benefits as they relate to employment outcomes. This aspect of VfM should be possible to estimate, furthermore, in respect of both participants and wider society, including taxpayers. Allied to the careful use of the DWP model in conjunction with other departmental CBA frameworks if available and as appropriate, other key likely effects of the programme, such as those relating to re-offending and skills development, should have the potential to be monetised and assessed within a CBA approach.

As noted, where other results of the programme are considered significant to incorporate into a VfM assessment, potentially those relating to wellbeing for example, additional approaches such as CEA can be considered. It should be recognised,

²⁵ Further details on the Department's approach to CBA in this context can be found in Fujiwara, D. (2010), *The Department for Work and Pensions Social Cost-Benefit Analysis Framework: Methodologies for estimating and incorporating the wider social and economic impacts of work in Cost-Benefit Analysis of employment programmes*, DWP Working Paper No. 86

²⁶ Ibid.

²⁷ Ibid.

²⁸ A discussion of these effects, including supply and demand side substitution effects, displacement effects and equilibrium wage effects, is included at p.20 and p.21 of Fujiwara, D. (2010), *The Department for Work and Pensions Social Cost-Benefit Analysis Framework: Methodologies for estimating and incorporating the wider social and economic impacts of work in Cost-Benefit Analysis of employment programmes*, DWP Working Paper No. 86.

ESF impact evaluation: research design and scoping study

however, that while the potential CBA approach outlined can be used to derive a VfM ratio or ratios, this would not be able to be directly combined with equivalent ratios derived from (potential) CEA approaches for effects that are unable to be monetised.²⁹

As alluded to above, there are also decisions that need to be taken in designing the approach to assessing VfM concerning the relative amount of resources available for this aspect of the evaluation. Equally, considerations around the potential research burden that could result from, for example, additional research and data collection to refine the VfM assessment, will need to be worked through. For instance, primary data collection at the provider level could be envisaged, through surveys or other methods. This could help determine a more complete picture of costs and also capture wider evidence around benefits that may not be available from data collected routinely as part of ESF programming. However, such an approach would need careful consideration given the significant levels of data capture already required from providers. Equally, the need for the VfM component to be proportional to the overall evaluation, in light of its relative importance, may also preclude such approaches.

²⁹ In particular, while CBA enables the monetisation of both costs and benefits to produce ratios, CEA would monetise costs but utilise other non-monetary units for benefits. Hence the approaches can not be combined in this sense and are, equally, not directly comparable.

4 Potential counterfactual impact evaluation approaches

This chapter examines the potential for including a counterfactual impact evaluation (CIE) approach or approaches within a European Social Fund (ESF) impact evaluation. It first conducts a high-level initial feasibility assessment of the potential for such approaches, reflecting on the programme's intended results and impacts identified in preceding chapters. An initial consideration of possible CIE designs is then undertaken. Key considerations relating to CIE designs are then examined, specifically in respect of how ESF participants are assigned to provision, the availability of data and likely comparison groups to facilitate a CIE approach, and issues around required sample sizes to estimate impacts reliably.

4.1 CIE – initial feasibility assessment

4.1.1 Defining CIE in the context of ESF impact evaluation

Before initially assessing the feasibility of conducting a CIE, it is important to define what we mean by this term in the context of an impact evaluation of the ESF. A CIE aims to compare the outcomes of an intervention with the outcomes that would have been achieved had the intervention not been in place. Impact evaluations using CIE methods thus aim to determine the causal effect of a specific programme, policy or intervention on an observed outcome, in this case the intended results and impacts of the ESF programme. Approaches of this type aim to answer two questions:

- Has a change occurred?
- Was the policy or intervention in question responsible for this change?³⁰

CIE approaches rely on defining a 'counterfactual' through which to test whether an intervention has a presumed effect, typically by comparing outcomes for the subjects of an intervention with a similar comparator, 'comparison group', or 'control group' not subject to that intervention. In the context of ESF, results are often measured, for example outcomes for programme participants, after programme participation. However, these outcomes can be affected by factors other than the programme itself. We might observe positive results driven by the selection of particular types of people into the programme, or we might observe positive outcomes that have nothing to do

³⁰ Adapted from HM Treasury (2011), *The Magenta Book: Guidance for Evaluation*, p.98

with programme participation, but are mainly driven by contextual factors. These can include, for example, an improving labour market. Impact evaluations aim to exclude such alternative explanations where possible to identify net effects of interventions.³¹

This exclusion of alternative factors or explanations is the role of the counterfactual. The difference between outcomes observed on the subjects of an intervention, compared to outcomes in this counterfactual scenario, is assumed to represent the real or net effect of that intervention. Different research designs use a variety of techniques to construct a counterfactual. In CIE approaches relying on the development of a 'comparison' or 'control' group, the aim is to choose participant (treatment) and comparison (control)³² groups that are as similar as possible with regard to all other factors: that is, factors relevant for the outcome such as socio-economic background or motivation.

It is also worth briefly examining what the counterfactual represents in this context. The aim of such an approach is to infer the counterfactual outcome for ESF participants; in other words, to estimate what would have happened to this group of individuals had they not received support. Typically, construction of a comparison group in contexts such as this aims to uncover this effect in terms of the 'average effect of treatment on the treated' (ATT), as was the case with a previous ESF impact evaluation undertaken by DWP in the 2007-2013 programming period.³³ It is important to understand, therefore, that following such an approach, as is typical for employment programme impact evaluations, does not involve comparing outcomes for ESF participants with outcomes for participants benefitting from other employment programmes, and defining these latter outcomes as a 'counterfactual'. Rather, it focuses specifically on estimating the impact on those individuals receiving support through the ESF in terms of what would happen without this support.

Approaches that seek to compare the impact of ESF support to, for example, DWP's Work and Health Programme are possible to develop in theory. However, it would be preferable in methodological terms to estimate the ATT of these programmes separately using a very similar, or the same approach, as for the ESF, and then compare the results in respect of an identical outcome (for example, employment entry). However, it should also be noted that this raises a potential issue of not comparing like with like, or at least having to reflect on and account for differences that arise from, for example, the mandatory or voluntary nature of programmes, different levels or types of support offered and so on.

Having broadly defined the concept of a CIE in this context, we now turn to consider its feasibility.

³¹ EC, (2013). Design and Commissioning of Counterfactual Impact Evaluations. A Practical Guidance for ESF Managing Authorities, Luxembourg: Publication Office of the European Union: p. 5

³² The term 'control group' is typically used for randomised control trials, while the term 'comparison group' is used for quasi-experimental designs that depend on mimicking such trials through constructing a comparator by which to estimate impact.

³³ Ainsworth, P. and Marlow, S. (2011), *Early Impacts of the European Social Fund 2007-2013*, DWP In-House Research Report No.3

4.1.2 The potential for assessing ESF impacts through CIE

Not every programme, policy or intervention lends itself to conducting reliable CIEs, as outlined in the Treasury's Magenta Book guidance.³⁴ Whether, and the extent to which, adopting such an approach is possible is a case-by-case decision. This is dependent upon the design of the measure, the type of results it aims to achieve, and, crucially, on data availability. Ideally, counterfactual analysis is applied when dealing with a well-defined intervention, targeted at a well-defined population or subject of intervention, with the aim of inducing a change in well-defined state and/or behaviour of participants. Consequently, some policy areas do not easily lend themselves to such analysis: for example, macro-level policies or effects, large infrastructure and public works projects, and interventions seeking to achieve structural changes.

Examining the intended impacts of the ESF programme in England, as articulated through the development of intervention logics in chapter two, enables us to consider their potential to be assessed using CIE. As the above discussion of the factors influencing the feasibility of CIE approaches indicates, scenarios where the intended results of ESF activities, or the activities themselves, are less well-defined lead to 'a priori' concerns over feasibility. In particular, impacts around structural changes in education and training systems, or seeking to achieve results that are more conceptual, such as 'addressing the root causes of poverty', are likely to be problematic from this perspective.

The same is true of intended results of programme activity around contributing to macro-level outcomes such as an 'improved employment rate', 'reduced skills gaps', or 'sustainable and inclusive economic growth'. The key problem here is the difficulty of proving causality, given that outcomes such as these are subject to a wide range of causal factors. Realistically, therefore, the specific contribution of the ESF to these outcomes would be impossible to disaggregate from other contributory factors through a CIE. Conversely, well-defined activities that can be directly linked to ESF interventions, such as those directly targeted at individuals to improve their employability or skills, and with well-defined results, such as entry to employment or improved skill levels evidenced by qualifications, are by definition more amenable to CIE approaches.

Reviewing the key intended impacts of the ESF programme in this manner enables us to make initial judgements over which impacts might be tested, or estimated, through CIE methodologies. Table 4.1 below provides a summary overview of which key impacts could be feasibly assessed using a CIE at the level of ESF Investment Priority (IP), provided that robust data can be collected for both the treatment and a potential comparison group. It should be noted that, following the logic above concerning the importance of focusing on impacts in relation to those directly subject to the ESF intervention – namely the individuals who participate – and the a priori concerns raised in respect of structural or macro-level effects, the focus of the overview in table 4.1 is on individual level results and impacts (where possible in light of the intended results of the IP concerned).

³⁴ HM Treasury, (2012). Quality in policy impact evaluation: understanding the effects of policy from other influences (supplementary Magenta Book guidance), December 2012

ESF impact evaluation: research design and scoping study

In making this assessment, the nature of the activities concerned, their intended results, and broader issues such as the likelihood of being able to detect effects through a CIE, given the size of the ESF intervention, have informed the judgement presented.

Table 4.1 High level examination of CIE feasibility against programme impacts

Investment priority	Key impacts	Potential to be assessed through CIE in principle
IP 1.1	Individuals moving into work	Yes – employment outcomes can be assessed at an individual level; the desired change in status is clear
	Fewer barriers to employment for individuals from disadvantaged groups	Partially – employment outcomes can be assessed at the individual level and the desired change in status is clear; however, barriers to employment would need to be defined unless employment entry was taken as implicitly indicating they had been addressed and/or such barriers were clear and measurable in impact terms (e.g. gaining a qualification to address a skills barrier)
	Reduced gender gap	Unlikely – the gender gap defined more broadly is a structural and macro-issue likely to be affected by multiple policies and contextual developments, hence not lending itself to a CIE.
IP 1.2	Increased number of young people in education, employment or training (interpreted in terms of young people entering education, employment or training)	Yes – these outcomes can be assessed at an individual level, the desired change in status is clear.
	Reduced number of young people who are Not in Employment, Education or Training (NEET) or at risk of being NEET (interpreted in terms of young people entering education, employment or training)	Partially – as above for NEET status as this is a converse of being in education, employment and training; ‘at risk’ of being NEET is more definitionally unclear and involves a more heterogeneous set of indicators, making assessment through a CIE in respect of individuals likely to be problematic / challenging.
IP 1.3	Contributing to a reduction in NEET levels and in youth unemployment (interpreted in terms of young people entering education, employment or training)	Yes – as above in that employment, education and training entry can be assessed at an individual level and the desired change in status is clear.

ESF impact evaluation: research design and scoping study

Investment priority	Key impacts	Potential to be assessed through CIE in principle
IP 1.4	<p>Contributing to addressing the root causes of poverty that are barriers to work (can be interpreted as individuals moving closer to work through addressing barriers / distance travelled towards work)</p> <p>Individuals are moved closer to – or into – work</p>	<p>Unlikely – unless restricted to employment entry or skills development effects – it is unclear what is meant by ‘root causes’ of poverty; while there may be a possibility to assess such causes at an individual level, the concept is definitionally unclear and outcomes to consider would be numerous and heterogeneous. Possible to interpret at the individual level as relating to moving individuals closer to work through addressing barriers / distance travelled – the above issues re definitional clarity and wide range of outcomes to consider again make this unlikely.</p> <p>Partially – moving into work can be addressed at an individual level and the desired change in status is clear; moving closer to work through addressing barriers is too heterogeneous in terms of numbers of likely relevant affects and definitionally unclear to assess through a CIE</p>
IP 1.5	<p>Sustained bottom-up regeneration and economic development</p>	<p>Unlikely – impacts around ‘bottom-up regeneration’ are definitionally unclear and likely to be numerous and heterogeneous; ‘economic development’ impacts are similar and if defined in GDP or GVA terms are unlikely to be detectable in the context of wider economic patterns for the reasons outlined in respect of macro-level effects above.</p>
IP 2.1	<p>Improved skills in England at all levels, including basic, intermediate and higher levels (if interpreted as improved skills for individuals)</p>	<p>Yes – desired change in status is clear at the individual level if interpreted in terms of skills development for individuals.</p>
IP 2.2	<p>Contributing to smart, sustainable and inclusive growth</p>	<p>No – this is a macro-economic impact which does not lend itself to a CIE in this context, in particular because growth levels in GDP terms, or GVA per head in ‘smart’ / productivity terms, are unlikely to be detectable in wider economic patterns as a result of ESF activity, given this could only have an extremely small impact on the overall economy.</p>

ESF impact evaluation: research design and scoping study

The above analysis suggests that only certain elements of the ESF intervention logic and intended impacts should be assessed using CIE methodologies. In particular, effects on employment entry and skills development in respect of individuals appear the most amenable to such an approach. Other intended impacts derived from the intervention logic are either structural, hard to define and/or open to significant interpretation, or unlikely to be feasibly examined in terms of the likely detectability of effects or specific attribution to ESF interventions. This latter issue is significant when taking the size of the ESF investment into account in light of the wider backdrop against which the detection of impacts would be sought. As outlined, this is a particular problem for any macro-economic impacts, such as presumed effects on Gross Domestic Product (GDP) levels or Gross Value Added (GVA) figures.

At this point, it is important to note that the above assessment does not fully preclude some of the effects that could theoretically be tested through a CIE. In particular, it is possible to conceive of a way to assess the 'root causes of poverty' and 'barriers to work', depending on how these concepts are defined or unpacked. For instance, as noted in the summary table above, addressing a lack of relevant skills could be measured through a CIE if this is considered a barrier to work or root cause of poverty. Likewise, in theory, effects on other commonly cited barriers including debt problems, homelessness, or health, could be tested through CIE approaches.

However, developing an approach to these latter factors is more methodologically challenging than impacts around skills development or employment entry. Moreover, the nature of these effects implies the need to gather treatment and comparison group data through means other than the data routinely collected as part of ESF delivery, or that which would be available in administrative datasets such as the National Benefits Database (NBD). In turn, this implies a need to gather primary data, in all likelihood through a survey-based approach.

The lack of suitability of CIE approaches to test the objective of securing smart, sustainable and inclusive growth, is clearer. In part, this relates to definitional issues in terms of precisely determining what it is that a CIE would look to test. Equally, as described above, macro-level structural impacts on the economy of this type are problematic in terms of how far they can be detected in light of the size of the ESF intervention. The same issues would affect any attempt to estimate impacts in relation to 'sustained bottom-up regeneration and economic development'. In both cases, the potential to use CIE to test such impacts is also much more difficult, and arguably infeasible, than examining effects on individuals.

In summary, the above a priori assessment indicates:

- CIE approaches are likely to be most feasible for employment results and skills impacts for individuals.
- Using CIE techniques may be possible, but methodologically challenging and likely to require significant resource, in terms of impacts around reducing barriers and the causes of poverty, along with the 'risk of becoming NEET'.
- CIE approaches are likely to be infeasible, insufficiently robust and/or insufficiently sensitive to test macro-economic or structural effects relating to systems change or effects on growth and productivity.

4.1.3 The potential for assessing other ESF results through CIE

While the above section considers the feasibility of using CIE approaches to assess some of the key intended impacts of the ESF programme, it is also important to consider whether any of the more intermediate results identified through the intervention logic models could be tested in this way. In terms of PA 1, most of the intended results of activity relate to reducing barriers to employment and enhancing employability, hence moving people closer to the labour market, and access to work experience and pre-employment training opportunities.

In terms of the former, the preceding discussion on using CIE to assess impacts on reducing barriers is directly relevant and the same considerations apply; that is, while this may be possible in principle, there are definitional challenges to overcome and, methodologically, using CIE approaches in this context is challenging and likely to require significant resources. The latter is more of an output than a result or impact, in the sense of access to work experience or pre-employment training / support being something that provision seeks to achieve to encourage a result or impact, rather than being an impact in itself. As such, this would be unsuitable for testing through a CIE; rather, outputs such as this are more linked to effectiveness, in terms of exploring whether the intended activities within provision have been delivered as anticipated. Such issues are better considered through a process evaluation.

In respect of PA 2, results such as improving the labour market relevance of skills and employment training, the provision of more apprenticeship and traineeship access opportunities, and improving the design of skills provision are structural changes. As such, they are not particularly suitable for CIE approaches for the reasons outlined in the previous section (in respect of intended impacts around changing education and skills systems or macro-level impacts concerning regeneration and growth). However, PA 2 also intends to lead to results for individuals with the potential to be assessed through a CIE – some of these relate to skills development and are, in principle, feasible for the reasons outlined in relation to impacts of this type in section 4.1.2. Results concerning in-work progression, whether through promotion or wage growth, and those relating to technical and job-specific skills, are more methodologically challenging however.

The main issue for this latter set of results is three-fold. Firstly, definitional considerations are again a problem, for example in terms of 'technical and job-related skills'. Secondly, data availability can be envisaged as problematic, in particular in terms of how to capture whether individuals have benefitted from promotion and, to a lesser degree, wage growth. Thirdly, even if these challenges can be overcome, identifying a comparison group of similar individuals and gathering data on whether they achieve promotion or develop technical and job-related skills, for example, brings with it obvious challenges of identification and data collection. While logistical and methodological considerations do not preclude a CIE for results such as these in principle, they do indicate the challenges of implementing one in practice.

4.2 Review of potential CIE designs

In building on the feasibility assessment of a CIE in light of the results and impacts the ESF programme aims to achieve, it is necessary to consider which types of CIE design could be feasible given the understanding of the programme developed in preceding sections and chapters. In part, this involves similar theoretical considerations to those informing section 4.1, in that the choice of approach significantly depends on the impacts concerned and the design of the programme.

Relevant issues also include the mechanism by which individuals are referred to the programme, the likely availability of secondary data, and the extent to which the collection of primary data can reasonably be envisaged. These issues are considered in more detail in sections 4.3 to 4.5. However, their initial consideration here is intended to develop an understanding of which common CIE designs have, in a priori terms, the potential for use in an ESF impact evaluation.

In table 4.2 below, we present a summary overview of several potential research designs that might be used to estimate the counterfactual. The table also outlines requirements, risks and cost implications. We also provide an a priori assessment as to whether the designs have the potential to be used in the proposed impact evaluation.

The selection of designs reflects those generally seen as likely to be applied in the context of the ESF. In adopting this approach, we follow the guidance to ESF Managing Authorities on CIEs issued by the European Commission (EC).³⁵ Following the principal focus on Propensity Score Matching (PSM) in the guidance's discussion of potential matching designs, the table that follows reflects this. It is important to note, however, that while PSM is the most commonly used matching design in ESF and related impact evaluation, it is not the only option. In refining any matching approach for use in an ESF evaluation, then examination of alternatives such as Coarsened Exact Matching (CEM) could usefully be undertaken, balancing the relative pros and cons of such alternatives against PSM.

It should also be noted that one of the designs discussed in the guidance, the Instrumental Variables Approach (IVA), is not considered for two main reasons. First, as the guidance notes, this approach depends on assignment to the programme resulting at least partly from an 'an exogenous factor (or shock)' which is unrelated to results other than through the 'treatment' (intervention). While it is possible to conceive of this in a very specific context, for example at the level of an individual ESF project with assignment being influenced by distance from a provider, such factors can not realistically be conceived at the programme-wide level. Second, for this and other reasons, as the guidance notes, there are no examples of IVA being used in ESF impact evaluations. We therefore focus on the main likely design options.

³⁵ European Commission (2013), Design and Commissioning of Counterfactual Impact Evaluations: A Practical Guidance for ESF Managing Authorities

Table 4.2 Summary overview of potential CIE designs

Design	Description	Requirements	Risks	Cost implications	A priori feasibility
Randomised Control Trial (RCT) (Experimental design)	<p>Randomly assigns individuals to a treatment and control group</p> <p>As allocation process is random, there should be no selection bias provided there is no effect on the control group – e.g. resentment at being denied help</p> <p>Any outcome difference between treatment and control group can be attributed to the programme</p>	<p>Needs to be built into delivery from the outset</p> <p>Can be used when</p> <p>a) the eligible population is larger than the number of spaces on the programme, or b) a programme is gradually phased in</p> <p>Assignment can be enforced/ monitored</p> <p>Cross-contamination between both groups is limited</p> <p>Sample sizes are sufficient</p>	<p>Low external validity, i.e. participants in the RCT may be different from the general population</p> <p>Cross-contamination of treatment and control group</p> <p>Drop-out of study participants</p> <p>Ethical concerns of denying comparison group equal treatment/ opportunities</p>	<p>Can be high due to resources required to ensure/monitor compliance with the assignment protocols to treatment and control group</p> <p>Also dependant on the type of data used, scale of intervention and exact analytical approach</p>	<p>Not feasible, as (self-) selection into ESF-funded provision has already started and engagement with provision is voluntary</p>

ESF impact evaluation: research design and scoping study

Design	Description	Requirements	Risks	Cost implications	A priori feasibility
<p>Regression Discontinuity design (Quasi-experimental design)</p>	<p>Mimics an experiment by exploiting an existing strict assignment rule into the programme.</p> <p>When the selection into the programme is exclusively driven by this assignment rule, an ‘experiment’ around the cut-off point takes place, i.e. a person just falling short of the threshold should not differ from a person just above the threshold.</p>	<p>The selection process must be exclusively driven by a strict assignment rule</p> <p>An analysis of cross-overs (comparison group members who receive treatment) and no-shows (treatment group members who don’t receive treatment) is possible</p> <p>Sample sizes around the cut-off point are sufficient</p>	<p>Low external validity, i.e. people around the threshold might be different from the general population</p>	<p>Will depend on data availability, if data is collected regularly on the participant and comparison group, cost implications are generally low</p> <p>Cost implications are higher if primary data is collected</p>	<p>Not feasible for the programme overall, as no clear cut-off exists. May be feasible for specific target groups and interventions: e.g. those over 50, but of limited potential in the context of a ‘whole programme’ evaluation.</p>

ESF impact evaluation: research design and scoping study

Design	Description	Requirements	Risks	Cost implications	A priori feasibility
<p>Difference-in-Differences design (DiD)</p> <p>(Quasi-experimental design)</p>	<p>Limits differences between the participant and comparison group by comparing their outcomes at two points in time – before and after programme participation. Any additional outcome improvement of the participant group compared to the comparison group can be attributed to participation.</p>	<p>Outcomes for participants and the comparison group have historically developed in parallel (parallel trend assumption)</p> <p>Data for the participant and comparison group is available for at least two points in time.</p>	<p>It can be difficult to prove the parallel trend assumption.</p> <p>When primary data is collected, engagement at two points in time is needed. There is limited incentive for the comparison group to participate.</p> <p>If primary data is collected, there is a risk of dropouts.</p>	<p>As above</p>	<p>A DID design with administrative data could be feasible in principle</p> <p>If primary data collection is involved, a collection of data for the full cohort of ESF participants is infeasible, as delivery has already started. It could be used for supplementary analysis on specific outcomes.</p>

ESF impact evaluation: research design and scoping study

Design	Description	Requirements	Risks	Cost implications	A priori feasibility
<p>Propensity Score Matching design (PSM) (Quasi-experimental design)</p>	<p>Identifies treatment and comparison groups by taking into account background characteristics which could impact on the outcome for participant and non-participant groups in an existing administrative or survey dataset, matching and balancing these, and compares outcomes. Can be combined with DID design on the basis of observing the status of a treatment and comparison group prior to programme commencement and tracing the differential outcomes for these groups to a later time point.</p>	<p>Strong data needs on outcome and background/matching for the intervention and comparison group; this ideally includes historical information</p> <p>Good theoretical knowledge of which variables are likely to affect the outcome.</p>	<p>Groups vary on relevant factors which cannot be accounted for, results are biased</p> <p>Likely to be expensive and potentially difficult to implement if primary data collection is required, difficult to engage comparison group.</p>	<p>As above</p>	<p>Matching design is the most feasible design if robust comparison groups can be identified. PSM generally seen as most likely form of matching in the context of an ESF evaluation.</p>

Source: Ecorys' own illustration

The above assessment indicates that the most likely CIE designs are some form of propensity score matching (PSM) design, along with, potentially, a difference-in-differences design (DiD). Neither a randomised control trial (RCT) or regression discontinuity design (RDD) are likely to be feasible in the context of the ESF programme. While this initial assessment offers an understanding of what CIE approaches offer the most potential, it is necessary to consider some of the key issues facing any CIE design. Such issues require examination as they might also influence the approach or approaches selected. This process is the focus of sections 4.3 to 4.5 below.

4.3 Assignment onto provision

Understanding how people are selected for or assigned onto ESF provision is important in designing a robust CIE. All approaches of this type must outline the selection or assignment mechanism and discuss how the research design implemented addresses the biases arising from this. In the context of this feasibility exercise, we primarily sought to understand selection into the ESF through reviewing relevant documentation. If implemented, the ESF impact evaluation should ideally allocate resources to conduct qualitative research into how selection works at a project level to confirm, justify and contextualise the CIE.

The overall eligibility rules for the ESF are broad, offering no clear cut-off points for eligibility that could, for example, be exploited in an impact evaluation design. General eligibility criteria for individuals benefiting from ESF provision are:

1. Be legally resident in the UK
2. Be able to take paid employment in a European Union member state.³⁶

Although there is no official upper age limit for ESF provision, the support is for individuals “who will contribute to the growth of the economy through employment or increased skills level”³⁷ and it is the responsibility of providers to demonstrate that participants meet these criteria. The only exception to this focus on individuals of working age is that ESF provision can be provided to young people age 15 who are at risk of becoming NEET.³⁸ In addition, certain elements of the ESF programme have restricted eligibility on geographical grounds. The main consideration here relates to the Youth Employment Initiative (IP 1.3) which, as a geographically targeted initiative, requires participants to be resident in Youth Employment Initiative (YEI) -eligible areas.³⁹

³⁶ DWP. 2017. European Social Fund Programme for England 2014-2020. National Eligibility Rules. (p5)

³⁷ Ibid (p5)

³⁸ Ibid (p5)

³⁹ See European Social Fund Programme for England 2014-2020. National Eligibility Rules, p.5 for details of eligible areas.

ESF impact evaluation: research design and scoping study

In contrast to those granted refugee status, asylum seekers are typically not eligible for ESF support, but since 2005 they have been able to apply for permission to work if they have been waiting over a year for an initial decision on their asylum claim.⁴⁰ Those with permission to work can access the full range of ESF support, while others may be supported through ESF pre-vocational provision such as initial English for Speakers of Other Languages (ESOL) support.

For provision contracted through co-financing organisations (CFOs), depending on the CFO, the IP and the particular geographic area, the eligibility criteria for participants vary. Table 4.3 below provides an overview of the eligibility criteria for participants across each of the CFOs and by PA and IP. Directly contracted provision is subject to the general ESF eligibility guidelines outlined above and set out in full in the relevant programme guidance.⁴¹ This means that eligible participants include the unemployed, economically inactive,⁴² refugees, asylum seekers subject to some restrictions, the self-employed subject to confirming their self-employed status, individuals in prison depending on the length, nature and length of time left on their sentence, and individuals at risk of redundancy.

⁴⁰ Ibid (p5)

⁴¹ Ibid.

⁴² That is, those out of work but not satisfying the International Labour Organisation (ILO) definition of 'unemployed' in terms of being available for, and actively seeking, work.

Table 4.3 Participant eligibility criteria by CFO, PA and IP

CFO	Department for Work and Pensions	Big Lottery Fund	Education and Skills Funding Agency (ESFA)	Her Majesty's Prison and Probation Service (HMPPS)
PA	PA1	PA1	PA1 and PA2	PA1
IP	1.1	1.4	1.1, 1.2, 1.4, 2.1, 2.2	1.4
Criteria for eligibility	<p>The participant must:</p> <ol style="list-style-type: none"> 1. Be legally resident and have the right to work in the UK 2. Be over the age of 16 3. Be unemployed or inactive and fall into one or more of these three categories: <ul style="list-style-type: none"> • Long-term unemployed (LTU) and/or inactive (26 weeks and over) • Basic skills need (Below NVQ level 1 or equivalent) • Have more than one barrier to employment, including (but not limited to... any other barriers must be agreed with DWP): lone parent; older worker (50+); ex-offender (someone who has completed a custodial sentence or community sentence); caring responsibilities (including those returning to work after caring responsibilities end); having physical disability of health condition, including sensory impairments; mental health; learning disability; drug/Alcohol dependency; an ethnic minority; low or no qualifications (Below NVA level 2); language barrier (English is not a first language) 	<p>The participant must:</p> <ol style="list-style-type: none"> 1. Be legally resident in the UK and able to take paid employment in EU member states 2. Be unemployed or economically inactive <p>Activities working with young people aged 16-29 can only be for those who are NEET, or for those 15-18 year olds who are NEET or at risk of becoming NEET. Particular target groups may be specified in different LEP areas, but will include:</p> <ul style="list-style-type: none"> • People with parental or carer responsibilities • Women • People with health issues and disabilities • Black, Asian and Minority Ethnic groups • Refugees and recent migrants • People aged over 50 with few or no qualifications and not in employment • Low-income, single earner households with children where one partner is not working • Disabled people • Those out of contact with the labour market • (Ex-)offenders • Homeless people • People with low or no skills 	<p>Providers must only engage participants in ESF provision who meet the service-specific eligibility criteria set out in the relevant service specification. Participants must:</p> <ol style="list-style-type: none"> 1. Be legally resident in the UK 2. Be able to take paid employment in an EU member state 3. 15 years old or older <p>Evidence must show that participants will contribute to the growth of the economy through employment or increased skills levels.</p>	<p>All participants must:</p> <ol style="list-style-type: none"> 1. Be resident in the UK with permission to work 2. Unemployed (all prison starters are considered unemployed at point of enrolment) 3. Over 16 years of age in three regions (North West, West Midlands, London) must be under Youth Offender Team (YOT) supervision 4. Serving a custodial sentence (with up to three years left to serve) (Women offenders currently in remand are also eligible) 5. Completing a post custody licence/supervision period or serving a community order <p>Particular target groups or priority groups will be targeted across each ESF Contract Area (ECA). These include: Sentenced prisoners with up to three years to serve or on remand (women only); Offenders completing a post-custodial licence/supervision period or serving a community order; People from minority ethnic communities; People with disabilities and health conditions, including those linked to the misuse of drugs/alcohol; Over 50s; Ex-service personnel; Young people; Women (including those on remand)</p>

ESF impact evaluation: research design and scoping study

CFO	Department for Work and Pensions	Big Lottery Fund	Education and Skills Funding Agency (ESFA)	Her Majesty's Prison and Probation Service (HMPPS)
Source	DWP. 2017. <i>DWP Provider Guidance ESF 2014-2020. Chapter 3 – Participant Identification, Eligibility and Referral Contents</i> (Page 2-3)	Big Lottery Fund. 2015. <i>Building Better Opportunities: Programme guide</i> (October 2015). (page 5-6)	ESFA. 2017. <i>Funding and performance management rules 2014 to 2020 European Social Fund (ESF) programme: Version 3.</i>	NOMS. 2015. <i>Schedule G Specification.</i>

ESF impact evaluation: research design and scoping study

As well as the broad eligibility criteria applied to the ESF, the above table also illustrates that the target group, or groups, for ESF provision is/are heterogeneous, even within specific investment priorities. Directly contracted provision (not through the national or local CFOs) is likely to be similarly heterogeneous. While this does not preclude CIE approaches, it does have implications for the design of any approach.

There are several considerations requiring examination here. First, the heterogeneous nature of potential participants means that the development of any comparison group, or groups, will need to be undertaken carefully and sensitively to reflect the heterogeneity of the participant group. This issue is examined further in section 4.4 below and also implies a potential need for subgroup analysis, or development of different comparison groups, in respect of particular target groups, or broad participant types. While this is possible, it should be noted that this will require careful and further thought in developing the specific approach when implementing an evaluation. In particular, the challenge here is that such groups are not discrete, in that for example, individuals considered to be part of a Black, Asian and Minority Ethnic (BAME) group are likely to have overlapping characteristics with other target groups, such as those with disabilities or the long-term unemployed.

Second, the range of target groups involved suggests that there are likely to be a number of referral routes through which individuals engage with the programme. Consultations with stakeholders strongly indicated that there remain uncertainties around precisely how people are selected for different provision (in the sense of that run through each national CFO and non-‘opt-in’ provision). An understanding of this will need to be gained from more in-depth consultation with providers, particularly in terms of whether certain types of individual or target group are more likely to be referred, or engage with provision, than others.

In addition, mirroring the eligibility of participants in developing a comparison group will have to take account of the fact that some eligibility criteria are readily observable in potential sources of data for such groups – for example, administrative datasets held by DWP – though not all. Thus, some of the barriers to employment listed for IP 1.1 in table 4.1 above as qualifying eligibility criteria, notably caring responsibilities, might not be easily identified in such data. However, the majority of the criteria listed are, at least in theory, available. In our judgement, therefore, this needs to be understood as a potential limitation to approaches depending on matching to generate a comparison group, but does not cast serious doubt on their feasibility.

In respect of PSM, for example, matching variables for the majority of eligibility and target group characteristics listed in the table are available in the datasets considered in section 4.4 below. In addition, accepting that some characteristics, such as motivation, are unable to be exactly identified or reflected in a comparison group is inherent in PSM (though proxies and indicative variables can be used to partly account for this as outlined further below). As reflected in a previous ESF impact evaluation undertaken by the Department, however, it does underline the importance of articulating how the proposed approach accounts for selection bias and acknowledging the limitations of this.⁴³ Any such approach will need to address this issue as noted at the outset of this section.

⁴³ See Ainsworth, P. and Marlow, S. (2011), *Early Impacts of the European Social Fund 2007-2013*, DWP In-House Research Report No.3, particularly sections 3.1 and 3.2

ESF impact evaluation: research design and scoping study

Issues concerning eligibility and selection, along with the need to reduce bias, also give rise to the need to address a central underpinning assumption of a design identified as most promising in the a priori assessment above – PSM. As with other designs, PSM depends on addressing and reducing as far as possible any bias resulting from systematic differences between participants that would form a treatment group, and non-participants within a comparison group, in respect of the potential for these differences to influence the outcomes we are looking to test (e.g. entry to employment). For an approach to be valid, what is known as the Conditional Independence Assumption (CIA) must be satisfied; that is, as far as possible we must be confident and justify that the outcomes of individuals from the treatment and comparison group would not differ in the absence of treatment. Or, in other words, potential outcomes for these groups are independent of treatment status. This is a prerequisite of approaches that seek to identify the average treatment effect on the treated (ATT), as would be the aim of a PSM approach in this context.

Previous ESF impact evaluations, including one undertaken by DWP analysts in the last ESF programming period,⁴⁴ have addressed this issue through identification and careful selection of the explanatory variables, or covariates, that jointly determine the likelihood of selection or participation into the programme and outcomes from it. As with the study cited, a further requirement is to ensure that there are no characteristics amongst the treatment and comparison group unable to be observed from available data, such as motivation, that would effect outcomes, or that these characteristics are highly correlated with those covariates that have been chosen.

In advancing PSM as a possible approach, we follow the contentions in the previous DWP ESF impact study in terms of the potential to identify and select a range of covariates that, through PSM, can control for observed characteristics between participants and non-participants. We anticipate that a similar approach to that used would be developed, including collection of demographic characteristics for a large number of participants and non-participants such as age, gender, ethnic group, disability, marital status etc., allied to use of proxy variables to capture unobserved characteristics such as motivation.⁴⁵ As with the approach taken in the cited study, collating geographical districts and Index of Multiple Deprivation (IMD) scores could be used to address the fact that local labour market conditions are also likely to affect outcomes. Implicit in this discussion is the expectation that if PSM were to be used for an ESF CIE, appropriate tests would be undertaken through the analysis to, for example, assess matching quality, confirm the presence of the availability of support in terms of eligibility and check the sensitivity of the analysis to different methods.

Discussions in the impact evaluation workshop, held to inform consideration of issues such as eligibility and selection, additionally suggested that while the selection mechanism will need to be understood to contextualise a CIE as explained above, and understand potential sources of bias, there is also the potential to explore systematic differences between those engaging and the general population. From this perspective, whilst selection is important, any such systematic differences between participants and non-participants are potentially more so. Exploration of this could be included within a CIE approach through, for example, comparing management information (MI) on participants, or data from the ESF Leavers Survey, with demographic and secondary survey data concerning the general population

⁴⁴ Ibid.

⁴⁵ Ibid, p.26

from sources such as the Annual Population Survey (APS), or the Understanding Society project.⁴⁶ Such potential might, for example, be used to ameliorate the issues noted above concerning the (full) visibility of eligibility criteria as characteristics in administrative datasets.

While the potential for such an approach will need to be explored further in refining and finalising the impact evaluation methodology, it does indicate that there are viable routes to identifying and minimising any bias resulting from selection. Identifying any differences amongst those participating with the general population should be complemented by an approach that seeks to minimise, or at least understand and make explicit, any effects from the likelihood of individuals to participate in the ESF. As outlined, one of the approaches considered above, propensity score matching (PSM), is specifically intended to address this issue through seeking to match participants with those having a similar propensity to engage with an intervention, in this case ESF. This consideration also relates to potential comparison groups and data availability, and it is to this that we now turn.

4.4 Data availability and potential comparison groups

All CIE designs need, as a pre-requisite, to consider the data available for both the treatment group, in this case those supported by the ESF, and a potential comparison group. As noted earlier, the key consideration here is around how a comparison group or groups can be constructed to mirror those receiving support through the ESF. This is required so as to develop a reliable and robust counterfactual able to facilitate a comparison of impacts between treatment and comparison groups.

4.4.1 Data availability for the treatment group

A key start point of any approach looking to use counterfactual methods is to identify the data that will be available on those participating in ESF support. A range of data is collected in the context of regular ESF monitoring and evaluation. This includes:

- MI data on participants' characteristics and situation when they join and leave the programme. Providers are required to submit this information in order to make a claim.
- The ESF leavers survey which captures outcome data on participants 6 and 12 months after leaving, amongst other data and variables.
- Data collection carried out in the context of independent evaluations of CFO or project level activities (e.g. participant surveys within the context of the national evaluation of the Big Lottery IP 1.4 Building Better Opportunities provision).

In terms of the above, consultations with stakeholders indicated that data being collected at the level of national CFO programme or project levels, over and above the standard data requirements for the ESF, is unlikely to be useful for a CIE. This conclusion relates to the lack of consistent or comparable variables, covering enough

⁴⁶ Understanding Society involves a set of surveys informing a longitudinal household study capturing information every year about the social and economic circumstances and attitudes of people living in the UK.

of the programme, to contribute to data on the treatment group. The basis for data on ESF participants is thus the programme MI and/or that captured through the Leavers Survey.

As well as data collected directly from ESF provider returns, ESF MI offers the potential for matching to other databases to gather additional data and variables, whether in terms of participant characteristics or the results/impacts pertaining to them. This approach is currently being implemented in the impact evaluation of the YEI commissioned by DWP; this has confirmed the viability in principle of linking participant MI data to benefits, education and tax data held by relevant Departments. Specifically, relevant administrative datasets here include those held by DWP (Customer Information System (CIS), National Benefits Database (NBD), and Universal Credit (UC) database); the joint DWP-Department for Education (DFE) Longitudinal Education Outcomes (LEO) dataset; the National Pupil Database (NPD) held by DFE; Individualised Learner Records (ILR) data held by the ESFA/DFE; Higher Education Statistics Agency (HESA) data; and Real Time Information (RTI) data on salary, tax and NI contributions held by Her Majesty's Revenue and Customs (HMRC).

The ESF Leavers Survey also provides data that can be used to understand participant characteristics and outcomes, this having been used in the past within an ESF CIE in the context of the Welsh ESF programme.⁴⁷ The Leavers Survey is also being used as a data source in the aforementioned YEI impact evaluation, though not for the CIE element of that study. In contrast to ESF MI and data held in the administrative datasets specified above, which in theory at least includes data on all participants,⁴⁸ data from the Leavers Survey is by definition only a sample of ESF participants and is self-reported by survey respondents. In addition, as with all surveys, some of those surveyed are more likely to take part than others; hence any use of the data for the purposes of the ESF treatment group must take account of non-response bias.⁴⁹

4.4.2 Potential comparison groups and data availability

Determining appropriate comparison groups for the CIE implies the need to identify the different target groups engaging with ESF support, in the sense that the nature of the treatment group or groups will be key in comparison group identification. Table 4.3 above, and the surrounding commentary, defined these target groups from the perspective of all individuals eligible to receive ESF support under each IP and type of nationally co-financed provision. It is also important, however, to consider the target groups likely to be supported in light of the aims of different provision within the

⁴⁷ See Davies, R. (2017) *The Effect of European Funded Labour Market Programmes on Enhancing Employment in Wales*, available at: https://www.sheffield.ac.uk/polopoly_fs/1.735781!/file/A3_2_Davies.pdf (accessed 01.12.17)

⁴⁸ In reality, there will be some participants for whom MI data is missing or unable to be validated. Equally, not all participant details from the MI will be able to be linked to administrative data or, in other words, not all ESF participants recorded in the MI will be able to be 'found' in or matched to the administrative data sets cited.

⁴⁹ Non-response bias refers to the potential for bias to occur due to some potential respondents of a survey being more likely to participate than others, whether due to the topic of the survey, its mode of delivery or characteristics inherent in the potential respondent population.

ESF impact evaluation: research design and scoping study

ESF programme. This process serves to inform which comparison groups might be identified, reflecting on the need for such groups to mirror, as far as possible, those participating in ESF provision.

Table 4.4. below provides a high-level overview of the main target groups by PA and IP, drawn from a review of ESF documentation, including lists of approved contracts. In reading this table, it is worth bearing in mind that the bulk of DWP provision is delivered through IP 1.1. ESFA provision makes up the bulk of nationally co-financed provision delivered through IPs 1.2, 2.1 and 2.2, while IP 1.4 mainly includes Big Lottery Fund and HMPPS provision. The YEI (IP 1.3) is delivered through direct bids, as is IP 1.5 which is focused on Community-Led Local Development (CCLD), whilst direct-bid provision is also found across the other IPs.

Table 4.4 Principal target groups of provision by IP

Investment Priority	Principal target group(s)
1.1	Unemployed and inactive closer to the labour market
1.2	Young people, particularly those who are not in employment, education or training (NEET), or at risk of being NEET
1.3	NEET young people in YEI-eligible areas
1.4	Unemployed and inactive more distant from the labour market and facing multiple disadvantages / barriers to work
1.5	Marginalised communities and unemployed / inactive individuals within them
2.1	Unemployed / inactive plus those in work but at risk due to skills deficiencies or facing redundancy
2.2	Education and training partnerships and systems

As the table above indicates, the likely comparison group for provision within each of the IPs includes the following who are not participating in the ESF: unemployed and inactive individuals who are closer to the labour market, whilst still facing disadvantages or barriers; unemployed young people including NEETs in particular; unemployed and inactive people with multiple disadvantages or barriers to work who are therefore more distant from the labour market; unemployed and inactive individuals in marginalised communities; and individuals in work with skills deficiencies and/or who are facing redundancy. As discussed earlier in this chapter, the IPs and provision, or aspects of them, targeting systems or communities are unlikely to be amenable to a CIE. Therefore, the focus is on defining, at a broad level, ‘ideal’ comparison groups of individuals not participating in the ESF and considering what data might be available with which to construct comparison groups.

ESF impact evaluation: research design and scoping study

Table 4.5 below provides a summary overview of this, drawing on the nature of the ESF target groups with the scope to be part of a CIE, the types of impact or result likely to be the focus of analysis, and the nature of the ESF provision concerned. In doing so, it draws on the analysis undertaken in earlier sections of the report and the recognition that the administrative datasets cited in respect of the treatment group may also offer the potential for comparison group identification. Alongside this, the table also draws on additional desk research, stakeholder discussions and the impact evaluation workshop to consider further potential data sources for the comparison group. In particular, these include existing survey datasets from which comparison groups might be drawn, specifically the Annual Population Survey (APS), Longitudinal Survey of Young People in England (LSYPE), and English Longitudinal Study of Ageing (ELSA). As the table reflects, it may also be possible to identify and gather comparison group data from primary survey research. In addition to potential sources, a preferred source and brief rationale for this is also outlined.

Table 4.5 Overview of comparison groups and potential data sources

Treatment Group	'Ideal' comparison group	Possible data sources	Preferred data sources
Unemployed and inactive closer to the labour market	Unemployed / inactive non-ESF participants with few indicators of labour market disadvantage	Administrative datasets cited above – RTI, UC and NBD for employment impacts, others (e.g. ILR / HESA) for skills impacts in terms of qualifications. Potential use of all administrative datasets for comparison group identification and matching. APS for employment and skills outcomes using secondary survey data. Primary survey data for 'distance travelled' measures amongst a comparison group identified from administrative data.	Administrative data for treatment and comparison group to offer maximum comparability of matching and result/ impact variables, greatest chance of being implemented, lower resource implications, avoidance of non-response bias issues with survey use and potential non-comparability of question sets if using ESF Leavers Survey for the treatment group and other surveys for a comparison group. Potential for supplementary distance travelled survey comparing Leavers Survey data with primary survey of a comparison group if resources available and can be feasibly implemented (likely to be difficult).

ESF impact evaluation: research design and scoping study

Treatment Group	'Ideal' comparison group	Possible data sources	Preferred data sources
Young people, particularly those who are, or are at risk of, being NEET	Unemployed and inactive young people not participating in the ESF, including NEETs and those with a number of indicators of future NEET status such as low prior educational attainment, high unauthorised school absence	Administrative datasets cited above – RTI, UC and NBD for employment impacts, others (e.g. NPD/ ILR/ HESA) for skills impacts in terms of qualifications and engagement with training or education. Potential use of all administrative datasets for comparison group identification and matching, including NPD for indicators of future NEET status. APS and/or LSYPE for employment, education and training outcomes using secondary survey data.	As above, administrative data would be preferable for the rationale outlined. Possible use of NPD to derive variables for 'at risk of being NEET' would need further exploration and verification, as would the potential for linking with other datasets to track whether treated individuals did become NEET along with a similar statistically derived comparison group. Alternative of comparing ESF Leavers Survey with a cohort group in the LSYPE for outcomes of those 'at risk of being NEET' can also be further explored as a fall-back. No identified need for primary data collection.
NEET young people in YEI-eligible areas	NEET young people not receiving ESF support in YEI-eligible or similar areas with regard to youth unemployment and/or NEET levels	Administrative data-sets cited above – RTI and NBD for employment impacts, others (e.g. NPD/ ILR/ HESA) for skills impacts in terms of qualifications and engagement with training or education. Potential use of all administrative datasets for comparison group identification and matching. APS and/or LSYPE for employment, education and training outcomes using secondary survey data.	Already the subject of a prior design and scoping study and a further scoping report within the YEI impact evaluation itself. Currently being implemented through the preferred approach of using administrative data to identify the treatment and comparison group using PSM, and to compare employment, training and education outcomes through this. No identified need for primary of secondary survey data.

ESF impact evaluation: research design and scoping study

Treatment Group	'Ideal' comparison group	Possible data sources	Preferred data sources
Unemployed and inactive more distant from the labour market and facing multiple disadvantages / barriers	Unemployed / inactive non-ESF participants with multiple indicators of labour market disadvantage – e.g. spells of long term unemployment	Administrative datasets cited above – RTI, UC and NBD for employment impacts, others (e.g. ILR / HESA) for skills impacts in terms of qualifications. Potential use of all administrative datasets for comparison group identification and matching. APS for employment and skills outcomes using secondary survey data. Primary survey data for 'distance travelled' measures amongst a comparison group identified from administrative data.	Administrative data for treatment and comparison group to offer maximum comparability of matching and result/ impact variables, greatest chance of being implemented, lower resource implications, avoidance of non-response bias issues with survey use and potential non-comparability of question sets if using ESF Leavers Survey for the treatment group and other surveys for a comparison group. Leavers Survey could be used to inform results relating to 'distance travelled' in a non-CIE design; CIE could be used for this purpose but collecting primary data through a survey of a comparison group would require significant resource and would be potentially difficult to successfully implement.
Unemployed / inactive individuals within marginalised communities	Unemployed / inactive individuals in similar marginalised communities	As above	As above

ESF impact evaluation: research design and scoping study

Treatment Group	'Ideal' comparison group	Possible data sources	Preferred data sources
Those in work but at risk due to skills deficiencies or facing redundancy	Individuals in work but with low skill levels and/or facing redundancy (or drawn from similar sectors to those supported)	Administrative data sets cited above – RTI, UC and NBD for employment impacts on the low skilled, including potential to use RTI and UC for wage impacts for the treatment and comparison group (defining the latter would be problematic due to no indicator of 'at risk' of redundancy being available in administrative data). APS / Understanding Society may be used to generate a comparison group to compare employment outcomes for those 'at risk of redundancy' if sectorally defined, with outcomes for the treatment group being identified through the ESF Leavers Survey.	While needing further exploration, the preferred option if feasible is to compare employment outcomes for those facing redundancy within the ESF Leavers Survey with a comparison group generated using secondary survey datasets (given that identifying such a group in administrative datasets appears infeasible due to such 'at risk' or sectoral elements not being able to be identified, whereas survey data at least gives the potential of comparing with individuals in certain sectors if the sectors as a whole are seen as being 'at risk'). Employment outcomes for those in employment with low skill levels could be compared with a similar group identified using administrative data, as could wage outcomes.

ESF impact evaluation: research design and scoping study

As the above table indicates, for reasons including the feasibility of implementation, resource issues, and likely comparability of variables across the treatment and comparison groups, using administrative data for a CIE is generally preferable to using survey data where this is feasible (taking the results/impacts concerned into account). This holds true for both secondary survey data and data collected through primary survey research. Moreover, an approach using administrative data should be feasible for the most common and key impacts across IPs and different provisions – namely those linked to employment status, training and education status, and results relating to qualifications gained.

As the above table also indicates, however, for some results and impacts, a CIE using administrative data is less likely to be feasible. It may be possible to compare results around distance travelled towards employment using a survey-based approach, as is the case for the support provided for those at risk of redundancy. However, the degree to which a primary survey data for ‘distance travelled’ could feasibly be gathered from a reliable comparison group is questionable. At the least, this is likely to be methodologically challenging, carry a significant risk of failure in terms of survey implementation, and require significant resources.

The challenge for estimating impacts for those at risk of unemployment primarily relates to the issue of defining a reliable comparison group of those similarly ‘at risk’. While, in theory, this could be done using administrative data if certain sectors were compared, such an approach would be prone to questions concerning the degree to which this accurately represents treatment and comparison groups. As detailed in the above table, in this case comparing ESF Leavers Survey data with a comparison group constructed from other secondary datasets may be the only viable fall-back option therefore. Impacts on a narrower group, those in work but ‘at risk’ due to low skill levels, is more feasible by drawing on administrative data. Equally, impacts on wages for this group, as a proxy for in-work progression, could conceivably be tested through a CIE, in that identification of a treatment and comparison group with similar skill levels should be possible.

Results for those ‘at risk’ of being NEET could, in theory, be the subject of a CIE drawing on administrative and/or secondary survey data to generate a treatment and comparison group. However, this would require further exploration of how far the variables that could be used as proxies for this status can be defined and justified, with reference to wider literature on the characteristics of such ‘at risk’ young people, allied to exploration through the ESF MI of the potential sample size for this group. At this stage the potential for such an approach should be treated cautiously therefore.

It is also worth reflecting briefly on the discussion of the likely unit or units of analysis in the preceding chapter. The likely treatment and comparison groups discussed above would, in ideal terms, fit with the preferred approach outlined. More specifically, where possible the approach to a CIE should be designed to produce an understanding of at least some impacts at the whole programme level, whilst also having the potential to facilitate analysis according to other dimensions such as types of provision (in this case national CFO programmes, direct bids, IP and PA levels) and within the heterogeneous nature of the target groups outlined through subgroup analysis. The most prevalent and significant impacts of the programme identified in chapter three, those of entry to employment, education and training status, and skills development, should have the potential to be analysed through a CIE approach, using PSM for example, at multiple levels and in respect of different target groups.

In respect of this latter point, however, as noted in section 4.3 above, this would require careful and further thought when further developing and implementing an evaluation. This relates to the challenge of reliably defining and disaggregating target groups when individuals comprising such groups have multiple and overlapping characteristics. Such an issue is less acute if target groups are defined at the level advanced in tables 4.4 and 4.5 above. However, if an approach is taken which seeks to identify treatment groups, and model comparison groups, according to more specific demographic characteristics or particular eligibility criteria (at the level of specific projects for example), then significant analytical challenges and confusion may arise.

Adopting a PSM approach can ameliorate some of these issues, in that the richness of ESF MI, and data available from administrative datasets, can facilitate subgroup analysis if particular populations of interest were carefully defined and matched on characteristics pertaining to those populations. This will require further scoping but, again, needs to be approached with caution the more specific such groups or populations of interest are made. The need to ensure adequate sample sizes also arises as a related consideration, as do potential challenges in respect of matching quality and balancing treatment and comparison groups – appropriate tests would thus be required in respect of these latter issues, as also discussed in section 4.3 above. Sample sizes and implications for subgroup analysis are considered further in section 4.5 below.

Other results and impacts highlighted above are less likely to be amenable to using PSM at multiple levels. If a CIE is to estimate impacts on those ‘at risk’ of being NEET, on distance travelled, and on in-work progression and job retention amongst the employed, the approach would necessarily be restricted to specific smaller elements of the programme. This should not preclude such approaches, in that their results could be reviewed in combination with those concerning more common intended impacts, around employment and skills, to generate insights into the overall impact of the programme. However, this issue does influence overall design considerations in terms of our recommended approach to the ESF impact evaluation, set out in the concluding chapter five below. Before this, we move to consider one further key issue in designing any CIE, that of available sample sizes for analysis.

4.5 Sample size considerations

When designing impact evaluations, consideration should always be given to the selection of a sample size and allocation⁵⁰ able to maximise the precision of impact estimates. In statistical terms, precision is expressed as the minimum detectable effect (MDE), which is the smallest true treatment effect that can be detected given specified levels of statistical power and statistical significance.⁵¹ In the case of an ESF impact evaluation, maximum sample sizes for the treatment group are set by the number of people participating in the programme. Sample sizes for the comparison group are constrained by the extent of access granted to existing administrative and survey data, and/or the resources available for additional primary data collection.

⁵⁰ That is, allocation between the treatment and comparison group

⁵¹ Standard levels of statistical powers are 80% and 0.05 for statistical significance

ESF impact evaluation: research design and scoping study

The above scenario indicates the key determinants and constraints in considering sample sizes for the impact evaluation. In considering the feasibility of a CIE in light of likely, and potential, sample sizes for treatment and comparison groups respectively, the start point is to review available data on anticipated participant numbers. Table 4.6 below specifies the participant numbers that can be expected for the ESF programme. While these output numbers give some indications of the potential sample sizes likely to be achieved, the following should be noted:

- Target numbers are to be achieved by 2023. A future impact evaluation must consider the time lag needed to assess sustained outcomes for participants. If a CIE is conducted before the end of 2023, this will decrease the sample sizes available. On current information it is likely that the evaluation will be undertaken in 2018 to 2019; hence sample sizes should be reviewed once the timescale is confirmed.
- Participants will have received different levels of exposure to the intervention, with some dropping out. It is unlikely that the evaluation can fully account for this.
- Participants may be double counted, as the ESF MI counts 'participations' rather than participants.⁵²
- For purposes of profiling, sample sizes for subgroup analysis will be smaller.

⁵² Our understanding is that individual participants might engage on different provisions, with each 'participation' being counted.

Table 4.6 Target participant numbers

Indicator	IP1.1	IP1.2	IP1.3	IP1.4	IP1.5	IP 2.1	IP 2.2	Total
Total Participants	764,900	180,000	112,000	302,600	42,100	927,300	/	2,331,900
Participants	764,900	/	/	302,600	42,100	/	/	1,373,600
Participants (below 25 years of age) who are unemployed or inactive	/	180,000	84,000	/	/	/	/	209,500
Participants (aged 25-29) who are unemployed or inactive (not in education or training)	/	/	28,000	/	/	/	/	28,000
Unemployed, including LTU	532,200	125,500	84,000	137,200	29,000	/	/	907,900
Inactive	194,400	45,600	/	150,400	11,100	/	/	401,500
Participants over 50 years of age	151,700	/	/	43,100	7,740	185,700	/	388,240
Participants from ethnic minorities	142,540	32,360	33,600	59,730	8,200	138,600	/	414,430
Participants with disabilities	198,200	18,250	11,200	68,200	9,900	72,800	/	378,550
Participants without basic skills	135,300	31,860	/	/	/	163,630	/	330,790
Participants who live in a single adult household with dependent children	100,600	8,560	5,000	/	/	46,240	/	160,400
Participants who are offenders or ex-offenders	/	/	/	96,500	/	/	/	96,500
Long-term unemployed participants (YEI)	/	/	28,000	/	/	/	/	28,000
Inactive participants not in education or training (YEI)	/	/	28,000	/	/	/	/	28,000
Number of supported micro, small and medium-sized enterprises (including cooperative enterprises, enterprises of the social economy)	/	/	/	/	/	/	17,900	17,900

Source: DWP. 2015. European Social Fund Operational Programme 2014-2020.

ESF impact evaluation: research design and scoping study

To determine the MDE accurately, we need to know the standard deviation and response distribution of the data. However, this will vary across the different outcomes the ESF is trying to achieve and is unable to be determined 'a priori'. Nonetheless, we can use a simple example to understand how likely it is that an effect can be identified by looking at a simple binary ESF outcome, such as being in employment (or not). In this, we assume that the size of the comparison and treatment group is equal, the statistical significance is set at 5%, statistical power is set at 80%, and the maximum possible variance is 25% (which is a conservative estimate). On this basis, the minimum detectable effect for the smallest sample size available (7,740 for participants 50+ years in IP 5) would be 0.02; i.e. 2%. If only half of this sample size is available for analysis, for the reasons outlined above, the MDE would be around 2.8%. It is therefore likely that a CIE will detect an effect if it exists. However, these calculations would need to be revisited when the actual sample sizes are available. In particular, the MDE may be higher in reality than that presented in the example, given that, as with any non-experimental (i.e. quasi-experimental) approach, there may be additional sources of uncertainty which could cause the MDE to rise.

Accepting the above, as the example provided indicates, we can also be relatively confident that available sample sizes will permit the estimation of impact in respect of key subgroups participating in the ESF programme. This should facilitate subgroup analysis by the particular 'ESF target groups' for whom participation and outcomes are measured, covering ethnic minorities, inactive individuals, the unemployed, the over 50s, those with disabilities, those with a low education level, ex-offenders, and young people under 25. As noted in sections 4.3 and 4.4, the use of PSM in these cases would need to be carefully applied in terms of the matching variables selected for treatment and comparison groups, particularly in recognition of the points made concerning the likely presence of overlapping characteristics across and between such 'ESF target groups'.

A further consideration worth briefly highlighting here is the potential need to address issues of multiplicity or the multiple testing problem that may occur. This arises when running multiple significance tests (i.e. testing multiple hypotheses) as would be the case in running analyses across the range of sub groups, and potentially ESF measures, noted above. While for individual hypothesis testing, the likelihood of detecting impacts caused by chance is relatively limited with a standard confidence level of 0.05, the chance of false positives or getting a significant result by chance increases in testing multiple hypotheses. In other words, the probability of getting a significant result due only to chance keeps rising as the number of simultaneous hypotheses being tested does. Addressing this will require consideration of adjusting the significance level being used and applying an appropriate correction method when conducting the analysis.

Having examined potential CIE designs above, and considered some of the key issues in respect of them, we now turn to present our recommended approach and highlight some further considerations in the concluding chapter of the report.

5 Recommended design for the impact evaluation

This final chapter brings our analysis together to present a recommended approach to the proposed European Social Fund (ESF) impact evaluation. To lay the groundwork for the evaluation further, the chapter also outlines a set of recommended key research questions for the study and considers some likely challenges, risks and limitations to the proposed approach.

5.1 Recommended approach and rationale

While any impact evaluation will require further development in preparation for, and in the course of, its implementation, the preparatory work undertaken in this design and scoping study is intended to provide the basis for this. The analysis strongly points to counterfactual impact evaluation (CIE) approaches to assessing impact in the context of the ESF programme being feasible, but only in respect of some results and impacts the programme seeks to generate. Equally, only certain designs are likely to be methodologically sound, defensible, and logistically feasible.

Our recommended approach thus involves a mixed-method evaluation of ESF impacts. This is required to assess the totality of potential impacts stemming from the programme, whilst also examining pertinent questions relating to the effectiveness of ESF provision in generating these impacts. The recommended approach combines CIE with theory-driven impact evaluation, also drawing on additional evidence from available sources such as ESF management information (ESF MI) and the ESF Leavers Survey. In our judgement, this approach has the greatest chance of providing robust evidence in line with the requirements for an ESF impact evaluation set out by the European Commission (EC), in addition to providing evidence to inform any future related interventions.

5.1.1 Recommended approach to the CIE

Based on the review of CIE designs and key issues requiring examination around their likely implementation, we recommend that propensity score matching (PSM), drawing on ESF and administrative datasets, forms the core of a CIE approach. As discussed in section 4.4, the main rationale for this is its potential to help ensure comparability between treatment and comparison groups through the development of a statistically generated comparison group with similar propensity to participate. The recommendation also rests on our judgement of the advantages of this approach over others considered, particularly in terms of the likelihood of it being successfully implemented in respect of the key intended employment and skills impacts of the ESF.

It also offers the potential for analysing impacts at the level of particular ESF target groups, through subgroup analysis, as well as the impact of particular groups or types of provision, in the sense of 'sub-treatment' analysis.

We recommend complementing the PSM with difference-in-differences (DiD) analysis when estimating impacts through the CIE. The rationale for this relates to the need to help address potential bias remaining from the application of the PSM approach. As far as possible, PSM corrects for differences between the proposed treatment and comparison groups on observed variables: for example, gender, age or educational level. DiD can help correct for unobserved variables, for example motivation, assuming that these are fixed over time and that outcomes of the beneficiary and comparison group have traditionally moved in parallel. In recommending this, we envisage that the ESF evaluation would follow the approach taken in the impact evaluation undertaken by DWP in the preceding ESF programming period.⁵³

In terms of the impacts that the CIE should focus on, we recommend that the above methodology be used principally to estimate the effect of the programme on entry to employment, (re-)engagement with education and training, and skills outcomes in respect of accredited levels and qualifications captured in education datasets. While further consideration can be given to the potential to use CIE approaches to estimate impacts on, for example, those 'at risk' of being not in employment, education or training (NEET), distance travelled towards employment, or support for in-work progression, the evidence reviewed suggests that this is likely to be methodologically challenging, resource intensive, and prone to possible implementation failure. In particular, those impacts assessed in section 4.4 as requiring survey data to support their implementation, particularly primary survey data, need further consideration from a resource and feasibility standpoint before they can be recommended as part of an impact evaluation approach.

5.1.2 Recommended approach to assessing other impacts

To ensure coverage of a broader set of impacts and results than those recommended for estimation through a CIE, we recommend that a theory-based evaluation design should complement the PSM-DiD approach suggested. Theory based evaluation in this context should be understood as involving the development of hypotheses concerning the presumed results of particular activities or types of support within ESF provision, with these being tested through development of an appropriate evidence base. We recommend that these hypotheses be developed with reference to the activity groupings and presumed causality between activities, results and impacts presented in section 2.3.3 of the report. In such a way, the theory-based evaluation approach would draw on the intervention logic and theory of change articulated in chapter two.

It should be noted that in recommending this approach we are not precluding the use of CIE to estimate the impact of particular groups or types of provision, but rather being realistic over the likely limitations to any 'sub-treatment' analysis within the PSM-DiD model proposed. As intimated in the preceding assessment, the main challenge if CIE is to be used to estimate the differential impact of provision 'types' is being able to adequately define and discretely analyse the effects of these. While

⁵³ See, Ainsworth, P. and Marlow, S. (2011), *Early Impacts of the European Social Fund 2007-2013*, DWP In-House Research Report No.3, p.32-33

this is possible at the Investment Priority (IP) level, or that of particular co-financed programmes, it would be much more difficult, and open to methodological challenge, if grouping individual ESF projects together at a lower level. There are two main concerns here. Firstly, from previous research it is evident that ESF provision typically bundles a range of support and activities together within single projects.⁵⁴ Any project typology is thus likely to be prone to significant overlap, and the potential for discrete analysis lessened. Secondly, the extent to which ESF MI could reliably be used to identify and allocate individuals to particular types of activity groups at this level, as opposed to IP or national co-financed provision, appears open to question.

In our judgement, a theory-based approach to exploring links between specific activities and results would provide a valid and viable alternative, or adjunct, to a CIE in cases where PSM could not reliably or feasibly be applied. Such an approach would rely on construction of an evidence base using qualitative data gathered from in-depth interviews with ESF providers, participants, and other stakeholders, allied to analysis of ESF MI and insights from Leavers Survey data. To test the hypotheses developed, attention should be paid to developing as robust an approach as possible to determining the causality of any effects observed. This process should focus on two main areas: first, determining the precise methodology and data sources through which to construct the evidence base in light of the hypotheses being tested; second, identifying a suitable methodology through which to consider questions of causality.

A range of different approaches to these areas can be conceived. Decisions on them will depend, in part, on the resources available for the evaluation and the methodological preferences of those commissioning and delivering the study. At a minimum, we recommend that qualitative fieldwork is undertaken covering all relevant ESF IPs and activity areas, potentially on a case study basis aimed at triangulating perspectives on particular types of provision from multiple perspectives. Findings from such fieldwork should be complemented and contextualised by analysis of ESF MI and the Leavers Survey as indicated above. Any fieldwork commissioned should be informed by a robust sampling strategy aimed at appropriate and proportionate coverage of the programme.

An appropriate approach to exploring causality should be integrated with this. Several options are available, though it is worth noting that contribution analysis has been used in related contexts previously and offers one viable methodological option.⁵⁵ Such an approach could also usefully draw on the contextual factors likely to affect ESF results and outcomes detailed in section 2.2.

As part of the theory-based approach, we recommend that effectiveness considerations form part of the focus. This is primarily to help understand which activities lead to which results and impacts, along with how and why, rather than

⁵⁴ See, for example, Atkinson, I et.al (2016), *Evaluation of the Department for Work and Pensions European Social Fund Support for Families with Multiple Problems*, DWP Research Report No 914 and Atkinson, I (2013), *Evaluation of European Social Fund: Priority 1 and Priority 4 (Employment and NEET) Provision*, DWP Research Report No 825

⁵⁵ 'Contribution analysis' is a technique that aims to build a credible 'performance story'; drawing upon available sources of evidence to consider whether an intervention, alongside other factors, contributed towards observed outcomes. See Mayne, J (2001) Addressing attribution through contribution analysis: using performance measures sensibly. *Canadian Journal of Programme Evaluation* [16]; p.1-24 for further elaboration of the approach.

focusing on results and impacts in isolation. Including effectiveness considerations is intended, in particular, to enable the ESF impact evaluation to offer lessons for future programmes, including a possible domestic successor to the ESF. Importantly, however, we recommend that care be taken in defining the evaluation scope so that effectiveness considerations do not lead to an over-emphasis on exploring process evaluation issues. We therefore suggest that the evaluation only considers process aspects where their likely effects on the intervention logics developed are direct and significant. From this perspective, examining effectiveness should focus solely on explaining why impacts occurred or did not.

5.1.3 Potential approach to incorporating Priority Axis 3

At an appropriate point, whether at the commissioning or early implementation stage of the evaluation, a decision should be taken concerning whether to incorporate Priority Axis (PA) 3 within the evaluation scope. If this is seen as necessary, we recommend that this forms part of the theory-based element of the suggested evaluation design outlined. Relevant questions pertaining to the presumed role of technical assistance in effectively supporting programme management and delivery should be developed. These questions should be incorporated in research tools for the fieldwork element described above, and contextual evidence from ESF MI and the Leavers Survey should be used to inform judgements of effectiveness and results. The suggested approach to the theory-based element of evaluating aspects of PA 1 and PA 2 should be reflected in the methodology adopted for PA 3.

5.1.4 Recommended additional data sources

As indicated, the theory-based and CIE approaches should be complemented by additional data from ESF MI and the ESF Leavers Survey where these have the potential to add to the evidence base. In particular, we recommend that the questions in the Leavers Survey around the effects of ESF on addressing barriers and contributing to distance travelled are incorporated in the assessment. Likewise, findings concerning the effect of the ESF on intended results, such as gaining qualifications and employment status, should be used to further contextualise the CIE. Similarly, the outcome and result indicators captured through the ESF MI should be used for this purpose.

5.1.5 Recommended approach to assessing value for money

We recommend that the DWP cost-benefit analysis (CBA) model be used as the starting point and main basis for a CBA in respect of the programme, and its component elements, where it is feasible to monetise both costs and benefits. We also suggest using the CBA frameworks or approaches developed by other government departments where appropriate and/or required in the context of likely benefits accruing from the ESF programme. For example, this might include estimating the returns resulting from skills development, or from activities to reduce re-offending, utilising Department for Education and Ministry of Justice/Her Majesty's

Prison and Probation Service approaches respectively. However, caution should be applied, and appropriate caveats used, to account for differences in how these approaches used by other departments vary from that used by DWP.

Using the DWP CBA framework as the basis for assessing the costs and benefits that can be ascribed to the ESF programme, and by extension the value for money (VfM) it offers, will enable a range of relevant benefits to be quantified. In the context of the ESF, these include, for example, most of the fiscal and social impacts of the programme as they relate to employment, including those concerning, for instance, changes in income, benefits payments and tax receipts. The results of the CIE, and data on participant characteristics gathered through the approach advocated should, moreover, be fed into the model where feasible to enhance the accuracy of estimates made.

Further consideration concerning the exact parameters of a CBA approach will be required when refining the design and implementation of the planned ESF evaluation. However, the above suggests that CBA should be feasible for considering VfM in respect of significant parts of the programme's likely costs and benefits as they relate to employment and other outcomes. Where other results of the programme are considered significant to incorporate into a VfM assessment, but are harder or impossible to monetise, potentially those relating to wellbeing for example, additional approaches such as cost-effectiveness analysis (CEA) should be considered and incorporated as appropriate into the VfM assessment.

5.1.6 Learning from related evaluations

In further supporting the rationale for the suggested approach, it is worth noting that adopting a combined PSM-DiD approach to the CIE, complemented by a theory-based approach, fits with, and can learn from, other ESF evaluation work commissioned by the Department. In particular, we recommend that insights gained from implementation of the Youth Employment Initiative (YEI) impact evaluation, for example around the feasibility of the PSM-DiD approach outlined, and the use of the required data to inform this, feed into further refinements of the suggested ESF impact evaluation approach.

5.1.7 Producing early evidence

One further question posed for this design and scoping study was the need to consider how an evaluation can be designed to produce early results and thereby inform ongoing development in DWP and the future of the ESF. The recommended approach presented above has been developed, in part, with this in mind. The timing of the CIE element of the recommended approach will necessarily be influenced by the need to allow time for the impacts of the programme to emerge. This suggests that the CIE should be implemented towards the end of the programme lifetime, so as to gain the most accurate estimate of its impacts. However, we recommend that an interim analysis of VfM is conducted part way through the evaluation, complemented by fieldwork to produce early findings in the context of the theory-based approach suggested. This could be undertaken at an appropriate point during the second half of programme delivery, and could be complemented by insights from available MI data and that from the ESF Leavers Survey.

DWP may also wish to consider a staged approach to the CIE element of the recommended approach, potentially producing estimations of programme effects at two points. However, we would recommend that such an option be treated with caution. In particular, too early a timing for the CIE element could lead to a false impression of the scale of programme impacts. Similarly, there is a need for realism over the time likely to be required to secure data access agreements and to implement the suggested design in the context of the remaining programme lifetime. We therefore suggest that an approach to developing early findings using activity groupings from a theory-based perspective, complemented by an interim VfM analysis drawing on ESF MI and unit cost assessment, would be preferable. As outlined, the thinking behind these approaches has, in itself, been influenced by considering how to generate evidence for future programmes. In our judgement, therefore, we recommend using these aspects for this purpose rather than implementing the CIE at different time points.

5.2 Proposed research questions for the impact evaluation

To help further inform the development of a specification for an ESF evaluation, this section presents a series of potential high-level research questions that the evaluation should look to address. The suggested questions are as follows:

1. Drawing on an appropriate counterfactual, what impact has the programme had on its key intended results around entry to employment, (re-)engagement with education and training, and skills development?
2. Drawing on an appropriate counterfactual, what impact has the programme had on supporting young people in terms of reducing NEET levels and increasing the number of young people in education, employment and training?
3. How, and in what ways, has the programme supported those at risk of being NEET and to what effect?
4. What impact has the programme had on social inclusion, particularly in terms of reducing barriers to employment, moving those furthest from the labour market closer to it, and addressing the root causes of poverty?
5. What impact has the programme had on structural effects such as improving the labour market relevance of skills provision and addressing skills shortages? How and in what ways are any impacts evident?
6. How, in what ways, and to what extent, has the programme supported in-work progression?
7. To what extent has the programme helped to address employers' skills needs – how and in what ways?
8. What contribution has the programme made to reducing the gender gap – how and in what ways?
9. How, and in what ways, has the ESF programme contributed to the ultimate objective of supporting smart, sustainable and inclusive growth?

10. To what extent does the ESF programme represent VfM in light of its costs and the impacts that can be attributed to it?
11. What is the relative cost-benefit, or cost-effectiveness, of the different common 'activity types' that can be identified within the ESF programme?
12. What are the key lessons from evaluating the ESF programme that can be drawn for any future domestic successor programme, and for related employment and skills development initiatives?

5.3 Challenges, risks and limitations

When finalising the specification and proposed approach to the ESF impact evaluation, it will be important from the outset to bear in mind the likely challenges, risks, and limitations to its successful implementation. Drawing on our understanding gained from the design and scoping study, along with previous experience of implementing related impact evaluations, this final section of the report highlights some likely challenges, risks and limitations that can be anticipated at this point. It is important to recognise that the list presented below is not exhaustive; inevitably, in the course of implementing the evaluation, further risks and challenges will emerge. There may also be limitations to what the evaluation can achieve that are not anticipated at this stage.

Accepting these points, the key challenges, risks and limitations to successful implementation of the design outlined can be summarised as follows:

1. Unavailability of key datasets and/or delays in accessing them, including ESF MI, the administrative datasets outlined, and the ESF Leavers Survey (risk)
2. Linked to this, a broader inability to implement parts of the proposed methodology due to unforeseen circumstances (risk)
3. Issues with the legal basis on which to access and process required data, particularly in light of the implementation of the new General Data Protection Regulation (GDPR) (risk)
4. Complaints by subjects engaged in the research (risk)
5. Changes in programme delivery arrangements, timescales, or lifetime due to the United Kingdom's exit from the European Union (EU) or changes in co-financing arrangements and participating organisations (risk)
6. Changing requirements for, and/or expectations or interpretations of, the requirements for an impact evaluation on the part of the European Commission (EC) (risk)
7. Synthesising evidence from the different methodological elements suggested to inform clear judgements on the programme as a whole (challenge)
8. Refining the approach suggested to ensure it can be successfully implemented in practice and in light of potentially changing circumstances regarding data availability or programme implementation (challenge)
9. Assessing and filling, where possible, any gaps in required data identified during further investigation and refinement phases (challenge)

ESF impact evaluation: research design and scoping study

10. Refining the approach to assessing VfM dependent on programme data availability and access, and in light of the resources available (challenge)
11. Recognising that only some aspects of the programme's results and impacts will be able to be quantified, principally those assessed through the CIE and those for which quantitative data is gathered through programme implementation (limitation)
12. Recognising the inherent limitations of CIE designs in terms of their ability to account for all potential differences between treatment and comparison groups (limitation)
13. Understanding that the results of the CIE will be estimates of impact and subject to a number of caveats stemming from issues with available data, decisions and potentially compromises needed in implementation, and any unforeseen issues affecting the reliability of results (limitation)
14. Recognising that the theory-based design elements proposed can only offer an understanding of causality as opposed to a definitive attribution of cause and effect (limitation).

As well as awareness of the above potential risks, challenges and limitations, effective project management will be required to ensure successful implementation of the proposed evaluation. Finally, to reduce these risks and challenges where possible, we recommend a further refinement stage in the early part of evaluation implementation. In particular, this should focus on confirming data availability and requirements for access, testing the extent to which the treatment and comparison group(s) can be identified in the administrative datasets outlined, and finalising the approach to estimating impacts that could be in scope for a CIE but which have significant methodological challenges and limitations to them.

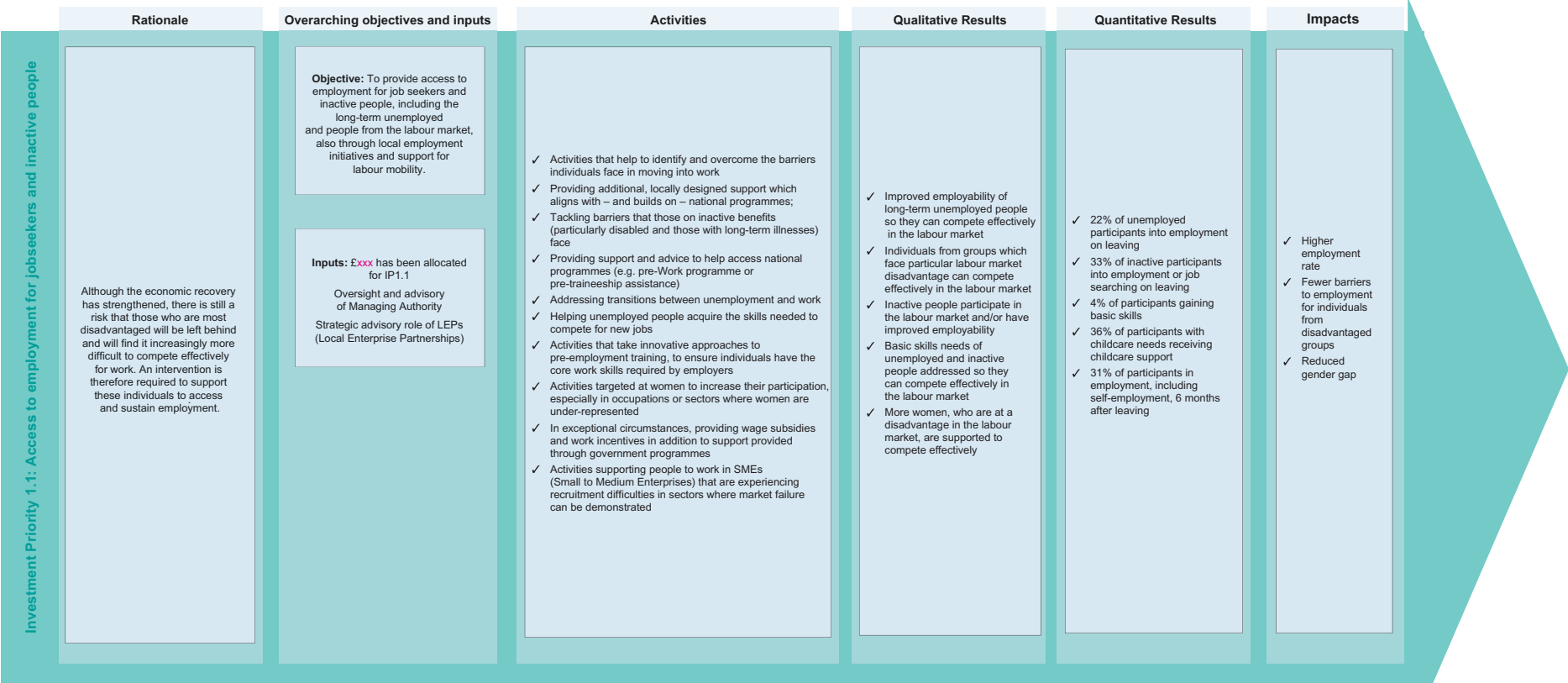
6 Annex one: additional technical detail

The sections that follow:

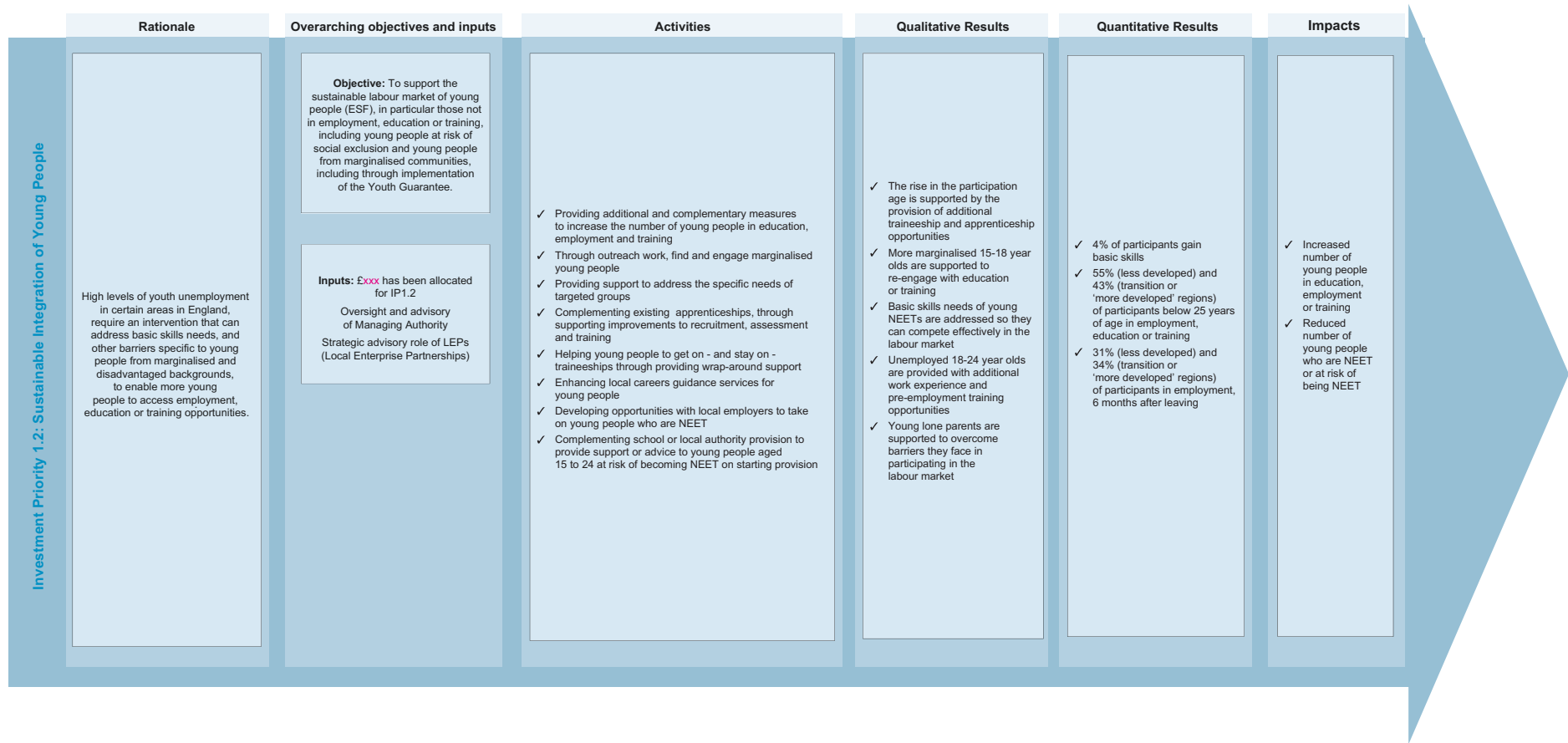
1. Present the additional intervention logics developed at the Investment Priority (IP) levels as part of articulating the theory of change behind the European Social Fund (ESF) programme.
2. Present the activity groupings derived from these IP level intervention logics that informed the development of a tracing of the ESF programme's causality detailed in chapter two of the main report and referred to elsewhere.

6.1 Intervention logics at the IP level

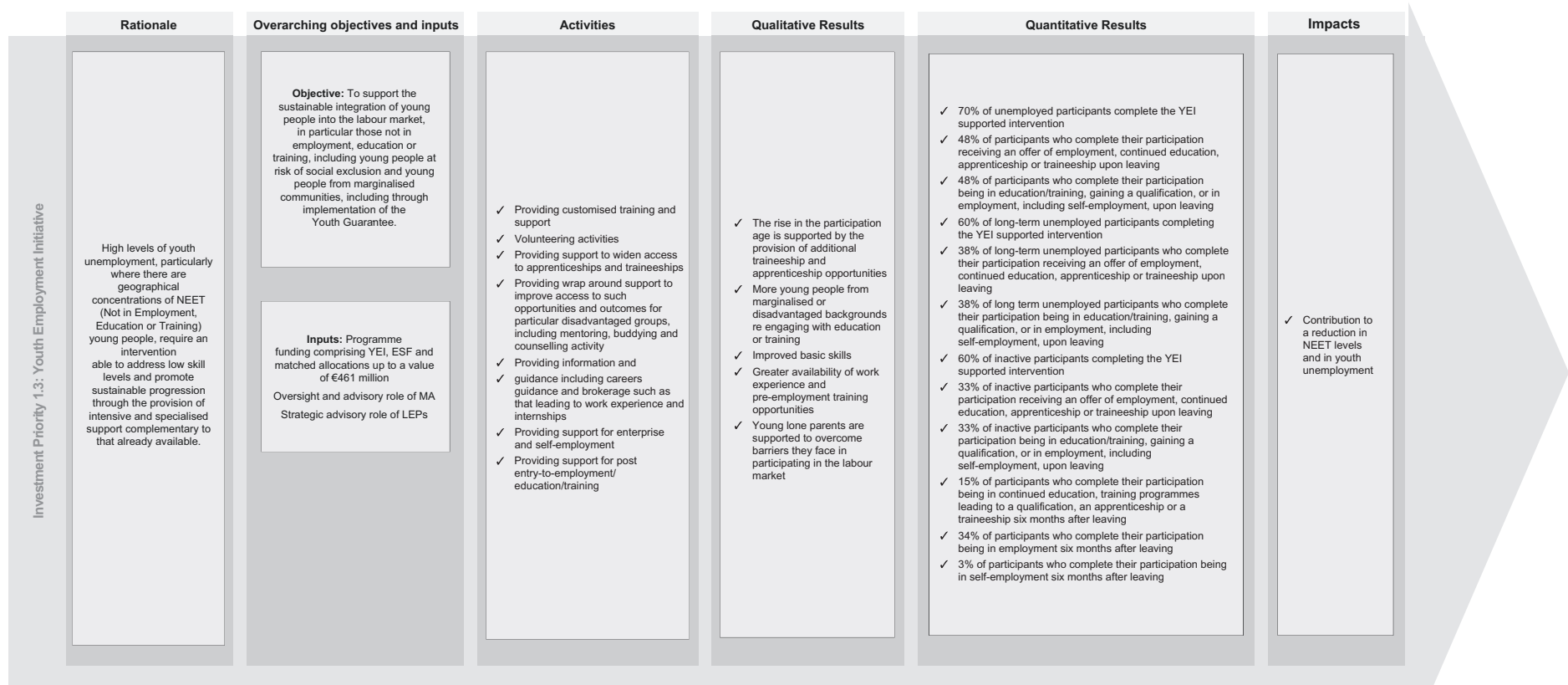
ESF impact evaluation: research design and scoping study



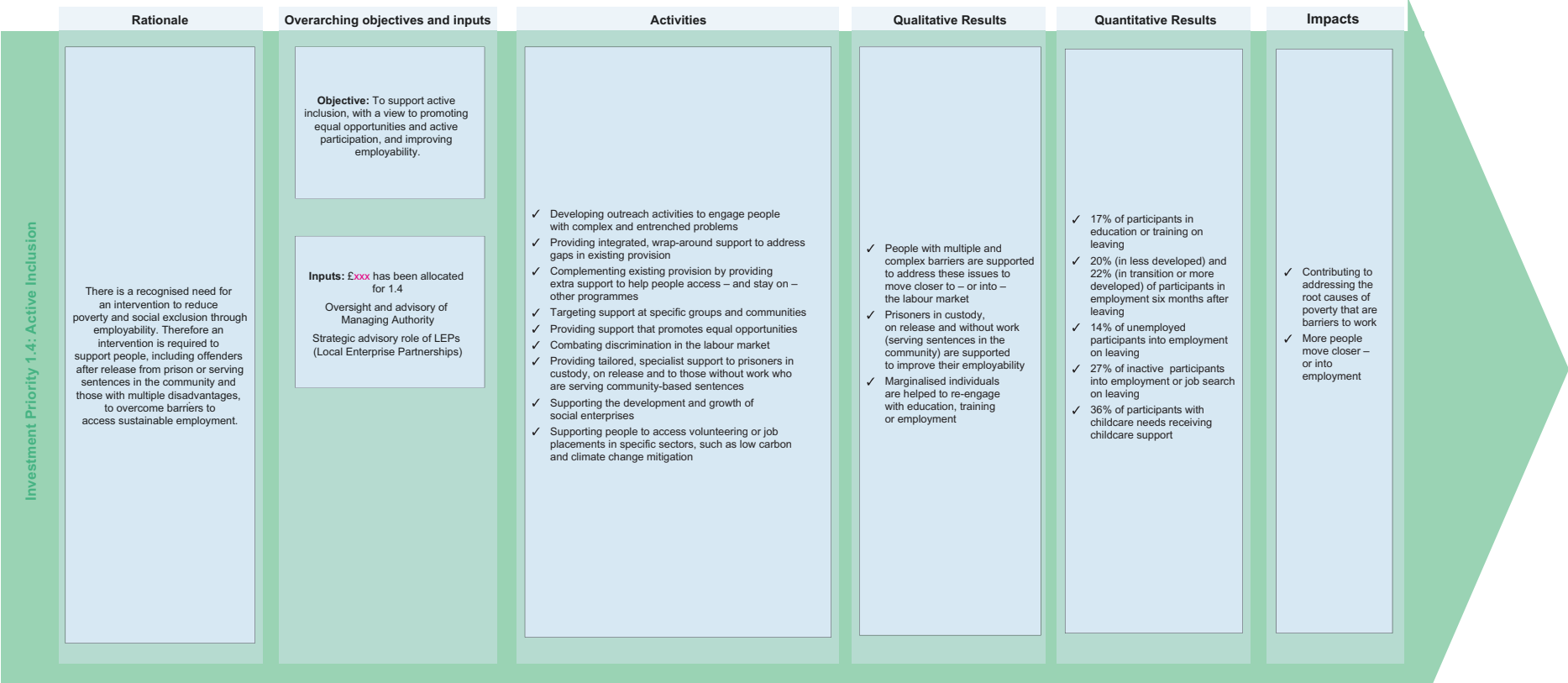
ESF impact evaluation: research design and scoping study



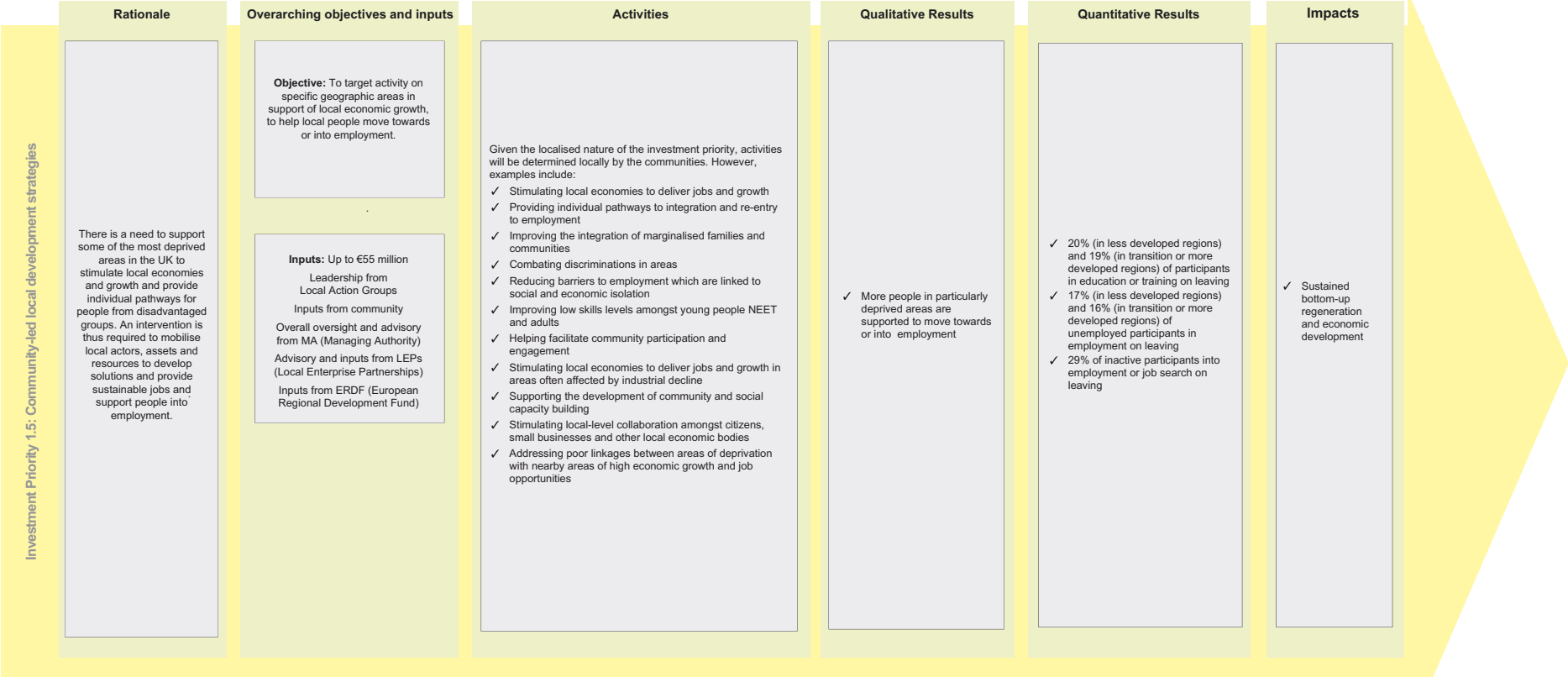
ESF impact evaluation: research design and scoping study



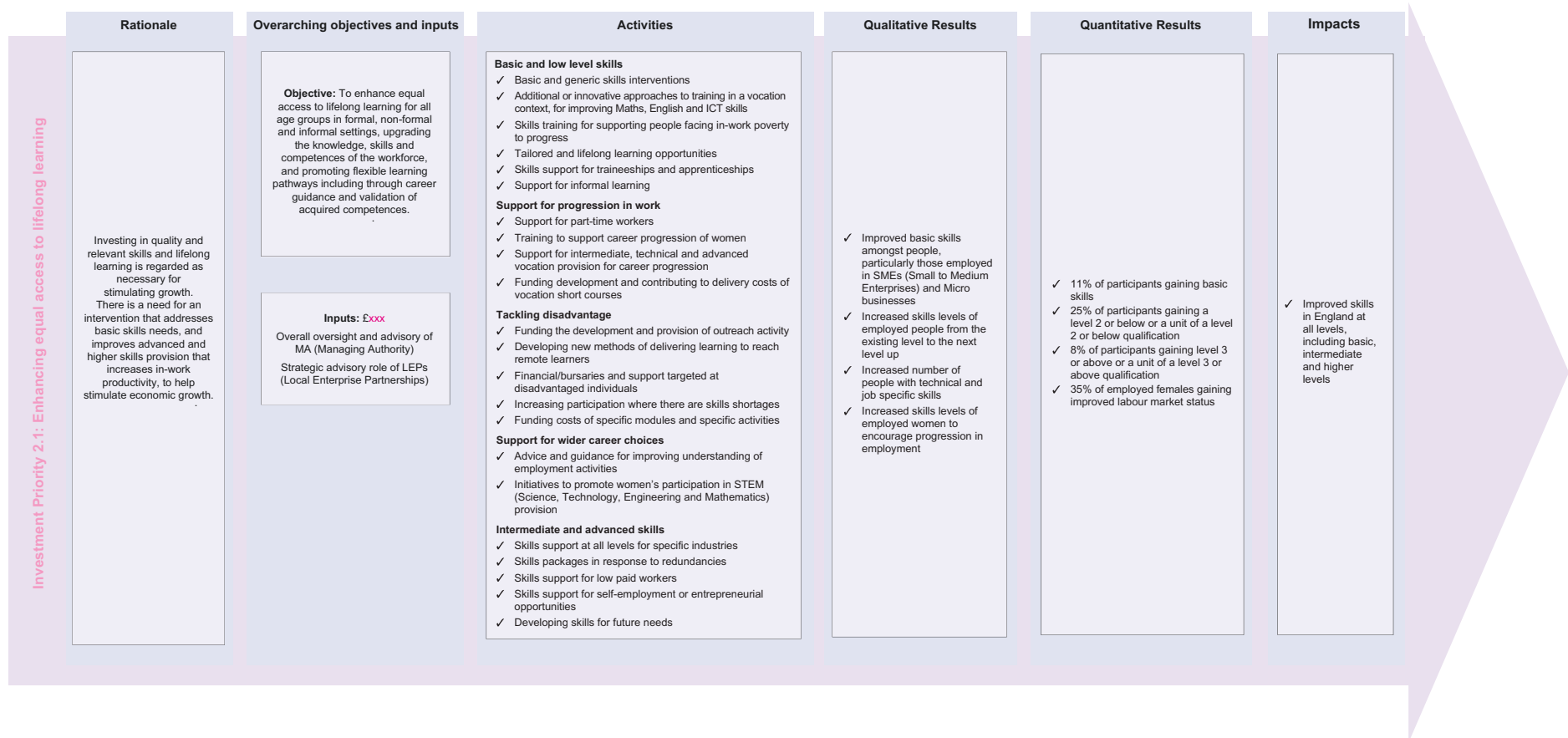
ESF impact evaluation: research design and scoping study



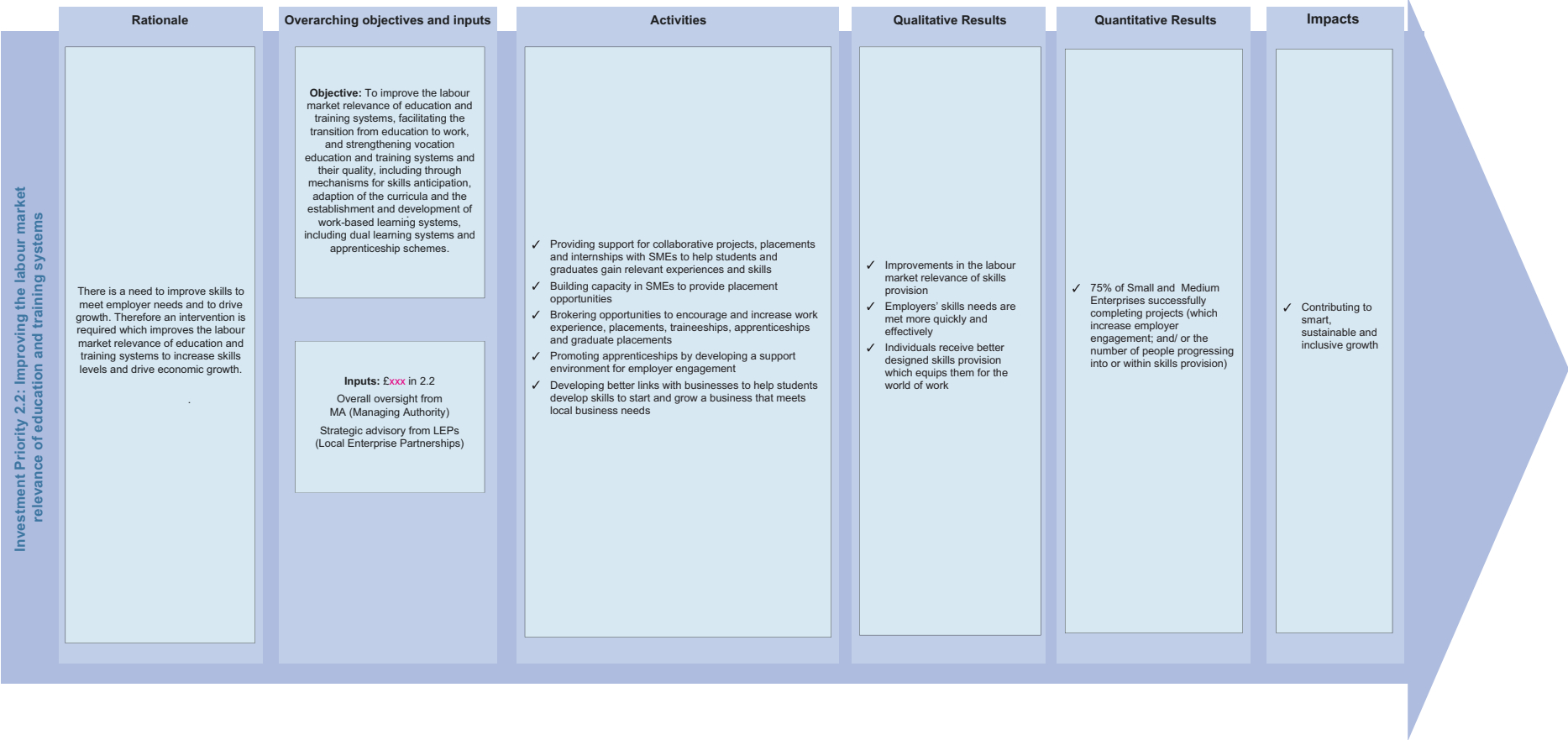
ESF impact evaluation: research design and scoping study



ESF impact evaluation: research design and scoping study



ESF impact evaluation: research design and scoping study



6.2 Activity groupings

A: Outreach activities to promote engagement of marginalised individuals

Through outreach work, find and engage marginalised young people
Developing outreach activities to engage people with complex and entrenched problems
Funding the development and provision of outreach activity

B: Provision of pre-employment support, guidance and advice to individuals

Activities that take innovative approaches to pre-employment training, to ensure individuals have the core work skills required by employers
Helping unemployed people acquire the skills needed to compete for new jobs
Addressing transitions between unemployment and work
Advice and guidance for improving understanding of employment activities

C: Providing tailored, wrap-around and holistic support to individuals (including financial assistance)

Providing wrap-around support to improve access to such opportunities and outcomes for particular disadvantaged groups, including mentoring, buddying and counselling activity
Providing integrated, wrap-around support to address gaps in existing provision
Providing individual pathways to integration and re-entry to employment
Providing support to address the specific needs of targeted groups
Helping young people to get on – and stay on – traineeships through providing wrap-around support
Providing customised training and support
Providing tailored, specialist support to prisoners in custody, on release and to those without work who are serving community-based sentences
Financial/bursaries and support targeted at disadvantaged individuals
In exceptional circumstances, providing wage subsidies and work incentives in addition to support provided through government programmes
Volunteering activities
Activities that help to identify and overcome the barriers individuals face in moving into work
Tackling barriers that those on inactive benefits (particularly disabled and those with long-term illnesses) face

D: Provision of post-entry to employment/education/training support

Providing support for post entry-to-employment/education/training

E: Provision of careers advice, guidance and brokerage services for individuals

Enhancing local careers guidance services for young people
Complementing school or local authority provision to provide support or advice to young people aged 15 to 24 at risk of becoming Not in Employment, Education or Training (NEET) on starting provision
Providing information and guidance including careers guidance and brokerage such as that leading to work experience and internships

F: Providing basic and generic skills support, including that to facilitate access to apprenticeships and traineeships

Complementing existing apprenticeships, through supporting improvements to recruitment, assessment and training
Providing support to widen access to apprenticeships and traineeships
Improving low skills levels amongst young people who are NEET and adults
Basic and generic skills interventions

ESF impact evaluation: research design and scoping study

Skills support for traineeships and apprenticeships

G: Support for informal and lifelong learning opportunities

Support for informal learning

Tailored and lifelong learning opportunities

H: Support for vocational and in-work skills development to promote progression and retention, including that targeted at particular sectors

Funding costs of specific modules and specific activities

Skills training for supporting people facing in-work poverty to progress

Support for part-time workers

Support for intermediate, technical and advanced vocation provision for career progression

Funding development and contributing to delivery costs of vocation short courses

Skills support at all levels for specific industries

Skills packages in response to redundancies

Skills support for low paid workers

Skills support for self-employment or entrepreneurial opportunities

Developing skills for future needs

Developing new methods of delivering learning to reach remote learners

Additional or innovative approaches to training in a vocation context, for improving Maths, English and ICT skills

I: Provision of support for self-employment and enterprise, including social enterprise

Providing support for enterprise and self-employment

Supporting the development and growth of social enterprises

J: Activities to promote equality and tackle labour market discrimination

Activities targeted at women to increase their participation, especially in occupations or sectors where women are under-represented

Providing support that promotes equal opportunities

Combating discrimination in the labour market

Combating discriminations in areas

Reducing barriers to employment which are linked to social and economic isolation

Training to support career progression of women

Initiatives to promote women's participation in Science Technology, Engineering and Mathematics (STEM) provision

K: Activity to link provision to and build upon mainstream provision

Providing additional, locally designed support which aligns with – and builds on – national programmes.

Providing support and advice to help access national programmes (e.g. pre-Work programme or pre-traineeship assistance)

Complementing existing provision by providing extra support to help people access – and stay on – other programmes

Providing additional and complementary measures to increase the number of young people in education, employment and training

L: Measures focused on demand-side stimulation including supporting access to particular sectors

Stimulating local economies to deliver jobs and growth

Stimulating local economies to deliver jobs and growth in areas often affected by industrial decline

ESF impact evaluation: research design and scoping study

Supporting people to access volunteering or job placements in specific sectors, such as low carbon and climate change mitigation

Increasing participation where there are skills shortages

Activities supporting people to work in SMEs that are experiencing recruitment difficulties in sectors where market failure can be demonstrated

M: Employer engagement to promote opportunities for individuals

Developing opportunities with local employers to take on young people who are NEET

Brokering opportunities to encourage and increase work experience, placements, traineeships, apprenticeships and graduate placements

Building capacity in SMEs to provide placement opportunities

Providing support for collaborative projects, placements and internships with SMEs to help students and graduates gain relevant experiences and skills

Promoting apprenticeships by developing a support environment for employer engagement

Developing better links with businesses to help students develop skills to start and grow a business that meets local business needs

N: Measures to support community development, engagement and linkages to opportunities

Helping facilitate community participation and engagement

Supporting the development of community and social capacity building

Stimulating local-level collaboration amongst citizens, small businesses and other local economic bodies

Addressing poor linkages between areas of deprivation with nearby areas of high economic growth and job opportunities

Improving the integration of marginalised families and communities

Targeting support at specific groups and communities

7 Annex two: Research tool for the stakeholder consultations

Interviewer notes / pre-amble

- Outline the aims of the study and purpose of the consultation.
- Inform the interviewee of the likely duration.
- Provide a brief overview of the scope of the issues the discussion will be looking to explore.
- Outline our treatment of data gathered and the approach to confidentiality.
- Read out the consent statement and gain explicit consent based on this.
- Check if the interviewee has any questions prior to commencing the discussion.

Background information

1. Check role of the interviewee in respect of the European Social Fund (ESF) programme – time in post; current remit / responsibility; previous roles
2. From your perspective, how would you describe the purpose of the ESF programme?

Intervention Logic Models

3. Gather any general views and overall comments on the intervention logic model or models shared in advance

[Programme / PA level for European Social Fund Division (ESFD) staff; PA / IP level as applicable for other stakeholders]

- a. Does the intervention logic adequately capture activity within the programme elements under review?
 - b. Is anything missing?
 - c. Does anything seem incorrect and, if so, why?
4. Discuss the following aspects of the logic models in turn and gather any additional comments:
 - a. Rationale (any additional factors influencing spend in the area being discussed?)
 - b. Objectives (any additional objectives not captured in the model?)
 - c. Inputs (any additional inputs to consider?)
 - d. Activities (additional activities not covered? Which are likely to be most significant in leading to change and why?)
 - e. Qualitative outcomes (are there any additional outcomes that should be reflected and why?)
 - f. Quantitative outcomes (are there any additional outcomes being captured / measured at project level and, if so, what?)
 - g. Impacts (are these the correct ultimate and longer term impacts as distinct from the shorter/medium term outcomes articulated – i.e. is this distinction appropriate and why / why not?)

ESF impact evaluation: research design and scoping study

5. What contextual considerations should be captured as part of developing a theory of change around the intervention logic / logics?
 - a. What are the key factors that may influence / affect the intervention logic?

Focus of the planned evaluation

[Tailor according to the role of the interviewee and area of programme being discussed]

6. What are the key outcomes and impacts the evaluation needs to test and why?
7. What factors are likely to influence how effective the programme / PA / IP is in generating outcomes and impacts?
8. In assessing value for money, what are the key considerations for the evaluation to focus on?
 - a. Simply costs / benefits? Wider considerations?
9. Do you have any other views on what any ESF evaluation should cover and/or how it should be undertaken?

Additional Questions for CFOs

10. Within the provision you manage, how are participants generally recruited and assigned to the programme?
11. Are you consistently capturing any data from funded projects in addition to the outputs and results indicators detailed in the Operational Programme under the relevant PA and IP, and which is additional to that required by the Commission / MA guidance?

If so...

 - a. What is this data?
 - b. How is it captured and validated?
 - c. How is it used and for what purpose?
 - d. Can you provide a written list of this data?
12. What evaluation activity are you undertaking in addition to that commissioned through the national ESF evaluation team?
 - a. What type of evaluation is involved (e.g. process, impact, methodological approach etc.)
 - b. What are the reporting schedules for the evaluation?
 - c. Does the evaluation involve the collection of any primary survey data? If so what are the planned sample sizes and sampling approach?

Additional questions for ESIF sub-committee representatives

13. Are there any plans you are aware of to conduct evaluations of ESF activity at the LEP / ESIF sub-committee geographical level?
 - a. What type of evaluation is involved (e.g. process, impact, methodological approach etc.)
 - b. What are the reporting schedules for the evaluation?
 - c. Does the evaluation involve the collection of any primary survey data? If so what are the planned sample sizes and sampling approach?
 - d. Do you use any other data or monitoring at the local level to assess the success of LEP activity in respect of the ESF or ESIF more broadly?