

**UNIVERSITÀ DEGLI STUDI
DI MODENA E REGGIO EMILIA**

Dottorato di ricerca in Lavoro, Sviluppo e Innovazione

Ciclo XXXII

EU human capital investments:
monitoring and evaluation, from theory to practice

Candidato: Andrea Giannetto

Relatore (Tutor): Prof. Massimo Baldini

Coordinatrice del Corso di Dottorato: Prof. Ssa Tindara Addabbo

Ringraziamenti

La lista dei ringraziamenti per chi ha reso possibile la realizzazione di questo sfidante progetto è, potenzialmente, infinita.

Anzitutto, ringrazio tanto l'Università di Modena, in particolare la Fondazione Marco Biagi, quanto la Fondazione Giacomo Brodolini, ente di ricerca presso il quale ho svolto le attività di ricerca, per aver reso possibile il compimento di questo percorso coniugando le reciproche esigenze e vocazioni. Ringrazio inoltre il Regional Economic Modelling team del Centro Comune di Ricerca della Commissione Europea (JRC di Siviglia) per la collaborazione nell'ideazione di parte della ricerca. Mi auguro che questo lavoro ripaghi gli sforzi di tutti i soggetti coinvolti.

Un ringraziamento particolare al Professor Massimo Baldini, che ha guidato la mia navigazione in questi tre anni in aree di ricerca non sempre convenzionali o consuetudinarie e compreso le difficoltà fronteggiate, mostrando interesse e supporto. Un ringraziamento alla Professoressa Graziella Bertocchi, per i rilevanti e puntuali commenti e gli spunti di riflessione offerti. Un ringraziamento particolare anche ai *referee*, Professor Giuseppe Croce e Professor Francesco Pastore, per le loro dettagliate valutazioni, che hanno contribuito in modo determinante a migliorare la qualità e chiarezza della ricerca. Naturalmente, il ringraziamento si estende alla Professoressa Tindara Addabbo ed alla segreteria del corso di dottorato, per la loro disponibilità e collaborazione, nonché ai miei colleghi per la pazienza ed il supporto.

Tuttavia, il ringraziamento più importante va a chi mi è stato vicino in questi anni. Alla nonna Maria, a cui vanno le mie scuse più sincere per la lontananza e la mia sconfinata gratitudine per il suo amore incondizionato. A chi non c'è più, e senza i quali non sarei nulla e non avrei realizzato nulla. Alla mamma e al papà, a Leila, per il loro amore e supporto totale. A Matteo, a Claudia, agli amici di sempre, a chi in un modo o nell'altro mi ha offerto un consiglio, del conforto o anche solo a chi ha avuto pazienza, comprendendo il senso di tutto questo. Il mio impegno è di esserci e di vivere di più, più pienamente, con voi.

Ma grazie soprattutto a te, Stefania, che sei la mia luce e il mio rifugio.

Sintesi

Questo progetto è il prodotto della ricerca individuale dell'autore combinata con l'esperienza lavorativa maturata con la Fondazione Giacomo Brodolini nel campo dell'analisi dei sistemi di monitoraggio e valutazione degli investimenti dell'Unione Europea orientati allo sviluppo del capitale umano.

La natura ibrida del progetto è funzionale all'obiettivo di contribuire al miglioramento dei meccanismi di trasmissione tra conoscenza accademica sui temi oggetto di studio e la valutazione di *policy* "in practice". Tale obiettivo diventa cogente poiché la valutazione di *policy*, a causa di vincoli relativi alle tempistiche o risorse disponibili, rischia di essere solo parzialmente aggrappata ad evidenze di provenienza accademica. Inoltre, le stesse tecniche in particolare di monitoraggio delle politiche per lo sviluppo del capitale umano, restano ai margini o sono del tutto omesse dal dibattito accademico, con inevitabili conseguenze in termini di qualità delle informazioni prodotte dai relativi studi di valutazione. Il fine di questa ricerca è quindi quello di analizzare gli standard, le pratiche e alcuni esempi concreti di funzionamento dei sistemi di monitoraggio e del modo in cui questi possono informare la valutazione, al fine di identificare e proporre dei miglioramenti incrementali che siano quanto più possibile basati sulla letteratura scientifica sui temi in oggetto. In particolare, l'autore ritiene che migliorare la capacità di analisi dei sistemi di monitoraggio sia funzionale a produrre informazioni di maggiore qualità e granularità, così da permettere poi alla valutazione ai vari livelli (micro, meso, macro) di studiare le effetti delle *policy* in modo adeguato. Considerando la natura ibrida dei temi oggetto di questa ricerca,

Pertanto, questa ricerca tende a discostarsi dagli standard puramente accademici principalmente per due ragioni:

- un perimetro dell'esercizio molto ampio, il che implica che una parte della ricerca sia di natura definitoria e tassonomica. Tale aspetto è considerato requisito necessario per meglio analizzare in ottica comparativa diversi concetti, approcci e metodologie e in qualche modo si riduce la possibilità di approfondimento di alcuni specifici aspetti della letteratura. Questo tipo di approccio ha dei risvolti sullo stile, la struttura e il grado di sintesi della ricerca presentata di seguito;
- l'analisi di argomenti che sono tipicamente ignorati dal dibattito accademico, il che determina l'assenza o scarsità di fonti comparabili in letteratura. A sua volta, questo elemento influenza la possibilità di effettuare analisi comparative o di sensibilità approfondite per alcune delle stime computate nell'ambito della sezione più sperimentale della ricerca, le quali, pertanto, necessiteranno di maggiore approfondimento in futuro. Al tempo stesso, tale aspetto solleva il tema delle ragioni per le quali argomenti di grandissima rilevanza per il policy making non siano affrontati sistematicamente dalla letteratura accademica. In tale ottica, questa progetto di ricerca prova a compiere un primo passo per colmare questa lacuna.

Ciò premesso, la presente ricerca verte in particolare su:

- (i) capitolo primo: la definizione di un – quanto mai necessario – strumento analitico fondato sulla conoscenza cumulativa nel campo del monitoraggio e della valutazione così come della misurazione del capitale umano. Tale strumento è finalizzato a supportare ricercatori e professionisti nell'esame dei meccanismi e delle disposizioni relative ai sistemi di

monitoraggio per i programmi che promuovono il capitale umano finanziati dall'Unione Europea. Il valore aggiunto di questo strumento è che sebbene esista una florida letteratura che discute il tema del capitale umano e, separatamente, degli studi e manuali sul ruolo e le caratteristiche di monitoraggio e valutazione (M&V) separatamente, manca una discussione organica e unitaria del tema del M&V del capitale umano soprattutto in ambito UE. Questa mancanza tende a spiegare anche la scarsità di analisi dedicate a questi aspetti, come maggiormente approfondito nel capitolo 2. Al fine di colmare questa mancanza, nel capitolo si presenta una analisi comparativa delle linee guida di organizzazioni supranazionali ed internazionali sul tema del M&V *vis-à-vis* la letteratura rilevante sul tema del capitale umano, al fine di assicurare che la conoscenza accademica sul tema del concetto, determinanti e misurazione del capitale umano siano correttamente incorporati negli strumenti che influenzano direttamente il ciclo di *policy* e i *policy makers*.

- (ii) capitolo secondo: due casi studio che: (i) esaminano, sulla base dello strumento analitico sviluppato nel capitolo uno, i sistemi di M&V di tre programmi chiave dell'UE che promuovono il capitale umano, in particolare: il Fondo Sociale Europeo (FSE), l'Iniziativa Occupazione Giovani (IOG) e il Fondo di aiuti europei agli indigenti (FEAD); e (ii) forniscono raccomandazioni per il futuro. Come anticipato, non esiste in letteratura una valutazione "pura" del sistema di monitoraggio dei programmi UE per lo sviluppo del capitale umano. Esistono degli studi che sono commissionati dalla stessa Commissione Europea, i quali, tuttavia, coprono solo indirettamente gli aspetti legati al monitoraggio (per esempio nell'ambito di studi di impatto dei programmi) oppure non si basano su *framework* analitici che siano sufficientemente comprensivi o adatti allo scopo. Pertanto, il livello di conoscenza su questi temi resta limitato;
- (iii) capitolo terzo: una ricerca sperimentale che suggerisce un nuovo *framework* analitico per la valutazione degli impatti macroeconomici del FSE/IOG. La ricerca in particolare esamina se sia possibile, ed attraverso quali metodi, utilizzare le informazioni provenienti dal sistema di M&V del FSE/IOG relative al numero e le caratteristiche dei partecipanti ai programmi assieme gli impatti su di essi (equilibrio parziale) per stimare, attraverso, il modello macroeconomico RHOMOLO del Centro Comune di Ricerca della Commissione Europea (JRC), gli impatti macroeconomici (equilibrio generale) delle politiche per l'occupazione dei giovani. Il motivo di questa sperimentazione è che, da un lato, si è identificato un significativo *gap* nella letteratura che relativa agli effetti macroeconomici di questi programmi che tenga al contempo conto della diversità delle politiche finanziate, della eterogeneità degli effetti sui partecipanti e della loro distribuzione e variazione nel tempo e nello spazio – che sono elementi essenziali per le politiche *place-based*. Dall'altro lato, lo strumento che è correntemente utilizzato dalla Commissione Europea per stimare gli effetti sopramenzionati nei termini necessari alle politiche *place-based*, poggia su assunzioni riguardanti la tipologia/intensità del supporto e le caratteristiche dei destinatari del FSE/IOG che non corrisponde alle attività effettivamente finanziate. Pertanto, il progetto propone un metodo innovativo per utilizzare le informazioni provenienti dal sistema di monitoraggio del FSE/IOG le quali, elaborate econometricamente e triangolate con le evidenze raccolte attraverso le valutazioni di impatto controfattuale sulle politiche per l'occupazione giovanile finanziate nel corso della programmazione 2014-2020, possono fungere da input per le attività di ricerca svolte dal JRC e dalla direzione generale Occupazione, Affari Sociali ed Inclusione (DG EMPL) concernenti gli impatti macroeconomici degli sopracitate politiche.

Questo progetto di ricerca sviluppa (si allontana dal) la letteratura esistente nei seguenti modi:

- capitolo primo: questa sezione non ha l'obiettivo di contrastare la letteratura accademica ma piuttosto di utilizzarla per sviluppare un *framework* analitico per il monitoraggio e la valutazione delle politiche per lo sviluppo del capitale umano. Così facendo, la ricerca mette in luce alcune lacune degli standard e linee guida dell'Unione Europea in relazione alla misurazione del capitale umano e più in generale alla prassi di monitoraggio e valutazione, che si ritiene possa essere migliorata sotto vari aspetti fra di loro complementari. Fra gli altri, la ricerca indica che sarebbe necessaria una maggiore enfasi sugli aspetti del monitoraggio relativi alla misurazione dei risultati e che, nel disegnare i sistemi di monitoraggio, si dovrebbe lavorare in modo più sinergico con le esigenze della valutazione. Inoltre, si considera necessaria la creazione di linee guida e standard che siano customizzati su determinati settori o temi, come ad esempio quello dello sviluppo del capitale umano. Infine, sarebbe necessaria una maggiore attenzione alle implicazioni per i sistemi di M&V in relazione ai diversi livelli a cui questi operano, ovvero micro, meso e macro;
- capitolo secondo: l'obiettivo precipuo di questo capitolo è di rispondere alla scarsità di studi che si occupino specificamente della valutazione dei sistemi di monitoraggio. In riferimento al sistema di monitoraggio del FSE/IOG, la ricerca evidenzia la mancanza di comprensività e profondità degli approcci analitici usati fin qui per studiarne l'adeguatezza, anche a causa del fatto che questi studi non possono poggiare su strumenti analitici consolidati ed evidenze accademiche a dispetto della rilevanza dei temi. Questi problemi sono particolarmente evidenti in relazione ai giudizi espressi da questi studi in relazione alle attuali regole per la raccolta e utilizzo di micro-dati, per la raccolta di informazioni sulle diverse forme di supporto offerte e sulla rilevanza degli indicatori di programma specifici per gli studi di valutazione. Per quanto riguarda il FEAD, in mancanza di analisi dedicate allo studio del suo sistema di monitoraggio, risulta impossibile effettuare una comparazione con la letteratura. Ad ogni modo anche per questo secondo programma la ricerca evidenzia una serie di limitazioni del sistema di monitoraggio e aree di possibile miglioramento. In entrambi i casi studio, la ricerca sottolinea come la pratica di M&V dei programmi sia solo parzialmente adeguata ai fini dello stesso M&V e, pertanto, non si riescano a generare informazioni adeguate a riguardo dello sviluppo del capitale umano promosso dai programmi;
- capitolo terzo: l'obiettivo del capitolo è quello di colmare un *gap* nella letteratura relativa agli effetti macroeconomici del FSE/IOG, specialmente in relazione al periodo di programmazione 2014-2020 e a degli studi che discutano l'eterogeneità degli effetti in relazione alle diversità dei contesti socio-economici delle regioni coinvolte nei programmi, dei gruppi *target*, e delle forme di supporto offerte. Risponde in particolare alle indicazioni formulate nello studio di Sakkas (2018) sulla necessità di avere "informazioni sugli effetti di livello micro delle misure di supporto all'occupazione finanziate dal Fondo Sociale Europeo" per migliorare le stime prodotte da RHOMOLO, e dell'indicazione di Pellegrini et. Al (2013) sulla necessità di modelli che riconcilino l'approccio micro e quello macro nella valutazione degli effetti delle politiche di coesione, in vista delle limitazioni di entrambi. Poiché i risultati finali del framework analitico sviluppato e proposto nel capitolo 3 vanno al di là del campo di questo progetto di ricerca in quanto potranno divenire disponibili in futuro in collaborazione con il JRC, il focus dell'analisi nel terzo capitolo è sulla percorribilità di tale approccio e della affidabilità delle stime intermedie (di equilibrio parziale) prodotte. Nonostante queste ultime non siano direttamente comparabili con la

letteratura, si ritiene che confermino il potenziale dell'approccio analitico proposto di migliorare significativamente le assunzioni correntemente utilizzate nelle stime attualmente prodotte tramite RHOMOLO ed utilizzate nel *policy making*.

In definitiva, le evidenze prodotte da questo progetto di ricerca suggeriscono che ci sia uno spazio significativo per migliorare il disegno dei sistemi di M&V delle politiche UE per lo sviluppo del capitale umano, in particolare in relazione a:

- gli standard e linee guida proposte dalla UE in merito;
- la comprensività, adeguatezza e rilevanza degli approcci in particolare in relazione alla misurazione del contributo dei programmi allo sviluppo del capitale umano inteso in senso olistico, come per la letteratura più aggiornata;
- la simulazione degli effetti macroeconomici tramite RHOMOLO, in particolare per ciò che concerne le assunzioni utilizzate e i dati di input, che dovrebbero essere sempre più basate su dati reali sugli effetti di equilibrio parziale piuttosto che su assunzioni che (i) sono irrealistiche rispetto alle forme di supporto offerte (ii) non tengono conto dei risultati effettivamente generati dai programmi ma solo il loro costo e gli effetti potenziali di un investimento in qualche modo paragonabile al supporto offerto. Al tempo stesso, la ricerca sottolinea che esistono dei problemi significativi a perseguire strade alternative connessi alla stima degli effetti di equilibrio parziale dei programmi a causa dell'assenza di informazioni controfattuali di livello micro con di sufficiente granularità e qualità.

In ultima istanza, questo progetto cerca di fornire un contributo innovativo al dibattito sul concreto funzionamento del sistema di M&V degli esistenti programmi per lo sviluppo del capitale umano in UE. Questa ambizione è corroborata dall'esperienza lavorativa maturata con Fondazione Giacomo Brodolini nell'ambito delle valutazioni *in itinere* dei programmi FSE, IOG, FEAD e REC and JUST, commissionati dalla Commissione Europea. Questa esperienza lavorativa ha largamente facilitato la comprensione, da un punto di vista privilegiato, delle difficoltà materiali con cui i sistemi di M&V devono confrontarsi nella ricerca di fornire migliori evidenze per il *policy making*, e che spesso implicano l'impossibilità di valutare con certezza se i programmi stiano davvero generando gli effetti desiderati.

Summary

This PhD project is the result of both individual research as well as work experience accrued with Fondazione Giacomo Brodolini in the field of the analysis of the monitoring and evaluation (M&E) systems of the European Union's Human Capital (HC) investments.

The hybrid nature of this PhD project meant that its main objective has been to try to bridge academic knowledge onto the policy evaluation practice, which is often insufficiently informed by the former due to time and budget constraints typical of applied (funded) research. Furthermore, particularly methods and arrangements for the monitoring of HC development policies are either overlooked by or remain at the border of academic debate, with inevitable consequences on the quality of evidence produced by the related evaluation studies. Thus, the rationale of this research is that of analysing the standards, practices and some concrete examples of the functioning of monitoring systems as well as the way in which these inform evaluation studies, to identify incremental improvements based to the extent possible on scientific literature on the subject-matter. In particular, the author believes that improving the assessment of monitoring systems is instrumental in producing information of better quality and granularity which, in turn, will allow adequate evaluation of policies at all levels (micro, meso and macro). Given the hybrid nature of this research project, the audience to whom it is addressed is that of researchers as well as practitioners having to support decision makers, especially at the EU level, in the policy evaluation process. Nonetheless, the goal is also to spur the debate on such topics across a more academic audience, in a quest for better and more robust assessments of M&E systems, given their pivotal role in policy making.

Hence, this research tends to deviate from the standard academic practice in two main ways:

- A broader than average scope of the exercise, which implies that part of the effort is that of creating taxonomies which can help compare and contrast different concepts/approaches/methodologies rather than discussing one specific field of evidence in the literature. This has implications on the style, structure and degree of synthesis of the research presented further below;
- The analysis of subjects generally overlooked by the academic discourse, which means that in some cases the literature is simply not existing and. This, as a result, reduces the comparability of the estimates produced in the experimental part of the research and the extent to which a sensitivity analysis can be carried out at this stage, given the lack of suitable benchmarks. This, on the one hand, means that more research should be devoted to the theme in the future to test the findings of this project and, on the other hand, should spur the discussion as to the reasons why such relevant topics are not consistently addressed in the academic literature. This work is intended as a primer into filling such gap.

Against this premise, this research is in particular concerned with:

- (i) Chapter one: the definition of a much needed analytical tool aimed to help researchers and practitioners examine the EU's monitoring arrangements in the field of HC development policies, which takes proper stock of cumulative knowledge in this domain. The added value of this tool is that despite a rather rich literature discussing the two topics separately might exist (especially for HC), an organic and unitary discussion of M&E for HC development policies in particular is missing especially in the EU domain. This might also

explain the lack of dedicated assessments on this field, as described in the subsequent chapter. To address this issue, the chapter presents a comparative analysis of the ways in which guidelines of supranational and international institutions deal with monitoring and evaluation issues for human capital policies vis-à-vis the relevant literature on the subject-matter, to ensure that the knowledge from the academia spills over to the tools through which the policy cycle is informed and policy decisions are taken;

- (ii) chapter two: two case studies which: (i) examine, based on the newly developed analytical tool of chapter 1, the M&E systems of three major EU programmes that deal with human capital development, namely the European Social Fund (ESF), the Youth Employment Initiative and the Fund for European Aid to the most Deprived (FEAD); and (ii) produce recommendations for the way forward. As anticipated, a pure assessment of the monitoring systems of EU programmes on HC development is missing in the academic literature. On this, there exist studies commissioned by the EC themselves, which, however, either touch only indirectly monitoring issues (i.e. in the frame of impact evaluation) or without relying on comprehensive and dedicated analytical tools. In any event, evidence in this field is weak;
- (iii) chapter three: an experimental research that suggests a new framework for the evaluation of macroeconomic impacts of the ESF/YEI. The research investigates, in particular, whether and how information from the ESF/YEI M&E system on participants reached and the effects the support had on them (partial equilibrium effects), could be used to estimate, in combination with the existing RHOMOLO model of the Joint Research Centre of the European Commission (JRC), macroeconomic impacts of youth employment policies (general equilibrium effects). The reason for this research is that on the one hand, there is a large gap in the literature focusing on macroeconomic effects of the ESF and YEI which also addresses the issues of the existing heterogeneity of effects, heterogeneity of support offered policies as well as their distribution across time and regions – which is essential in the field of place-based policies. On the other hand, the tool currently employed by the European Commission to estimate the abovementioned effects (RHOMOLO model), which could potentially fill this gap, relies on assumptions on the support offered by the ESF/YEI which do not correspond to the activities effectively funded. Hence, the research explores an innovative way of using ESF /YEI monitoring information on youth employment investments which, elaborated econometrically and triangulated with evidence from Counterfactual Impact Evaluation studies, should serve as an input into current research activities carried out by the JRC and the Commission's Directorate-General for Employment, Social Affairs and Inclusion (DG EMPL) on the macroeconomic impacts of the above mentioned policies.

This research project departs from or builds upon the existing literature in the following ways:

- chapter 1 does not contend the academic literature but rather builds upon it to develop an analytical framework for the monitoring of human capital policies along with a tool for its assessment. In doing so, it sheds some light on the shortcomings of the current monitoring practice of the EU, which is not quite suitable to measuring human capital, and on the fact that the EU standards for monitoring and evaluation might be improved in several complementary ways. These include: (i) more emphasis on result-based monitoring and on the synergies between monitoring and evaluation (ii) the need for sector or theme specific

guidance as in the case of human capital, and (iii) a stronger need for differentiation across the micro, meso and macro levels;

- chapter 2 fully aims at filling a gap in the literature which is at best scarce on the assessment of monitoring and evaluation systems; in particular, with respect to the assessment of the ESF/YEI monitoring system, the analysis highlights that the analytical designs used thus far were insufficiently comprehensive or simply lacked depth. This is also due to the fact that these subjects are left to practitioners to deal with despite their scientific relevance. The ramifications of such tendency are particularly evident in the way existing studies have dealt with the issues of ESF/YEI micro-data collection and availability, of the collection of information on typologies of support offered and of the relevance of programme specific indicators or other arrangements to measure support to human capital formation. When it comes to the FEAD, as no dedicated analysis exists, there is no clear benchmark in the literature. Also in this case, some shortcomings of the M&E system are highlighted along with room for improvements. In both cases, it is stressed that the M&E arrangements are only partly fit for purpose and might fail to provide information on HC development.
- Chapter 3 aims to fill a gap in the literature on the macroeconomic effects of the ESF/YEI, especially when it comes to the 2014-2020 programming period and with respect to studies which take into account the heterogeneity of effects across regions, target groups and typologies of support offered. It particularly follows the indication from Sakkas (2018) on the need for “precise information on the micro-level effects of employment support from the ESF” for a better functioning of RHOMOLO, as well as that of Pellegrini et. Al (2013) on the need to better bridge the micro and macro approach to the evaluation of cohesion policy, given their mutual limitations. As the ultimate results of the simulation based on the analytical framework proposed go beyond this research project and will only become available in future research in collaboration with the JRC, the focus of the experimentation is on the viability of the analytical framework suggested and on the reliability of the intermediate (partial-equilibrium) estimates produced. Although these have no clear benchmark in the literature, this research contends that overall the potential of the underlying approach to strongly improve on the assumption used thus far in RHOMOLO is confirmed, especially as information on partial equilibrium effects of better quality is likely to become available in the near future.

Ultimately, the findings of this research project suggest that there’s ample room for improving the design of the M&E systems of EU HC development programmes, especially concerning:

- the EU standards and guidance documents, as indicated above;
- the comprehensiveness, adequacy, relevance and especially suitability to measure the funds’ contribution to human capital development in a holistic (and up-to-date) sense;
- the simulation of macroeconomic effects through RHOMOLO, especially for what concerns the input data used, which should be increasingly based upon actual data on the partial equilibrium/ medium term effects of the programmes rather than on assumptions which (i) can hardly match the support actually offered (ii) cannot take into account the results actually generated by the programmes but only consider their costs and the potential effects of comparable investments. At the same time, the research highlights the significant

struggles and limitations which are linked to the estimation of partial equilibrium effects of the programmes given the current lack of data of sufficient quality and granularity.

Finally, this project seeks to provide a novel contribution to the discourse on the concrete functioning of the monitoring and evaluation systems of existing programmes for human capital investments in the EU. Such an ambition is corroborated by the working experience accrued with Fondazione Giacomo Brodolini on the actual mid-term evaluations of the ESF, YEI, FEAD and REC and JUST programmes commissioned by the European Commission. This aided the understanding, from a privileged standing point, of the concrete issues which mean that the costly M&E systems of EU programmes might fail to provide decisive information on whether programmes deliver on their intended outcomes.

Table of Contents

RINGRAZIAMENTI	3
SINTESI	5
SUMMARY	9
TABLE OF CONTENTS	13
LIST OF ACRONYMS	17
LIST OF EU-28 MEMBER STATES	21
INTRODUCTION	22
1 CHAPTER 1 – MONITORING AND EVALUATION OF HUMAN CAPITAL POLICIES: A FRAMEWORK TO CHECK WHETHER IT IS FIT FOR PURPOSE.....	25
Abstract	25
Introduction	25
Motivation for the research and gaps in the literature	25
Structure of the chapter	26
1.1 The purpose and key defining elements of monitoring and evaluation	27
1.1.1 Monitoring.....	27
1.1.2 Evaluation.....	30
1.1.3 Results-based M&E: how monitoring systems differ from/interact with impact evaluations	37
1.1.4 Summing up: key principles and features of monitoring and evaluation	39
1.2 Human Capital: what is it and how should we measure it.....	41
1.2.1 What is human capital?	42
1.2.2 How can we measure human capital?	45
1.2.3 Human capital Indexes	48
1.3 Monitoring and evaluation of HC policies	52
1.3.1 Monitoring.....	53
1.3.2 Evaluation.....	55
1.4 Concluding remarks and monitoring of HC checklist	57
1.4.1 Assessment grid for the monitoring systems of Human Capital Policies.....	57
1.4.2 Concluding remarks	60
1.5 Main Bibliography	62
2 CHAPTER 2: CASE STUDY ANALYSIS ON THE MONITORING AND EVALUATION SYSTEMS OF SELECTED EU PROGRAMMES FOR HUMAN CAPITAL DEVELOPMENT.....	70
Abstract	70
Introduction	70
Motivation for the research and gaps identified in the literature.....	70

	Structure of the chapter	71
	Specificities of this chapter.....	71
2.1	HC in the EU and selection of case studies	72
2.2	The ESF/ YEI monitoring system.....	74
2.2.1	Introduction and key info on the structure and intervention logic of the ESF/ YEI	74
2.2.2	ESF /YEI monitoring system: main features and indicators	76
2.2.3	Intervention logic of the ESF and related monitoring arrangements.....	90
2.2.4	Focus: the ESF and the support it can provide to measure progress towards the goals of the New Skills Agenda for Europe (NSA).....	91
2.2.5	Main issues with the ESF/ YEI monitoring system	94
2.2.6	How does the ESF/YEI monitoring system fare vis-à-vis the Monitoring of HC checklists.....	97
2.2.7	The future of the ESF monitoring system	103
2.2.8	Conclusions and recommendations	105
2.3	Monitoring of the Fund for the Fund for European Aid to the Most Deprived (FEAD).....	107
2.3.1	Main goals and structure of the fund.....	107
2.3.2	Monitoring and evaluation arrangements.....	109
2.3.1	FEAD's intervention logic	112
2.3.2	Main issues with the FEAD Monitoring and Evaluation system as emerging from the literature and work in the MTE evaluation	115
2.3.3	The FEAD vis-à-vis the Monitoring of HC checklist	118
2.3.4	FEAD in the ESF+ and its foreseen monitoring.....	123
2.3.5	Key outstanding issues and recommendations	124
2.4	Conclusions	125
2.5	Main Bibliography.....	128
3	CHAPTER 3: A NEW FRAMEWORK FOR THE ESTIMATION OF MACROECONOMIC IMPACTS OF THE EUROPEAN SOCIAL FUND AND YOUTH EMPLOYMENT INITIATIVE: STRENGTHS, WEAKNESSES AND ROOM FOR IMPROVEMENTS.....	132
	Abstract	132
	Introduction and structure of the chapter.....	132
	Genesis of the research.....	132
	Motives and key elements of the research.....	134
	Structure of the chapter	136
3.1	Background: review of the literature and description of the RHOMOLO model	137
3.1.1	Brief survey on the literature on the impacts of cohesion policy in general, and on the ESF/YEI and active labour market policies in particular	137
3.1.2	Key features of RHOMOLO, JRC's Regional HOListic MOdel.....	142
3.2	The analytical framework for this simulation exercise.....	147

3.2.1	How to model HC investments in RHOMOLO	147
3.2.2	ESF monitoring data and main issues with a view to the simulation	149
3.2.3	Analytical framework and methods	152
3.3	Methods and methodological steps.....	156
3.3.1	Overview	156
3.3.2	Data used and how to get them	157
3.3.3	Econometric estimation of conditional distributions of participants characteristics.....	158
3.3.4	Quali-quantitative identification of a proxy for the level of implementation of the OPs	173
3.3.5	Transformation of ESF data in NUTS2 level info.....	179
3.3.6	Identifying coefficients to estimate net effects from Counterfactual Impact Assessment studies	179
3.3.7	Discounting values by LM transition data	188
3.4	Discussing what the estimated data (input for RHOMOLO) looks like	189
3.4.1	Data on participation	190
3.4.2	Data on changes of labour market status.....	192
3.5	Concluding remarks.....	197
3.5.1	Strengths of the approach and how this improves upon the literature.....	198
3.5.2	Weaknesses and room for development.....	200
3.5.3	Policy recommendations	201
3.6	Bibliography	203
4	CONCLUSIONS	209
5	ANNEX I - EU POLICIES FOR HUMAN CAPITAL DEVELOPMENT – AN OVERVIEW	211
5.1	Introduction and an operative definition of human capital.....	211
5.2	The scope for EU action in the field of human capital development.....	213
5.3	Europe 2020.....	216
5.3.1	EU 2020 main targets and flagship initiatives.....	216
5.3.2	Additional targets within the frame of the EU 2020 strategy.....	217
5.4	The current political guidelines for the EC: 'A New Start for Europe: My Agenda for Jobs, Growth Fairness and Democratic Change'	219
5.5	A new skills agenda for Europe and the European Social Fund.....	220
5.5.1	A New Skills Agenda for Europe - introduction.....	220
5.6	The European Pillar of Social Rights and the Social Scoreboard.....	222
5.6.1	Key principles and human capital development.....	222
5.6.2	The JRC's social scoreboard	223
5.7	Additional policy level arrangements for the monitoring of HC initiatives	224
5.8	The Multiannual Financial Framework and its programmes: how is human capital development funded in the 2014-2020 programming period	225

6	ANNEX II - MONITORING OF THE EU HUMAN CAPITAL DEVELOPMENT POLICY: A CASE STUDY ON THE NEW SKILLS AGENDA FOR EUROPE AND THE ESF'S MONITORING SYSTEM	234
6.1	Introduction	234
6.1.1	Where the NSA and the ESF aims align	234
6.1.2	Structure of the NSA and relevant actions	235
6.2	Analysis of ESF Indicators and their suitability to monitor the NSA.....	239
6.2.1	Approach to the analysis and keywords selected	239
6.2.2	Findings.....	242
6.3	Conclusions	254
7	ANNEX III – ADDITIONAL DETAILS ON THE EXPERIMENTATION AND ITS BACKGROUND.....	257
7.1	A brief description of the RHOMOLO model.....	257
7.1.1	Background: a brief taxonomy of General Equilibrium models	257
7.1.2	The JRC's RHOMOLO model - main features	258
7.2	Detailed procedure for bringing ESF monitoring data at the NUTS2 level.....	261
7.2.1	Reconstructing eligibility of EU NUTS2 regions for the cohesion policy 2014-2020.....	262
7.2.2	Dealing with different NUTS classification issues	262
7.2.3	Retrieving AIR Table 7 data on the spend at the NUTS0 to 3 data	262
7.2.4	Identifying those Operational Programmes that showed expenditure at the NUTS0 or 1 level	263
7.2.5	Identifying criteria to breakdown spend at the NUTS2 level.....	263
7.2.6	Writing algorithms (excel macros, VBA) which could split automatically distribute expenditure at the NUTS2 level based on the above criteria	263
7.2.7	Testing the resulting figures at the NUTS2 level (so that they matched the original values).....	263
7.2.8	Re-aggregating expenditure by NUTS2 region (different OPs contribute to the same region) and computing shares.....	264
7.2.9	Calculating output and results based on the relevant NUTS2 shares.....	264

List of Acronyms

List of Acronyms	
ADB	Asian Development Bank
AIR	Annual Implementation Report
ALMP	Active Labour Market Policies
AR	Achievement Rate
AROPE	At Risk of Poverty or Social Exclusion
BLUP	Best Linear Unbiased Predictor
BRG	Better Regulation Guidelines
C-VET	Continuing education and training
CEDEFOP	European Centre for the Development of Vocational Training
CES	Common Elasticity of Substitution
CGE	Computational General Equilibrium
CIA	Common Independence Assumption
CIE	Counterfactual impact evaluation
COI	Common Output Indicators
COM	European Commission
CP	Cohesion Policy
CR	Common Result
CRI	Common Result Indicators
DB	database
DEReC	DAC Evaluation Resource Centre
DFID	British Department for International Development
DG EMPL	Directorate-General for Employment, Social Affairs and Inclusion
DG REGIO	Directorate-General for Regional and Urban Policies
DGs	Directorates-General of the European Commission
DSGE	Dynamic Stochastic General Equilibrium
E&T	Education and Training
EaSI	EU Programme for Employment and Social Innovation
EC	European Commission
ECA	European Court of Auditors
EGF	European Globalisation Adjustment Fund
EP	European Parliament
EPSR	European Pillar for Social Rights
ESF	European Social Fund
ESF+	Umbrella Fund for the 2021-2027 MFF including ESF, YEI, FEAD, EaSI, Integration of Migrants and Health programme
ESIF	European Structural Investment Funds
ESL	Early School Leaving

List of Acronyms

EU	European Union
EU 2020	Europe 2020 strategy
EU-28	the 28 Member States of the European Union as in 2019
EVALSED	Evaluation of Socio-Economic Development
FEAD	Fund for European Aid to the most Deprived
GDP	Gross Domestic Product
GE	General Equilibrium
HC	Human Capital
HCI	Human Capital Index
HS	High Skilled
I-VET	Initial education and training
IL	Intervention Logic
ILO	International Labour Organisation
IP	Investment Priority
ISCED	International Standard Classification of Education
IZA	Institute for Labour Economics
JRC	Joint Research Centre of the European Commission
JUST	Justice Programme
LLL	Lifelong Learning
LM	Labour Market
LMP	Labour Market Policies
LS	Low Skilled
LTU	Long-Term Unemployed
M&E	Monitoring and Evaluation
MA	Managing Authority
MFF	Multiannual Financial Framework
MS	Medium Skilled
MTE	Mid-Term Evaluation
NACE	<i>Nomenclature statistique des activités économiques dans la Communauté européenne</i> (Statistical classification of economic activities in the European Community)
NEET	Not in Employment, Education or Training
NGO	Non-Governmental Organisation
NSA	New Skills Agenda
NUTS	Nomenclature of Territorial Units for Statistics
NUTS 1	Major socio-economic regions
NUTS 2	Basic regions for the application of regional policies
NUTS 3	Small regions for specific diagnoses
OECD	Organisation for Economic Cooperation and Development

List of Acronyms

OECD-DAC	OECD Development Assistance Committee
OP	Operational Programme
PA	Public Administration
PAX	Priority Axis
PES	Public Employment Service
PIAAC	Programme for the International Assessment of Adult Competencies)
PISA	Programme for International Student Assessment
PO	Partners Organisation
QUEST III	Quarterly European Simulation Tool
R&D	Research and Development
RCT	Randomised Control Trials
RDD	Regression Discontinuity Design
REC	Rights Equality and Citizenship Programme
RHOMOLO	Regional HOlistic MOdeL
RSB	Regulatory Scrutiny Board
SFC	System for Fund Management
SI	Share of Inactive
SLS	Share of Low Skilled
SME	Small and Medium Enterprise
SNA	System of National Accounts
SO	Specific Objective
SRE	Success Rate Employment
SRJS	Success Rate Job Seeking
STEM	Science, Technology, Engineering and Mathematics
SUTVA	Stable Unit Treatment Value Assumption
TBIE	Theory-based impact evaluation
TFEU	Treaty on the Functioning of the European Union
TFP	Total Factor Productivity
TG	Target Group
TKP	Total Capital Productivity
TLP	Total labour Productivity
TO	Thematic Objective
TO8	Thematic Objective 8 ‘Promoting sustainable and quality employment and supporting labour mobility’
TO9	Thematic Objective 9 ‘Promoting social inclusion, combating poverty and any discrimination’
TO10	Thematic Objective 10 ‘Investing in education, training and vocational training for skills and lifelong learning’
ToC	Theory of Change

List of Acronyms

TRANSTOOLS	TOOLS for TRansport Forecasting ANd Scenario testing
UNDP	United Nations Development Programme
UNECE	United National Economic Committee for Europe
UP	Upskilling Pathways
VBA	Visual Basic for Applications
VET	Vocational and Educational Training
WB	World Bank
WBL	Work-based learning
WEF	World Economic Forum
YE	Youth Employment
YEI	Youth Employment Initiative

List of EU-28 Member States

List of EU-28 Member States

AT	Austria
BE	Belgium
BG	Bulgaria
CY	Cyprus
CZ	Czechia
DE	Germany
DK	Denmark
EE	Estonia
EL	Greece
ES	Spain
FI	Finland
FR	France
HU	Hungary
HR	Croatia
IE	Ireland
IT	Italy
LT	Lithuania
LU	Luxembourg
LV	Latvia
MT	Malta
NL	the Netherlands
PL	Poland
PT	Portugal
RO	Romania
SE	Sweden
SI	Slovenia
SK	Slovakia
UK	United Kingdom

Introduction

This project is concerned with EU policy initiatives which aim at fostering human capital in the EU. In particular, its focus lies with the EU Monitoring and Evaluation (M&E) practice of their related policy initiatives and funding instruments.

The overarching goal of this project is to ultimately assess whether the EU monitoring and evaluation practice in the field of human capital development is fit for purpose, including identifying where and how improvements could be made.

In terms of geographical scope, it focuses on the EU-28, but looks beyond the EU to take stock of the monitoring and evaluation practice especially of other international organisations.

Time-wise, it focuses on the Multiannual Financial Framework (MFF) 2014-2020 (thus approximately from 2014 to 2023), but the analysis of the relevant literature includes previous programming periods.

In essence, this project presents:

- In chapter 1, a review of the concept of human capital (HC) along with the concepts and standards of M&E which is aimed at defining a novel and comprehensive framework for the analysis of the monitoring systems of HC policies that, although being addressed by practitioners typically dealing with such issues, is fully rooted in academic knowledge;
- In chapter 2, a novel analysis of the monitoring systems of two of the major funds for HC development in the EU, that is based on the analytical framework developed in chapter one and aims to make up for the scarcity of dedicated studies and literature on the themes. At the same time, it wishes to propose policy recommendations aimed at steering the current M&E practice towards an improved understanding of Human Capital measurement issues;
- In chapter 3, an experimental research on the contribution of EU support to youth employment, which in view of the shortcomings of micro and macro-econometric approaches aims to propose and test a new framework for the analysis of ESF/YEI macroeconomic impacts through the general equilibrium model RHOMOLO, bridging information on microeconomic effects of the funds onto it.

The Project structure is envisaged as follows:

- Chapter 1 – how to best monitor and evaluate human capital policies
 - Monitoring and evaluation: purpose and approaches
 - Brief overview of the concept of human capital in the literature, as well as review of existing methods to measure it
 - Review of M&E guidelines and related academic literature with a focus on the M&E of human capital policies
 - Definition of a checklist for the monitoring of HC initiatives, to serve as an analytical tool for the case study analysis in chapter 2 and conclusions
- Chapter 2 – how is HC development monitored and evaluated in the EU: case study analysis

- Human capital development in the EU: identification of an operational definition of HC to discuss the legal basis, key strategies and EU monitoring arrangements in this field
- The European Social Fund (ESF)/Youth Employment Initiative (YEI) M&E systems
- The Fund for European Aid to the Most Deprived (FEAD) M&E system
- Conclusions
- Chapter 3 – assessing the macroeconomic implications of human capital investments:
 - Literature review on the existing evaluations of cohesion policy in the EU with a focus on the ESF/YEI and brief description of the EC’s RHOMOLO model
 - Definition of an analytical framework for the assessment of macroeconomic impacts through RHOMOLO and identification of the experimental part of this research (limited to preparing input data for RHOMOLO on youth employment investments)
 - Description of the steps, assumptions and data elaboration process, including the related econometric analysis and a review of available Counterfactual Impact Evaluation studies on YE measures under the ESF/YEI
 - Description of the resulting input data for RHOMOLO (data on participants’ features and their related net partial equilibrium effects)
 - Discussion on the strengths and weakness of the approach, room for developments and possible alternative methods

This structure is instrumental to answering the following research questions:

- Chapter 1 (ground-clearing exercise and literature review):
 - How should we monitor and evaluate policies which aim to develop human capital, especially in the EU?
- Chapter 2 (case study analysis):
 - How is the EU dealing with the development of Human Capital and its monitoring at the policy level?
 - Is the M&E system of the ESF/YEI fit for purpose?
 - Can the M&E system of the ESF/YEI help understand whether we are making progress towards the objectives of the New Skills Agenda for Europe?
 - Is the M&E system of the FEAD fit for purpose?
- Chapter 3: A new framework for the estimation of macroeconomic impacts of the European Social Fund and Youth Employment Initiative: strengths, weaknesses and room for improvements
 - What does the literature find on the macroeconomic effects of cohesion policy, and through which methods? What are their main shortcomings?

- How were macroeconomic impacts of the ESF/YEI estimated so far and what are the main shortcomings?
- Could ESF/YEI M&E data be used in combination with the RHOMOLO model to better assess the macroeconomic impacts of the ESF/YEI? An experiment with YE policies.

The underlying rationale for this structure is that in order to deal with such broad and complex concepts there's a need to begin with understanding: (i) what is to be expected from monitoring and evaluation systems; (ii) what is exactly human capital development and, as a result (iii) which are the relevant objectives and policies that should form part of this assessment and they key policies in the field. Ultimately (iv) how to monitor in particular HC investments.

Each chapter is in fact intrinsically linked one another: only based on chapter 1 can a proper assessment of the M&E practice of the ESF/YEI and FEAD be made. And only on in-depth knowledge of the ESF/YEI can an experiment with its monitoring and evaluation data be based. At the same time, trying to understand whether and how information from the M&E system is suitable for an assessment of the ultimate overall impacts of one of the key priorities of such funds (youth employment) fully feeds back into the assessment of whether information generated by the M&E system is fit for purpose. Several shortcomings in fact appear when ESF/YEI M&E data is used in combination with the tools and methods used for impact evaluation, especially at the macroeconomic level.

Finally, chapter 4 presents some concluding remarks better highlighting interlinks among chapters and some streamlined suggestions for the way forward.

1 Chapter 1 – Monitoring and Evaluation of human capital policies: a framework to check whether it is fit for purpose

Abstract

Monitoring and evaluation (M&E) is fundamental to good policy, hence the great heed being paid by public institutions and researchers alike to developing its practice. Human capital (HC) development, in comparison, has attracted even more attention and research, in the light of its essential role in the creation of wealth as well as social cohesion. Despite the proliferation of guidance documents on how to best design and implement M&E systems and research as well as on the measurement of human capital separately, scarce attention was devoted, to the best of the author's knowledge, to the specificities of M&E in the specific field of HC policies. This can affect, in turn, the suitability of existing M&E systems to examine success of HC policies. But in order to improve on existing tools for the design and analysis of a M&E system, an in-depth consideration of what is human capital and how to measure it is needed. Thus, a comprehensive survey of the literature on both M&E and HC is presented. This paves the way for the creation of a tool intended to swiftly support practitioners and researchers in the design/analysis of M&E systems which deal with HC development. The tool, which is an assessment grid/ checklist devised with a view to the EU HC policies and to support the work carried out in chapter 2, builds upon the EU's better regulation guidelines, innovating and complementing its contents with suggestions from both other international institutions as well as a brand-new section on the measurement of human capital based on an analysis of the relevant debate and recent "human capital indexes" initiatives. Finally, the analysis highlights a few overlooked shortcomings in the EU standards for M&E, especially with respect to HC policies. This has consequences, as shown in chapter 2, on the evidence which can be produced about such initiatives.

Introduction

Monitoring and Evaluation is largely acknowledged as a key component of policy making. Result-orientation is pivotal to ensuring that available resources are used at their optimum to tackle societal challenges of all sorts and kinds. Achieving results means not only to address issues that affect the lives of those concerned, but also to reinforce confidence in public action, strengthen social cohesion and, ultimately, ensure prosperity.

This is why significant shares of public resources are invested into creating and managing monitoring and evaluation systems which should ultimately help us understand what works, where and for whom.

Motivation for the research and gaps in the literature

Answering these questions though is anything but trivial. Environmental factors, including the inherent complexity of some policy objectives or domains, mean that data collection is essential not only to managers who need to know on time whether things are going according to plan but especially to evaluators that need to examine causality and understand whether the investments are worth continuing.

Human capital development is one prominent example of this. Over the last three centuries, researchers from all over the world have studied, examined and discussed the concept of human capital, tried to measure its distribution and role in contributing to individual, firm and societal well-being, as well as sought to assess which policies work best to attract and develop human capital. This has led to increased awareness on the importance of human capital, but also to a progressive broadening of its scope, to eventually include “*the knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic well-being*” (OECD, 2001).

The multidimensionality of the concept of human capital means that also knowledge on how to monitor and evaluate policies which aim to develop it is fragmented and difficult to systematise. This is conducive to sub-optimal quality of information feeding the related policy cycle, with the negative repercussions alluded to above. Such dynamic is particularly apparent if one thinks of the process through which, e.g., EU programmes are devised. Needs change rapidly and it can be very hard for M&E to detect, understand and find solutions on how to ensure proper implementation whilst finding how to adapt to these, in a way that is coordinated with the policy cycle and sustainable. That is where the need for easy-to-understand and easy-to-handle tools for the analysis of a M&E system is more acute. But existing frameworks and guidelines are typically not developed in a way that is immediately relevant to human capital development. In addition, the creation of analytical tools such as assessment grids is often left to the companies to which studies are contracted out, typically under time and resource constraints.

Against this background, this chapter aims to create a tool that, taking stock of the vast literature on the theme of HC and M&E, can swiftly support practitioners and researchers in the design/analysis of M&E systems which deal with human capital development. In doing so, In doing so, it sheds some light on the shortcomings of the current monitoring practice of the EU, which is not quite suitable to measuring human capital, and on the fact that the EU standards for monitoring and evaluation might be improved in several complementary ways. These include: (i) more emphasis on result-based monitoring and on the synergies between monitoring and evaluation (ii) the need for sector or theme specific guidance as in the case of human capital, and (iii) a stronger need for differentiation across the micro, meso and macro levels.

Structure of the chapter

This chapter is a ground clearing exercise. It lies the foundation for the whole additional research carried out in this project. Hence, the starting point (section 1.1) is that of properly defining monitoring and evaluation, looking both within and beyond the EU practice and standards in the field. Such knowledge is elaborated and summarised in particular in section 1.1.4.

Next, the focus is shifted to the concept and key dimensions of human capital, including focusing on its measurement and the most updated indexes which have been developed (section 1.2). Here, the multidimensionality of the HC concept is apparent and so is the fact that this is not fully reflected in the EU M&E handbooks and guidance documents, which tend to overlook the issue.

As the issues of M&E and that of Human Capital measurement/assessment have seldom been discussed jointly from a holistic perspective, section 1.3 seeks to fill this gap based on the

elements discussed in section 1.1 and 1.2. Here, an analysis of the elements and key lessons which should be taken into account and properly addressed by a M&E system that covers HC development is provided.

Lastly, section 1.4 presents the dedicated tool for the assessment of HC monitoring systems in the EU. It builds upon the EU better regulation guidelines but innovates them in several ways based on the experience gained working with the monitoring systems (including the lessons learned from chapter 2 and 3) and by creating an ad-hoc module dedicated to the suitability of the monitoring system to collect information suitable to human capital development. It also summarises the key conclusions on the current standards for monitoring and evaluation of HC policies in the EU.

1.1 The purpose and key defining elements of monitoring and evaluation

The starting point for a reasoning on Monitoring and Evaluation system is inevitably that of understanding the purpose of monitoring and evaluation. Over the last decades the issue was widely debated and there is some consensus now as to these concepts. This is not to say, however, that their key principles are consistently applied in policy making. Below the two are described based on a number of guidance documents (including the European Commission's Better Regulation guidelines (hereafter BRG), the OECD Development Assistance Committee (OECD-DAC)/ International Labour Organisation (ILO), World Bank, but also a range of policy documents of the OECD national or thematic agencies) screened with a view to identifying some consolidated definitions as well as key required elements of their design.

1.1.1 Monitoring

1.1.1.1 European Commission's Better Regulation Guidelines

The reference document for monitoring and evaluation in Europe are the Better Regulation Guidelines, which both internal and external professionals need to comply to in the conceptualisation, implementation and review of the monitoring and evaluation of EU policies. They establish the following understanding of monitoring and its purpose:

“Monitoring generates evidence on an intervention's activities and impacts over time in a continuous and systematic way. A monitoring system is a necessary and integral part of Better Regulation helping to:

- Identify whether a policy is being applied on the ground as expected;
- Addressing any implementation problems of an intervention; and/or
- Identifying whether further action is required to ensure that it can achieve its intended objectives.” (European Commission, 2015a:44)

Given the important role of monitoring in generating timely and relevant evidence for evaluating policies, this needs to be fit for purpose. The Better Regulation Guidelines establish **three key questions** that need to be considered for the design of the monitoring system, namely:

1. *What evidence needs to be collected?*
2. *When and how should evidence be collected?*
3. *Who will collect the evidence and from whom?* (European Commission, 2015a: 45)

In answering these questions, an in depth consideration of the “intervention logic“ of the policy needs to be established. The intervention logic is where the link between needs identified, activities supported by the policies, its objectives and how these impact on the original needs is clarified.

Secondly, bearing these questions in mind, when designing the monitoring system, several **governing principles** must also be adhered to according to the Better Regulation Guidelines and in particular that the system be:

- **Comprehensive:** which means to gather data on all the relevant progress which is logically linked with the objectives of the policy, especially the specific and operational objectives (whilst for the broad ones it is likely that data already exists). This is to avoid missing crucial parts of the contribution the policy is bringing about.
- **Proportionate:** a monitoring comes at a cost, an effort is needed to focus on the most relevant aspects of the policy.
- **Minimise overlap:** to reduce administrative burden, the data which is already existing needs to be used, if compatible, instead of additional data collection.
- **Timely:** a lean monitoring system means also knowing when one should measure what, bearing in mind the timescales of the interventions, the evaluation or publicity needs and so forth.
- **Accessible:** data needs to be made accessible as long as it complies with privacy regulation as it both increases transparency and allow new knowledge to be produced which can efficiently improve policy making.

When designing and conducting monitoring in practice, these governing principles and the three key questions cited above, must be taken together, with the following implications:

- **Question 1: What evidence needs to be collected?**

This includes clarifying the intervention logic and what needs to be measured. Once appropriate indicators are defined, their metrics need to be decided (unit of measure, baseline) together with information on when data needs to be collected and how. Targets need also to be defined, at all levels (input, output, outcome).

Aspects to be considered include, as above, comprehensiveness and proportionality, whilst taking into account both the needs of the managers (who need to know whether things are going according to plan) and of the evaluators (who need to have the relevant data to measure effects). This is not just important from an administrative burden perspective but also, as clarified by the EVALSED sourcebook (DG Regio, 2013:87), because “*it is generally accepted that each actor requires an operating report with a small number of indicators, selected as the most relevant in relation to the nature of the decisions that have to be made. It has been shown that in a situation of decision-making, a person cannot take into account more than about ten indicators at once. When there are too many indicators decision-makers are swamped with an excess of information*”.

- **Question 2: When and how should evidence be collected?**

Not all data needs to be collected necessarily at all times. Again, data needs ought to be explicitly considered, also in choosing between new/ continuous data collection and systematic

(at given point in times)/ no new data collection at all. The goal is to strike a balance between informativeness of the data collection and administrative burden.

- Question 3: Who will collect evidence from whom?

The nature of information needed may strongly affect procedures to collect the data, its quality, timeliness, the intensity of administrative burden which can be accepted and so forth. So it is paramount also to consider this aspect of the monitoring system.

1.1.1.2 Development Assistance Committee of the OECD (OECD DAC) and International Labour Organisation (ILO)

Within the OECD DAC material for practitioners it is possible to identify some key definitions of monitoring, which are generally consistent with what seen in the case of the BRG. In particular monitoring is defined as “*A continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds.*” (OECD/DAC, 2002:27)

The British Department for International Development (DFID), in interpreting the OECD/DAC principles places some more emphasis on the distinct but complementary roles of monitoring and evaluation.

“Monitoring is a continuous process of systematic data collection to inform managers and key stakeholders on progress in relation to planned inputs, activities and results, as well as the use of allocated resources. Monitoring is structured around indicators, which are the measures of performance of the input, activity or results (output or outcome). Indicators target provide the benchmarks against which progress is monitored. Monitoring takes places to enable managers to rapidly identify problems and make the necessary correction to ensure proper implementation. Information from systematic monitoring serves as a critical input to evaluation but does not look in-depth at the causes behind why results are being delivered or not”. (DFID, 2005:22)

1.1.1.3 United Nations Development Programme - UNDP

Within the UNDP, monitoring is considered as also “outcome” monitoring, in the sense that “*Outcome evaluation depends on good monitoring. In the past, monitoring focused on UNDP inputs and activities. In the future, monitoring information given to outcome evaluators will be focused more at the level of outputs and outcomes*” (UNDP, 2002: 11). In particular for the UNDP, “*“good monitoring” means that monitoring is continuous, involves partners, and is focused on progress towards outcomes*”. The UNDP also stresses the importance of monitoring being participated, and clarifies it that this entails “*obtaining feedback from partners and beneficiaries on progress and proposed actions*”. In the monitoring practice, not only the relationship between input, output and outcome should be sound, but it should also be traced whether there are any factors contributing or impeding achievement of the outcomes.

1.1.1.4 World Bank

In World Bank’s “Ten steps for a results-based monitoring and evaluation system” handbook (Kusek & Rist, 2004) this emphasis on the “results” aspects of the monitoring is reiterated and some other relevant elements are stressed, including:

- the fact that monitoring differs depending on the level at which this is applied, and notably project, programme or policy level;
- the difference between M&E, through the definition “Monitoring gives information on where a policy, program, or project is at any given time (and over time) relative to respective targets and outcomes. It is descriptive in intent. Evaluation gives evidence of why targets and outcomes are or are not being achieved. It seeks to address issues of causality.” (Kusek & Rist, 2004:13)

The authors also report the following needs for constructing indicators:

- to involve many experts and stakeholders that can assess their feasibility from both a technical and substantial point of view;
- to stay close to policy objectives;
- to take into account different perspectives, first and second order conditions (so not only direct but also indirect consequences of the policy, even beyond its statutory beneficiaries);
- to keep the system stable: indicators can (and sometimes must) change over time but they should not be changed or switched too often (and never on a whim), as this can lead to chaos in the overall data collection system; and
- To take into account the fact that “constructing indicators takes work” and comes at a cost.

So overall, whilst there is agreement on the key aspects of monitoring, there are some differences emerging among international institutions, especially with respect to the emphasis given to certain aspects of it. There are in fact some important novelties with respect to the EU better regulation guidelines, which will be considered for the development of the monitoring checklist, particularly with respect to the separate but interlinked roles of monitoring and evaluation, the emphasis on result orientation and the need to take into account both direct and indirect impacts, at the different levels at which the policy impacts.

1.1.2 Evaluation

Policy evaluation is a key tool for policy makers around the world whose aim is to ultimately understand what works, for whom and where, with a view to improving the effectiveness, efficiency, coherence, relevance, added value and sustainability of policy action.

As the policy cycle is a fundamentally iterative process, evaluations can produce evidence which may be built upon to amend, adapt, repeal or renovate rules and policies on a continuous basis.

Policy Evaluation in the EU is informed by the principles enshrined in the Better Regulation Guidelines, amongst other, and further underpinned by a range of handbooks and guidance documents that have been produced throughout the years by the different Directorate Generals (DGs), including the DG for Regional and Urban Policy - DG REGIO.¹

¹ See for example EVALSED guidelines and toolbox

Although the features of a policy evaluation can vary a lot depending upon (i) the time at which the evaluation is carried out (e.g., *in-itinere* vs. *ex post*), (ii) the size of the intervention (supra-national, national, regional, programme/project etc.) (iii) the topic of the evaluation and constraints such as data availability, there are essentially two main types of evaluations acknowledged in the BRG (and in the policy evaluation practice at large): theory based impact evaluation (TBIE methods) and counterfactual impact evaluation (CIE methods).

TBIE methods focus on ‘why’ a certain change is triggered by a given policy aiming at it, that is, they seek to reconstruct the logical chain of elements which are set in motion by the policy and should ultimately contribute to the desired change. The focus here is on the contribution of the policy rather than on the attribution question. This type of evaluations are key for two main reasons:

- these are the only methods that look into why a given change is produced although they cannot measure to what extent a given change is attributable to a certain cause (i.e. policy intervention); and
- often times these are the only option on the ground as the statistical assumptions on which rigorous and reliable counterfactual methods are based fail to hold in real-world applied research.

CIE methods, which are meant to answer an essentially quantitative question, focus on ‘how much’ – if anything – a certain change that is observed can be attributed to a given policy (programme, initiative, etc.). As opposed to TBIE, CIE do not tell anything about why such changed occurred but may help reliably single out the net effects of an intervention, which is also the basis for cost-effectiveness analysis.

1.1.2.1 Theory based- impact evaluation (TBIE)

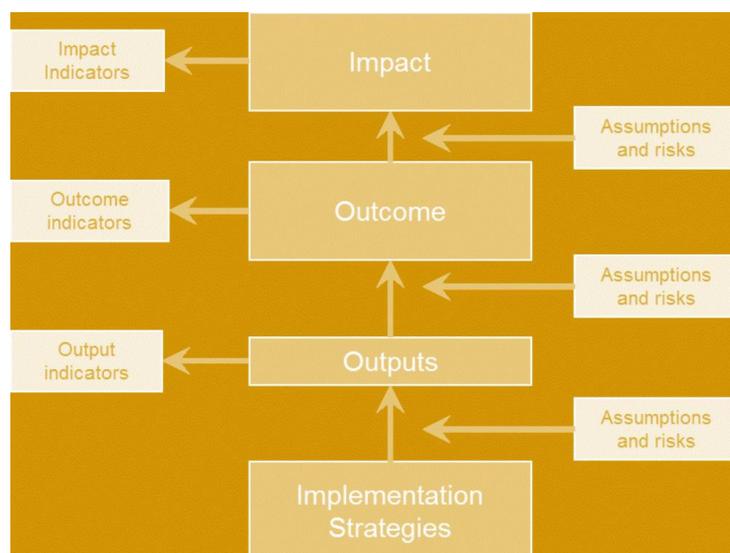
TBIE methods are very common place in real word research, as they are flexible instruments which can help gather evidence across a range of different topics and policy interventions. Within the family of TBIE, many are the specific techniques that may be applied, including Theory of Change, Realist Evaluation, General Elimination Methodology, Process Tracing, Contribution Analysis. The common ground for these methods is that of understanding the causal chain connecting the intervention under analysis to observed changes. For this reason, below is briefly presented what it is meant by Intervention Logic (IL) and Theory of Change (ToC), as these constitute the backbone of TBIE methods.

The application of ToC and IL is aimed at identifying the logical connections between activities, outputs, results and impacts in the form of informed assumptions on how and why a Programme is expected to achieve its stated results and impacts by means of the implementation of a series of activities leading to certain outputs. These allow researchers to identify the reasons for these results, and the factors that contribute to the success or failure (or limited success) of certain approaches in different situations. These logical connections can be tested through the collection of evidence from quali-quantitative analysis and allow draw some conclusions on what works and what does not work – or has not worked to date - for different types of activities and target groups.

ToC can be developed for any level of intervention – a project, a Programme strand, a Programme or a policy/strategy. It is a key tool of the Theory-based evaluation (TBIE) which

explains how and why an intervention will work and will lead to the intended outcomes (see the Figure below).

Figure 1 - Theory of change diagram



Source: Rogers P. (2014) *A theory of change*

In the case of a European programme, the Theory of change can explain how allocating funds (inputs) will produce outputs and how these in turn will produce intended short and medium term results (linked to operational objectives as defined in the different calls/ Annual work plans) and impacts (linked to the specific and general objectives).

TBIE follows each step of the intervention logic identifying causal links and mechanisms of change. Thus TBIE entails an ‘effect logic’, i.e. what happens to outputs with certain inputs and how do outputs lead to results and impacts. The implementation strategies are the drivers linked to the inputs which may include financial support for the implementation of actions falling under a given specific objective or for supporting the work of organisations or other entities pursuing objectives of a programme. The beneficiaries of financial support adapt to these drivers by undertaking a series of activities (eligible according to the type of support provided). These will then lead to changes in people’s behaviours/situations, expressed as results. The final step in this sequence is the impact in the area of analysis as defined by the IL.

In any event, interviews with privileged actors (i.e. those having access/ exposure to key information) and more generally a comprehensive stakeholders’ consultation are the central tools for TBIE. Techniques that systematically review and collate evidence together are then applied to come to clear and objective conclusions and recommendations.

1.1.2.2 Counterfactual impact evaluation (CIE)

The general need for reconstructing a counter-factual situation that can help distil changes that are caused by a specific policy intervention from changes that would have occurred even in absence of the given intervention is rooted in the core methodology for Impact Evaluation. The underlying rationale is that a simple observation of changes generated by an intervention through monitoring outcomes may be affected by a number of factors that are unrelated to the intervention. The variation of an indicator can well capture what happened *while* the

intervention was taking place but can hardly provide evidence to identify what happened *because of* the intervention (Barca, 2009). Thus, no evidence on causality, and therefore impacts, can be drawn from indicators, unless a counterfactual situation is reproduced.

Because quantitative techniques need robust, homogenous and reliable data in order to be applied without violating their core statistical assumptions, the choice of which method is a best fit for an evaluation question is affected to a large extent by data availability. This is confirmed by the BRG toolbox, which affirm that in the design of an evaluation, *“data availability and the quality of the data will play a key role in deciding which methods can reliably be applied; at the extreme, it may also influence the scope of an evaluation.”* (European Commission, 2015b:341).

The starting point is that in order to devise any sort of quantitative research method one ought to start from identifying independent and dependent variables, i.e. a variable of interest (independent variable) that is being investigated as a cause of an outcome variable (dependent variable). In the counterfactual design, the independent variable is the policy intervention under observation, that needs to be homogeneous in order to allow estimating its impacts. In order for impact estimation to be statistically feasible, however, one key issue should be considered that is: minimum sample sizes. Statistical inference, at the basis of all CIE quantitative methods, can be applied only if a reasonably large pool of units is observed. In any event, the smaller the sample size, the lower the probability of retrieving results that are statistically significant – and, as a result, of identifying and measuring impacts.

Next, a number of assumptions need to be respected to ensure that the estimates are unbiased.

- one central assumption refers to the **homogeneity of effects**, which describes the assumption that the treatment has the same effect on all individuals. It goes without saying that a pre-condition for this assumption to hold true is that the treatment should be unique or the very same for everyone. This is, however, not the case if a simple aggregation of similar measures is done at the EU level;
- **Common Independence Assumption (CIA)**: common variables that affect treatment assignment and treatment-specific outcomes must be observable (and observed). The dependence between treatment assignment and treatment-specific outcomes can be removed by conditioning on these observable variables. This assumption is also known as a selection-on-observables assumption because its central tenet is the observability of the common variables that generate the dependence;
- **Stable Unit Treatment Value Assumption (SUTVA)**, which means that "the [potential outcome] observation on one unit should be unaffected by the particular assignment of treatments to the other units" (Cox, 1958: 2.4); should this not be the case, then the comparison between a beneficiary and a comparable non-beneficiary would not alone provide a measure of the net effect of the intervention, as also the performance of the latter would be indirectly influenced by the intervention;
- the presence of **spontaneous dynamics**, due to the fact that target population differs from control population for the trend of the result variable. A solution is the introduction of an identification hypothesis to take in consideration the spontaneous dynamics of the result variable trend (this is the case, e.g., of Difference-in-difference methods which, on the other hand, help overcome the CIA assumption).

A range of quantitative techniques exist to support researchers in adapting their research frameworks to available data while satisfying to the extent possible the above assumptions. Below the main are listed together with a brief discussion on their limitations categorised through a hierarchy of evidence, i.e., grouped in categories sorted by the decreasing strength of the evidence they typically provide (see for example Roberts and Yeager, 2004), although it should be recalled that each individual case and technique should be considered in the research design to infer anything on the reliability of the actual findings thereof.

Level 1 – experiments:

- **Randomised Control Trials**
 - Key features: this is the gold standard in research, used mostly in medicine. By randomly selecting units of the control group, it is fair to assume that all the counterfactual assumptions are respected. This happens when, e.g. an incentive or funding is made available to firms/organizations on the basis of a random process, or, more traditionally, a drug to patients. In this situation, the randomisation allows to assume away all structural differences between those who are treated (and receive funds or support) and those who are not, so that one can use the non-supported organizations as a control group for comparison with the former group. A simple difference in mean in the outcome variable will give information on the extent to which a given policy has been successful, or not.
 - Key limitations: this method should be devised before the intervention takes place as individuals need to be randomly selected. Ethical concerns may arise and there is a strong internal validity which may be offset by low external validity (results may be very specific to the context in which they are studied or towards that specific population).

Level 2 - natural experiments, observational studies and quasi experiments:

- **Regression Discontinuity Design (RDD)**
 - Key features: this method, first introduced by Thistlethwaite and Campbell (1960), can be applied to situations in which it is possible to identify a clear cut-off level for treatment access and in which treatment status is based on observable characteristics. Typically, the cut-off is defined by the eligibility rules of the incentives (or participation to a course/training/learning experience) so that the treatment group is made up by organisations or individuals that just satisfy these criteria (and hence have access to the grant/support), whereas the control group is composed of organisations (individuals) that are just below the cut-off level and do not have access to the grant (support). In such a circumstance, it is reasonable to assume that the control group and the treated groups are very similar against most criteria correlated with the performance towards the objectives of the intervention, and that the small difference in the variables guaranteeing access to treatment are not sufficient to justify a different value of the outcome variable, so that a difference in the latter can be entirely attributed to treatment. RDD designs can be either sharp (full compliance, all those selected for support actually take it and all those not selected do not) or fuzzy (partial compliance). In the second case, two stages analysis such as those applied with an instrumental variable can be applied.

- Key limitations: Although RDD sits rather high in the hierarchy of evidence, it can be only be applied on homogeneous interventions and where a continuous measure or rating (running variable) which determines treatment – e.g. age – exists. This implies, for example, that if the cut-off is binary (e.g. being a migrant), it cannot be used. Other issues are linked to the fact that findings can hardly be generalised as they are drawn from a rather limited pool of entities.
- **Instrumental variables and natural experiments:**
 - Key features: this method is relevant when the exposure to the policy is to a certain degree determined by an external force which does not affect the outcome of the policy directly, but only indirectly, through its influence on the exposure. Angrist and Krueger (2001) define this situation as natural experiment, i.e. “*where the forces of nature or government policy have conspired to produce an environment somewhat akin to a randomized experiment.*”
 - Key limitations: unfortunately, this method is not often feasible in applied research as it does not work when treatment exposure (i.e. the application of the policy) is not mandatory and depends upon some selection process that needs to be controlled for. Another major weakness of the approach is that it can be difficult to find an instrument that is both relevant and exogenous.
- **Synthetic control:**
 - Key features: the creation of a synthetic control is another strategy that can help understand whether a given intervention has worked for a given entity, as explained for example in Abadie and Gardeazabal (2003). The main idea is to identify a group of entities which are to the extent possible similar to the one that has received the treatment but have not done so. In order to test similarity, the group of controls is weighted on a vector of predicting variables (i.e. socio-economic features that are considered as determinant of the outcome variable) so that it shows collectively a pre-treatment performance of a given outcome variable that is as similar as possible to that of the treated entity. In the case of a MS, it may be that the synthetic control is the weighted average of a group of other MS which had performed pre-treatment (when weighted together) just like the MS that received the treatment. If properly defined, the post-treatment performance of the synthetic control works as a counterfactual for the MS that received the treatment, i.e. it shows what sort of performance the MS would have had, in case it had not received the treatment.
 - Key limitations: this is a powerful and intuitively simple technique that however is data-hungry in the sense that it requires several observations across the years on observable variables which should help predict the trend in the outcome variable. One should also be sure that no other similar policies have been enacted in the synthetic control group.
- **Statistical Matching Methods:**
 - Key features: “matching” can be broadly defined as any method that aims at identifying a control group for treated units whose distribution of relevant observable characteristics is as close as possible to that of the those who benefitted from the intervention. Matching methods are particularly relevant in the field of non-experimental studies (see for example Rubin, 2007), i.e., where there is no ex-

ante experimental design (with the related data collection). Their use is, accordingly, rather widespread in social sciences. The main idea of matching is that it is possible to identify, ex-post, a control group showing characteristics affecting the probability to receive funding from the intervention which are similar to those of the group who actually received the funding. This would allow solving the issue of self-selection (i.e. only best performers have received funding, and therefore they would have overperformed even in absence of the intervention).

- Key limitations: the main issue of this method is that it requires comparatively large pools to obtain a proper balance between the treated and non-treated in the control group, meaning that it is difficult to find to groups of individuals that are similar enough under all the dimension that may affect either the chance of receiving the funding or the performance towards the objectives of the calls (common support). Thus, the CIA assumption may fail to hold true.

Level 3 – weak quasi experiments

- **Difference-in-differences:**

- Key features: this method provides for a way to overcome the limitations of the CIA through exploiting panel data. The main idea is that if values of the dependent variable are observed before and after the intervention across both the group of those who have received the intervention and those who have not, then it is fair to assume that by correcting the difference in the post intervention values for the differences in pre-intervention values what's left is the net effect of the intervention. In other words, the impact of a policy can be estimated by computing a double difference, one over time (before and after the treatment) and one across subjects (between treated and non-treated). For a survey on the literature on DiD method see for instance Lechner (2010).
- Key limitations: this, while overcoming the CIA assumption, implies another strong assumption: that both the factual and control group would have kept their trend over time had not the intervention taken place. Such assumption would hardly hold true in most environments, therefore this method, when not complemented with other analytical strategies (e.g. Difference-in-difference-in-difference or Conditional Difference-in-differences, or synthetic controls), which however require even more data, lies at a low level in the so called hierarchy of evidence.

As it can be seen, each counterfactual technique has its strengths as well as limitations.

The key cross-technique limitation when it comes to the study of the effects of a programme, as it will be done in chapter 3, is that these are inherently partial equilibrium techniques. In other words, these are techniques devised to understand which are the direct impacts on those receiving some form of support, but seldom inform about the indirect effects on those not receiving support, or what happens to the economy as a whole as a result of the intervention. Assuming absence of spill overs and interactions with the environment can be quite misleading, as in case of large programmes, especially with incentives attached, employers might be interested to hire, e.g., some of the individuals treated instead of the non-treated, so the increased employment chances for some, might mean reduced chances for others. In addition, support comes at a cost and need to be funded through taxation. This has negative effects on

the economy in the short-term, and one should always consider whether benefits outweigh costs, or not. These and other aspects will be discussed more in details when it comes to the discussion of assessing macro-economic impacts in chapter 3. Issues with the evaluation of active labour market policies in particular are discussed in section 1.3.2.

1.1.2.3 TBIE, CIE and the better regulation guidelines

Whilst the CIE is considered particularly reliable in answering to the attribution question (is the policy itself that produced the desired change on beneficiaries?), the need for using qualitative techniques in addition to or even instead of quantitative analysis is largely acknowledged in real world research and also in the Better Regulation Toolbox. This clarifies that “*when evaluating EU legislation, it is particularly difficult to identify a robust counter-factual situation (i.e. what the situation would be if EU laws had not been adopted), making absolute quantitative analysis problematic. Often, EU evaluations have to rely on qualitative, reasoned arguments (backed by the appropriate evidence) about the likely role/contribution of an EU intervention to the changes observed*” (European Commission, 2015b:344). Furthermore, typically CIE methods are accompanied by TBIE-type analysis, as a sound understanding of the logical chain of the intervention is key to gain intelligence over the issue at stake and can say a lot as to whether conclusions can be transferred, generalised etc.

As a result, studies usually employ a combination of techniques aimed at gathering and analysing objective and subjective data – statistical reports, monitoring or modelling data and opinions (interviews/survey) respectively – to “*establish a relationship between an intervention and the observed changes in the issues which the intervention addressed.*”²

However, as anticipated in section 1.1.2.2 an important distinction should be made between: a) micro level effects, i.e. the intended change that should be achieved by a specific project, assessing the achievement of operational objectives, and b) macro level effects, i.e. the intended change from a bundle of projects/programmes on a more aggregated level to assess contribution to the achievement of specific objectives in terms of results and, where relevant, impacts. These aspects, which are however not directly addressed in the Better Regulation Guidelines, will be discussed separately in section 1.3 below.

1.1.3 Results-based M&E: how monitoring systems differ from/interact with impact evaluations

First of all, it should be noted that there has been an evolution in the field of M&E when it comes to more result-based approaches. As clarified in the World Bank guidance document (Zall Kusek, and Rist, 2004:XI) “*governments and organizations may successfully implement programs or policies, but have they produced the actual, intended results. Have governments and organizations truly delivered on promises made to their stakeholders? For example, it is not enough to simply implement health programs and assume that successful implementation is equivalent to actual improvements in public health. One must also examine outcomes and impacts. [...] These systems help to answer the all-important “so what” question, and respond to stakeholders’ growing demands for results.*”

² Ibid

It was notably this shift in perspective which pushed experts and practitioners to further reflect also about the differences between monitoring and evaluation and on the limits of just looking at monitoring indicators to measure impacts and understand what works. This is well clarified in the guidance for M&E produced by the British Department for International Development (DFID, 2002:10). *“Monitoring and evaluation are often mentioned jointly under the acronym M&E. Although monitoring and evaluation are complementary analytical functions, they have clearly distinguishable characteristics and vary in depth and purpose. Monitoring is a continuous internal process, conducted by managers, to check on the progress of development interventions against pre-defined objectives and plans — ‘keeping the ship on course’. [...] The evaluation will answer specific questions related to the relevance, effectiveness, efficiency, impact and sustainability of the completed development activity. Impact is assessed in terms of intended and unintended, positive and negative effects, and whether these can be attributed to the project or other forces operating in the same context. Unintended negative effects would also be noted where these are attributable to the project or programme. Evaluations uncover wider lessons that can be applied elsewhere.”*

So the attribution question becomes centre stage in the discourse on M&E. Nevertheless, the situation in the EU seems slightly more nuanced, and “impact indicators” have long been used within the context of cohesion policy.

This ambiguity was already spelled out by Martini (2009) in his *“How counterfactuals got lost on the way to Brussels”*, who contended that the use of the term “impact indicator” was rather spurious, as there is no such a thing as an indicator that can, by itself, measure an impact. Martini ascribed the disproportionate emphasis placed at the EU level on indicators rather than counterfactual evaluation to the EU’s overriding concerns on accountability, which bordered onto hampering approaches aiming at understanding “what works and for whom”.

After, a.o., Martini’s critique, DG REGIO in their EVALSED sourcebook of September 2013 (DG REGIO, 2013) had in fact amended and updated significantly the EC’s position on the role of indicators.

When describing “result indicators” the EVALSED guidelines had clarified that *“Result indicators represent change sought by the programme or priority to which the programme or priority should contribute.”* (DG REGIO, 2013:86). So it is clear that they represent change that is sought by – not achieved because of – the programme. The guidelines went on to emphasise *“Readers’ attention is drawn to the changing concepts of “results” and “impacts” in the evaluation literature.”* (DG REGIO, 2013:86). The EVALSED ended up clarifying that *“Impact indicators have been removed from Structural Funds programmes, except those dealing with fisheries and rural development. For the DGs for Regional Policy and Employment and Social Affairs impact is the contribution of the intervention to change. It cannot be monitored; it requires evaluation.”* (DG REGIO, 2013:86). However, the wording used in the Better Regulation Guidelines seems to leave some larger room for interpretation about the role of monitoring.

On the one hand, the BRG make it clear that *“while monitoring looks at “what” changes have occurred since the entry into force of a policy intervention”*, i.e. changes during the intervention *“evaluation looks at “whether” the intervention has been effective in reaching its objectives,*

*and whether the objectives have been met efficiently (i.e. at least cost), as well as the reasons for the success or otherwise of an intervention”, i.e. **because of** the intervention. So it is clear that, at the outset, the BRG ascribe no causal link between what monitoring tracks and the intervention that may have caused the changed observed. They further specify “It is important to note, however, that the data collected will reflect changes due to the EU intervention and those caused by other factors.” (European Commission, 2015b:303)*

Nevertheless, when describing the monitoring indicators, the BRG first state that these should “[.] allow you to monitor the changes in terms of the main objectives” and, when providing examples of monitoring on results and impacts, clarify that “*Indicators aim at monitoring what concretely the policy intervention intends to achieve, i.e. raison d'être of your policy. **They represent changes over the short, medium and long term which can be directly linked to the application of the intervention. These indicators should include monitoring both the direct, as well as any significant indirect or unintended impacts of an intervention. They should be closely related to the identified problems and their drivers.***” (European Commission, 2015b:307)

Although the wording on the “direct link” of the BRG should probably be interpreted in the sense that indicators should measure all relevant dimension of a policy subject-matter, it seems to allude to the possibility of measuring change through indicators as they should allegedly “represent changes over the short, medium and long term”.

Quite significantly, in the ESF/ YEI monitoring guidance (further discussed in chapter 2), it is mentioned that “*Result indicators **capture the expected effects on participants or entities brought about by an operation.** Result indicators shall correspond to the specific objectives set out for each investment priority selected. Result indicators go beyond output indicators in so far as **they capture a change in the situation, in most cases related to supported entities or participants, e.g. in their employment situation.** In order to minimise external factors influencing the value reported under the result indicators, it is advisable to set indicators which are as close as possible to the activities conducted under the respective investment priority. This implies that the effects on participants or supported entities are captured and not the overall effects on a certain group of society or categories of entities which might include people or entities who did not benefit from ESF support. Results can be immediate or longer-term. Data for result indicators are collected through various means and in varying intervals.*” (DG EMPL, 2018:7)

Furthermore, data from output or results indicators are commonly used in reporting and visibility material. This question will be further discussed in chapter 2, together with a more in depth analysis of the monitoring system of selected programmes.

1.1.4 Summing up: key principles and features of monitoring and evaluation

All in all, it is clear that a monitoring and evaluation serve different purposes, but that are strictly interlinked one another and have a “complementary role”. In summary:

- Monitoring should:
 - **provide timely information to the those concerned with the implementation of a policy, programme or project whilst not by itself delve into the issue of causality**

to check whether any deviations from the original implementation plan are materialising. It should also be **relevant** in the sense that it should not overwhelm decision-makers with several non-decisive indicators which would steer attention away from the key ones

- **provide to the extent possible data which is then used in impact assessment** and therefore should:
 - be **comprehensive, grounded on an intervention logic**, thus measure all possible dimensions of a phenomenon that is addressed by a policy, including looking at factors that might have hampered implementation
 - be **cost-effective**, thus avoid duplication with other monitoring frameworks and measure key elements of an intervention
 - be **timely**, which does not mean that all data should be collected at all times, but that information is there when it is most relevant (e.g. to feed in evaluations, to monitor milestones etc.)
 - be **accessible**, so that the general public, including researchers, can make use of the data and produce new evidence
 - be **participated**, so that all relevant stakeholders can provide feedback
- Evaluation should:
 - Answer to at least 5 evaluation criteria (effectiveness, efficiency, coherence, EU added value, relevance)
 - Be grounded on an intervention logic and deploy appropriate methods to assess the evaluation criteria, including:
 - Theory based methods, which focus on “why” a given change has been produced (**attribution** question)
 - Counterfactual Impact Evaluation methods, which focus on “if and how much” change was produced by a given intervention (**contribution** question), including taking into account whether macro-economic dynamics (spillovers) might play a role
 - Be focused on causality

It is also important to notice that both monitoring and evaluation will necessarily take different forms depending on the level at which they operate. In general one should consider that:

- At the project/intervention level:
 - Monitoring: it can more easily be done on a case-by-case basis and monitoring data is typically collected ad-hoc and through the project’s monitoring system. It is easier to have indicators for the whole intervention logic, but proportionality should be taken into account. Furthermore, if the project is nested within a programme, then monitoring arrangements should be streamlined to the extent possible to allow aggregation and keep administrative costs and burdens proportionate.
 - Evaluation: the pool of the treated should be sufficiently large to allow for statistical assumption to hold when using counterfactual impact assessment, but it is much easier to make some key assumptions (e.g. no interactions between the treated and the general environment, homogeneity of effects on the treated, homogeneity of the treatment itself). It is also easier to control for self-selection in some cases, as more

detailed information on the features of the support and related beneficiaries can be retrieved.

- At the programme level:
 - Monitoring: less room for ad-hoc monitoring and greater need to harmonise data collection systems and standards.
 - Evaluation: intervention's heterogeneity and spillover /general equilibrium effects become increasingly problematic when looking at programmes which, by definition, aim to change and improve the interventions' environment.
- At the policy level:
 - Monitoring cannot be done on an individual basis. It will typically rely on sources which go beyond the implementation of the policy itself and are likely less timely and specific than those at the programme or project level
 - Evaluation: counterfactual impact evaluation risk becoming impossible as not only partial equilibrium, but also general equilibrium effects should be considered (i.e. some of the assumptions, including that of no spillover effects and of the SUTVA can hardly be fulfilled in a general equilibrium environment). For some policies meta-analysis of secondary CIE studies may be used, but generalising results remain hard.

Whilst more specific criteria and methodological aspects will be tackled in the next chapters, especially for sectoral (HC related) policy and evaluation arrangements/issues, these key elements will provide a backdrop against which to structure this research project.

1.2 Human Capital: what is it and how should we measure it

Having shed some light on what is it meant by monitoring and evaluation and what are the key issues to be taken into account, it remains to be discussed what is it meant by Human Capital (HC).

According to Kiker (1966) and many of the authors who investigated the topic in subsequent work, the concept of human capital can be traced back to the 17th century, when Sir William Petty placed a value on laborers, and to the *'[...]inquiry into the nature and causes of the wealth of the nations'*, the cornerstone work of Smith (1776) recognising that talents are part of one's fortune and likewise of the society's. Yet, according to Goldin (2014) the issue of skills development has gained momentum in the academic and policy debate especially with the studies of human capital as an investment decision, following from the work of Mincer (1958) and the first human capital model associated with Becker (1962 and 1964).

Macroeconomists such as Mankiw, Romer & Weil (1992), based on the observation that not all labour can be considered equal, employed human capital in order to explain what was contained in the 'residual' factor in Solow's (1956) model on economic growth, as well as productivity differentials among economies.

These studies are quantitative and have sought to gauge the role of capital accumulation in economic growth, by disaggregating the latter into its main driving factors.

Although, an accurate proxy for – and possibly definition of – human capital is yet to be found. The very reason for this is that there are several possible definitions and measurements of human capital, for it being a too broad concept to be summarised by just one (or a few) indicators.

Whilst being clear that human capital cannot be measured by way of years of formal education, even more recent and articulated attempts at encompassing additional elements potentially forming part of human capital (quality of education, health factors, family background and so forth) have fallen short of being conclusive.

In other terms, no single factor – nor composite vector – has yet been found that explains what Solow’s ‘residual’ is made of in a model of long term growth and, likewise, that explains differentials in labour productivity and remuneration, especially intra-group.

Amongst factors that should be assumedly accounted for, researchers (see for example Franzini and Raitano, 2011) have pointed to labour market segmentation, the role of social networks and family expectations, cultural and institutional factors, the level of union membership in a given sector and so forth. Research has also pointed to the relevance of factors such as openness, team working aptitude, cognitive and non-cognitive skills, adaptability.

Furthermore, while human capital has been hitherto discussed as essentially linked to productivity – and particularly as a means to increase it – many have rightly contended that its social relevance is far broader and associated with more complex and long-lasting benefits. Human capital raise is indeed usually coupled with better health conditions, social inclusion and opportunities.

Against this backdrop, the struggle continues to lie in comprehensively defining and observing/measuring human capital.

Yet human capital, however defined, seems inevitably a key element of long-term, sustainable and inclusive well-being (see for instance OECD (2001), Riley (2012), Lucas (1988), De La Fuente and Domenéch (2000 and 2006) and Stiglitz, Sen & Fitoussi (2009)). This is why, a.o., the EU invests substantially in human capital development, as a means to achieve its overarching goals of a smart, sustainable and inclusive growth. But without a clear idea of what human capital is, there’s a risk that the EU will fail to properly monitor and evaluate whether such investments are achieving their intended goals.

The starting point is to look at how researchers, experts and international institutions have been tackling the problem of defining and measuring over recent years. The goal is to understand if any of this can inform the design of future monitoring systems of initiatives aimed at promoting human capital in the EU.

1.2.1 What is human capital?

Over the years, many have been the attempts to organically describe the concept of human capital and categorise some of its distinguishing elements. But the problem is that these vary depending on the perspective from which human capital is studied. In its broadest and most commonly accepted definition, human capital is defined by the OECD (2001) as “*the knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic well-being.*”

In 2009, a three-day conference organised by the OECD³ was held in Busan (Korea) which brought together a range of academics and practitioners around the key question of what could human capital be and how to measure it. The conference discussed some of the key perspectives from which human capital was traditionally measured, thus presenting a taxonomy of human capital measures, together with suggesting where future work should focus on the theme.⁴

This taxonomy, further complemented with new research on the theme, identifies the following perspectives from which human capital has been observed:

- **Individual** point of view:
 - something akin to property, Schultz (1961);
 - human capacity as knowledge, skills (Beach, 2009) education and abilities (Garavan et. Al., 2001; Youndt et al., 2004);
- **Process of accumulation** point of view:
 - knowledge and skills (competencies, expertise and wisdom) obtained throughout education activities such as compulsory education, postsecondary education, and vocational education (Alan, Altman & Roussel, 2008);
- **Production oriented** point of view:
 - human capital as ‘*a fundamental source of economic productivity*’ (Romer, 1990);
 - an investment that people make in themselves to increase their productivity;
 - an amalgam of factors such as education, experience, training, intelligence, energy, work habits, trustworthiness, and initiative that affect the value of a worker's marginal product (Frank & Bernanke, 2007);
 - ‘the stock of skills and knowledge embodied in the ability to perform labour so as to produce economic value’ (Sheffrin, 2003); and
 - ‘the knowledge, skills, competencies and attributes in individuals that facilitate the creation of personal, social and economic well-being’ (Rodriguez & Loomis, 2007).

The choice of the perspective, is not irrelevant to the arrangements one might want to use to monitor human capital development. For instance, the individual perspective will favour approaches aimed at measuring the individual know-how and abilities in an objective perspective. The process of accumulation will look at some measure such as number of years of schooling, somewhat irrespective of the actual skills acquired. The last one (production oriented) will look at the increases of productivity coming as a result of increased human capital, both at the individual or societal level. Obviously, if one just measures productivity, it might be that individuals which become more self-aware than others, or just better citizens, will

³ 3rd OECD World Forum on “Statistics, Knowledge and Policy”

⁴ Much like many other subsequent papers on the theme, this section takes the preparatory work of Kwon, D. B. (2009), Human capital and its measurement. The 3rd OECD World Forum on “Statistics, Knowledge and Policy” Charting Progress, Building Visions, Improving Life Busan, Korea - 27-30 October 2009, from <http://www.oecdworldforum2009.org> as a reference. It goes beyond it, complementing it with some elements of the debate which followed that 2009 paper. In order to do so, the 238 papers which cite the 2009 paper have been taken into account. Amongst them, around 10 have been selected as they explicitly tackle the methodological points raised by the paper on the measures and determinants of human capital, together with some of the references through a snowball technique.

not display clear increases in their productivity, or will do so only indirectly. But maybe there will be an added value for the society as a whole. Hence, these different approaches might carry different forms of bias and should be considered also with a view to assessing whether monitoring is fit for purpose, or, in other words, adequate to measure progress towards the achievement of a programme's objective.

In addition to the different perspectives from which HC was studied there are some other key aspects that have been investigated, including:

- **Indigenous characteristics of HC**, which is expandable, self-generating, transportable, and shareable (Crawford, 1991)
- **The level at which human capital development impacts⁵**
 - Individual:
 - a) it means comparatively higher rewards in the internal/external labour market (Edwards, 1979)
 - b) it positively influences socio-psychological individual health (Blakey, et al., 2002), (Veenstra, 2001), (Veenstra, et al., 2005), (Wilson, et al., 2004)
 - firm/organisation:
 - a) it increases human resources' productivity in the workplace (Lucas, 1988), (Rosen, 1999)
 - b) organization/company/firms core competences as a determinant factor to improve the competitive advantage (Lepak & Snell, 1999)
 - c) it can be seen also as organisational culture and relational capital (Draghicia & Harpana, 2013)
 - societal:
 - a) it affects national economic growth (Romer, 1990)
 - b) it spurs civic participation, social cohesion (Stiglitz, Sen & Fitoussi, 2009)
- **Typologies of human capital**
 - **General** “*generic knowledge and skill, not specific to a task or a company, usually accumulated through working experiences and education*” (Alan et al., 2008)
 - **Specific** “*the specific human capital is rarely transferable to be applied to other jobs, firm, and industry, and thus it is impossible to transfer much income in the labor market*” (Becker, 1964 and 1976)

Other researchers have focused on the determinants of human capital creation, such as for instance:

- **Density, agglomeration, demographic trends**
 - Big cities attract human capital (Rauch, 1993)
 - The lower the population growth the lower the share of high-skilled, (Bjerke, 2012)
 - The higher the density, the stronger knowledge spillovers (Qian et al., 2012)

⁵ Based on Draghicia, A. & Harpana, I. (2013), *Debate on the multilevel model of the human capital measurement*, Procedia - Social and Behavioral Sciences 124 (2014) 170 – 177

- **Expenditure on education increases human capital** (Jung and Thorbecke, 2003; Idrees and Siddiqi, 2013)
- **Cultural diversity and tolerance**, which attract talents with very diverse background (Florida, 2002; Mellander and Florida, 2006; and Florida et al., 2008)
- **Urban amenities** attract talents (Glaeser et al., 2001).
- **Industry structure** (OECD, 1999)
- **Proximity to higher education institutions** (Moretti, 2004)

Another strand focused on the heterogeneity of the impact of education, depending both on the form of human capital development supported as well as the target group:

- **Vocational education vs. general education and vs. higher (academic) education:** different impacts on lifelong earnings, with VET being more effective in the short term and general education increasing prospects in the longer term (Golsteyn and Stenberg, 2017). Furthermore, according to Psacharopoulos and Patrinos (2018) the returns of academic schooling are higher than vocational education, also due to the latter's higher unit costs
- **Primary education vs secondary and tertiary education**, with returns of the former being higher (Psacharopoulos, 1985)
- **Women vs men**, with returns of the former being higher (Psacharopoulos, 1985)
- **Less developed vs more developed countries**, with returns of the former being higher (Psacharopoulos, 1985)
- **Employees in the private sector vs employees in the public sector**, with returns of the former being higher (Psacharopoulos, 1994)

As it can be seen, there are several elements of human capital which should be taken into account and many angles from which assessing it. For a monitoring and evaluation perspective this is also relevant as, for instance, understanding which are the determinants of human capital can help assess whether there are any relevant intermediate outcomes that could be monitored or taken into account in the intervention logic of the policy/programme. Similarly, it can help understand whether factors which affect the achievement of the objectives are measured. As abovementioned, the different perspectives from which human capital is measured have also a bearing on the capacity of given monitoring arrangements to actually track developments which are relevant to the aim of a given programme.

1.2.2 How can we measure human capital?

After having looked at some key features how the concept of human capital as well as its determinants, attention is turned to the issue of how can HC be measured. Traditionally, three main approaches have been consistently used to measure human capital, and notably:

- **output based:** typically based on indicators which are a proxy of the human capital in a given country. It can be ratio of high skill adults over the total (Romer, 1990), average

schooling years and educational attainment levels (Barro and Lee, 2013) but also quality adjusted measures such as PIAAC levels;

- **cost based:** this is based on the cost for education, Vocational and Educational Training (VET) and Lifelong Learning (LLL) activities; and
- **income based:** this is linked to increases in the individual productivity in a lifelong perspective.

A different and more articulated taxonomy is presented in the United National Economic Committee for Europe (UNECE, 2016) guide that is potentially more fitting with a consideration of human capital in its broader sense. This taxonomy distinguishes between monetary and non-monetary approaches. In particular:

- **Non-monetary approaches:**
 - **output/indicator based methods, including quality adjusted methods.**⁶ These benefit from a rather simple definition (e.g. just average years of schooling of the population) but have no common metrics (which is notably money for monetary methods), which hampers assessing their value over other forms of capital and thus over time. There is also no consideration of issues such as depreciation over time which, again, would be important if one wanted to consider this vis-à-vis other forms of investment. Interestingly, one could divide these into two additional typologies of indicators:
 - **of flow:**
 - enrolment rates in schools;
 - attendance (such as Early school leaving);
 - pupil skills or test scores (PISA);
 - **of stock:**
 - average years of education (one of the more common measures of human capital is indeed the average years of education of the labour force (Kyriacou, 1991, Barro and Lee, 1993, 2000)). This has however two key drawbacks:
 - skill may differ to a large extent even where educational attainment is the same (Moretti, 2004, e.g. workers from a city vs from the country side; students in different countries (Behrman and Birdsall, 1983);
 - not all years matter the same, decreasing returns to education (non-linear), (Judson, 2002) (Psacharopoulos, 1994);

⁶ From the UNECE guide on measuring human capital “The indicators based approach estimates human capital based on educational output indicators. It is necessary to highlight here that the indicator approach is understood as referring to relatively simple indicators, which may not attempt to measure human capital directly. They are generally proxies that are more-or-less directly observable rather than model-based, and/or do not apply economic accounting principles (discounting, depreciation, interest rates). This therefore excludes costs and income which could also be understood as indicators. [...] this approach relies on several indicators that, though rich in information, lack a common metric therefore making them less suitable for other uses such as the assessment of the ‘sustainability’ of a development path, which requires comparing changes in the aggregate stock of human capital with the stocks of other types of assets “However, they can be used in conjunction with other approaches and play a complementary role”

- scholastic attainments (Glaeser et al., 1995; Glaeser and Maré, 2001; Glaeser, 2005, Barro and Lee 2013); and
- adult skills indicators (PIAAC).

As there has been “a tendency to expand human capital measurement towards both financial potential and quality of life (one’s wellbeing, social networks, motivation) in order to capture all aspects.” (Kwon, 2009) some additional measures might be taken into account, which are still reliant on indicators:

- **typology of occupation and demand of workers in given sectors:** number of people working in specific sectors, e.g., creative sector. This is suggested in Florida (2002a, 2002b, 2002c and 2008). The ‘creative class’ measures individuals who “engage in complex problem solving that involves a great deal of independent judgment and requires high levels of education or human capital” (Florida, 2008). He argues that the economy is transforming and is in need of creative people, or people who work in creative occupations;
- **health based:** this is based on the observation that health as a positive impact on aggregate output (Bloom, Canning and Sevilla, 2001). As observed by Bleakley (2010) and especially Victora et al. (2008) there’s however the issue of endogeneity, in the sense that is higher HC that conduces to better health. However, this is key from a human development perspective, as confirmed if one looks at the HCI index from the World Bank (see further below);
- **monetary methods:** “A key advantage of both the lifetime income-based approach and the cost-based approach is that they combine different aspects that contribute to human capital in a single metric (money); this is not true of the indicators approach which results in a dashboard of indicators. A dashboard can however provide important information for policy and decision making, information which can be hidden by a single indicator of the stock.” (UNECE, 2016:40) In addition, monetary methods can be biased in that they rely on strong assumptions on: (i) interest/ discount rate (ii) market segmentation, power, trade unions which affect the wage setting (iii) spillovers which may benefit the individual productivity and increase the return to education due to general equilibrium dynamics. There are two main monetary methods known:
 - **cost based methods:**
 - costs for obtaining knowledge, this can be disaggregated by education level (Judson, 2002). According to the UNECE (2016:40) “*The cost-based approach is in line with the valuation of the majority of economic capital in the System of National Accounts which is estimated using a PIM method. It also provides an estimate of the resources invested in the education and other human capital related sectors, which can be useful for cost-benefit analyses.*”;
 - **lifelong income approach:**
 - increases in the individuals’ remuneration over their lifespans thanks to their human capital, as in Mulligan and Sala-i-Martin (1995). According

to the UNECE (2016:41) *“The use of the lifetime income-based approach also allows investment in human capital to be broken down into several components: investment from births, depreciation from deaths, investment from education, depreciation from ageing, and with investment from immigration and depreciation from emigration [see Gu and Wong, 2010]. The lifetime income-based approach also has the advantage that it allows outputs to be measured independently of inputs. In contrast the cost-based approach relies on the assumption that the value of inputs is equal to the value of outputs. Accordingly, using the lifetime income-based approach an estimate of productivity in the education sector can be made.”*.

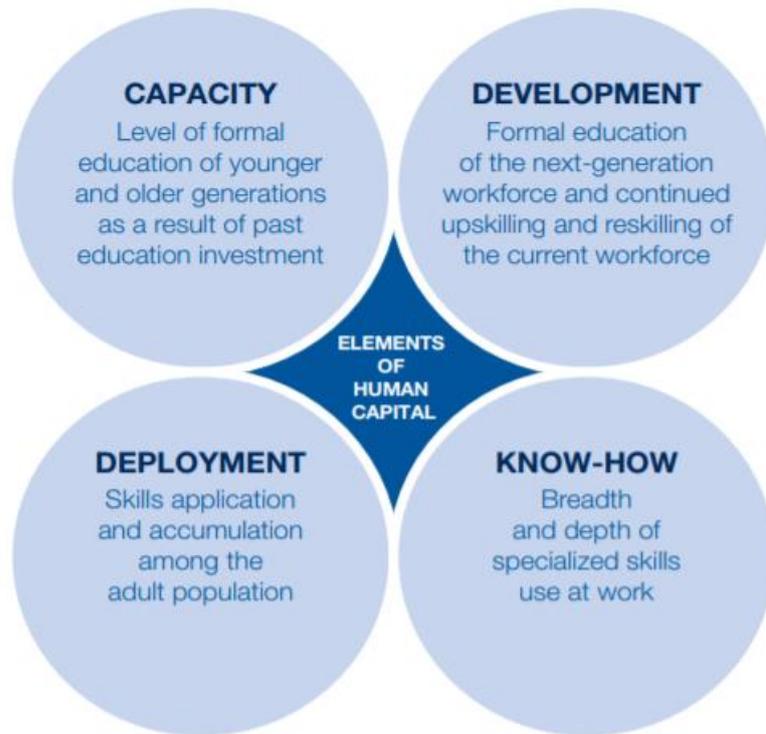
1.2.3 Human capital Indexes

Most clearly, human capital is a multi-dimensional concept. On the one hand, monetary approaches which merge several aspects of human capital into one common metric can be very useful for evaluations that compare countries, policy mixes, changes over time etc. On the other hand, non-monetary approaches can create dashboards which are useful for policy makers and help focus on the different components of human capital, including some which cannot be monetize but are still relevant for the quality of life (social cohesion, health, self-confidence, etc.).

This is why major international institutions such as the World Economic Forum (WEF) and the World Bank (WB) have developed their multi-dimensional indexes to track the development of Human capital over time and across countries. These however rely on non-monetary, indicator-based approaches. In particular:

- **World Bank’s HCI** is made up of five indicators, namely
 - the probability of survival to age five;
 - a child’s expected years of schooling;
 - quality of learning;
 - adult survival rate; and
 - the proportion of children who are not stunted.
- **WEF’s HCI** is an attempt to *“measure HC, holistically according to individuals’ ability to acquire, develop and deploy skills throughout their working life rather than simply during the formative years”* according to WEF (2017), and around 4 key dimensions (capacity, deployment, development and know-how)

Figure 2 - Structure of the World Economic Forum's human capital index



Source: WEF (2017) - *The Global Human Capital Report 2017*, p. 3

WEF's HCI is particularly interesting when it comes to its "deployment" sub index, which acknowledges that "*Beyond formal learning, human capital is enhanced in the workplace through learning-by-doing, tacit knowledge, exchange with colleagues and formal on-the-job learning.*" (WEF, 2017:5)

Overall, the following indicators are used (which are divided by age class, see table below)

Figure 3 – indicators of the World Economic Forum’s Human capital index

COMPONENT (Subindex weighting)	INDICATOR	AGE GROUP Share of total population (range of country values) ¹				
		0–14 (13%–48%)	15–24 (9%–22%)	25–54 (26%–70%)	55–64 (3%–14%)	65+ (1%–27%)
Capacity (25% of total Index score)	Literacy and numeracy					
	Primary education attainment rate					
	Secondary education attainment rate					
	Tertiary education attainment rate					
Deployment (25% of total Index score)	Labour force participation rate					
	Employment gender gap					
	Unemployment rate					
	Underemployment rate					
Development (25% of total Index score)	Primary education enrolment rate					
	Quality of primary schools					
	Secondary education enrolment rate					
	Secondary enrolment gender gap					
	Vocational education enrolment rate					
	Tertiary education enrolment rate					
	Skill diversity of graduates					
	Quality of education system					
	Extent of staff training					
Know-how (25% of total Index score)	High-skilled employment share					
	Medium-skilled employment share					
	Economic complexity					
	Availability of skilled employees					

¹ Individual countries' age group distributions are used for weighting indicators in the Capacity and Deployment subindexes to arrive at an overall subindex score.

Source: WEF (2017) - *The Global Human Capital Report 2017*, p. 4

It is worth noting that despite including relevant indicators of participation in the labour market, no indication of some of determinants of human capital are included (such as, for instance, health, number of people employed in the creative sector etc.)

- **Institute for Health Metrics and Evaluation** (Lim et al. 2018): here there is an approach which better comprises some elements linked to the functional health of individuals, including:
 - educational attainment;⁷
 - learning;⁸
 - functional health status relevant to economic productivity;⁹

⁷ Ibid. p.1219 “Estimates of average years of education were based on a compilation of 2522 censuses and household surveys.”

⁸ Ibid. p 1220 “student testing data from major international assessments and national continuing assessments of education progress”

⁹ Ibid. pp. 1220-21 “the prevalence of seven diseases and impairments identified in policy trials or observational studies to be related to learning or productivity (appendix). These include wasting, measured as the proportion of the population younger than 5 years below two SDs of the reference mean weight for height; stunting, measured as the proportion of the population younger than 5 years below two SDs of the reference height for age; anaemia, measured as the proportion of each age–sex group with a haemoglobin concentration defined by WHO as mild, moderate, or severe anaemia; cognitive impairment, measured as the proportion of the population with moderate, severe, or profound developmental delay; vision

○ Survival

As opposed to synthetic indexes, the OECD currently measures Human Capital mostly in terms of educational investments, participation and outcomes, as per the list of indicators shown in the Figure below.

Figure 4 – human capital indicators in the OECD framework

Factors	Details (range 1)	Details (range 2)
1. Investment in HC	High-level qualification	Growth in university-level qualifications (Growth in attainment levels in different fields)
	Graduation and enrolment rates	Trend in university-level graduation output Contribution of international students to university graduate output Entry rates into tertiary-type A education Entry rates at tertiary education compared to population leaving without completing tertiary education
	Time invested in education	Instruction time per year Number of hours per week spent on self-study or homework
	Investment in education	Expenditure per student at different level of education Percentage of GDP spent on educational institutions Private and public expenditure Public subsidies for education to households Expenditure on core service, ancillary services, and research and development Change in student numbers, expenditure, demographic forecasts etc.
2. Quality adjustment in HC investments	PISA assessments	
	PUIAAC (Program for the international assessment of adult competencies)	
3. Results of education	Matching of education to occupation	
	Labor market outcomes by age, gender, and educational attainment	
	Rates of return to education	

Source: Draghicia, A & Harpana, I (2013:175)

This approach spans across the cost and output based perspectives, including quality measures, whilst remaining out of the individual lifespan income perspective and from some interesting non-monetary measures linked to health, social inclusion, etc.

Ultimately, one can see that there is no one-size-fits-all approach to the measurement of human capital and that the different perspectives, i.e. the monetary and non-monetary one, tend to serve different analytic purposes, given their respective strengths and weaknesses.

If the aim is to assess the key components of long-term growth in a comparative fashion then probably monetary measures are the only way to go, and especially the lifetime income perspective. This is particularly relevant with a view to including measures of human capital in

loss, defined as the proportion of the population with moderate or severe vision impairment or blindness; hearing loss, defined by WHO as the proportion of the population with hearing loss greater than 40 dB in the better-hearing ear (30 dB in children); and infectious disease prevalence, with the use of three infectious disease aggregations from GBD 2016 classification, which includes HIV/AIDS, tuberculosis, malaria, neglected tropical diseases, diarrhoea, and several other common infectious diseases.”

the System of National Accounts (SNA, 2008) in the future, although there are reasons to say that this is still an open area of research.

On the other hand, if the aim is to have the most encompassing measurement of human capital, which goes beyond its monetary dimension, then indicators can better serve that purpose. According to the World Economic Forum (WEF, 2017), “*Using human capital index is a relatively new measure for capturing and tracking the state of human capital development*”. The possibility of using a vaster array of measures than a single monetary measure might allow for better representing the multi-dimensional nature of human capital. Nevertheless, there are limitations as to how well these can measure human capital across countries and, especially, over time.

In any event, aggregating human capital measures and taking into account spillover effects remain particularly defying challenges.

More generally, even most recent and consolidated approaches have decided to take it “step-by-step”, that is, starting from a narrow definition of human capital to then try to see how much can be added. The UNECE guide (2016) clearly stresses that including non-monetary elements in the estimation of human capital can be a particularly “daunting task”. In their work, some of the spillovers or feedback loops are included in the economic returns from human capital. The rest is just assumed away. These elements will be taken into account when considering how the EU is dealing with measures of its human capital investments.

One key message however is that over the years the concept of human capital has become a very all-encompassing one, and one should arguably consider whether there is room to monitor policies which aim at its development following this comprehensive “human development” approach. Yet, constraints that are linked to the strong administrative burden for data collection and storage are likely to hamper the process.

1.3 Monitoring and evaluation of HC policies

In addition to general indications on human capital and its measurement, an additional step of this literature review aimed at assessing whether in the existing guidelines or handbooks for M&E information could be garnered as to the principles for good M&E specifically in the field of human capital development policies. This strand of the literature review has included:

- the DEREc (DAC Evaluation Resource Centre) which lists 73 guidance documents from 20 Agencies;¹⁰
- the ILO guidelines;
- in the EU scenario, the better regulation guidelines and the EVALSED material;
- the IZA database;
- the OECD iLibrary; and
- Google scholar.

¹⁰ <https://www.oecd.org/derec/guidelines.htm>

The IZA and OECD libraries, and google scholar alike, saw a two-step strategy, with the first step being filtering documents for the word “monitoring”. Then, each document was screened checking for content related to a range of keywords, that included:

- education;
- human capital;
- human resources (development);
- training;
- skills/ upskilling/ reskilling;
- Vocational and Educational Training (VET);
- competences; and
- knowledge.

The first notable finding is that advise on sector-specific M&E is not very common, although some examples can be found which may include the educational or health sector. Especially in the IZA database very little could be found that was addressed to practitioners or researchers in the form of guidelines/ recommendations. The most useful source in this sense was the OECD iLibrary, somewhat unsurprisingly given the aims of the OECD.

The second worth noting element is that information on how to evaluate the effects of human capital policies can be found more easily, as understanding what works is the ultimate aim of the M&E systems taken as a whole. But proper evaluation cannot happen if the right data is not collected in advance.

1.3.1 Monitoring

Amongst the few documents which tackled directly the issue of principles of good monitoring in the field of human capital policies, there’s “*Monitoring Adult Learning Policies: A theoretical Framework and Indicators*”, Borkowsky (2013), written for the OECD. The paper acknowledges the importance of monitoring lifelong learning policies especially given trends such as the skills biased technical change, aging societies and artificial intelligence re-shaping the skills that are needed in the labour market. Based on a theoretical approach, Borkowsky aims to identify indicators which should be reported upon in the Education at a Glance publication of the OECD with reference to the Adult learning field. It reviews a number of policy objectives and themes which are linked to the adult education domain and seeks to identify good measures which can be suitably monitored on a regular basis so as to check nations’ developments in the field.

Even though it is intended to provide inspiration for the Education at a Glance publication, it ultimately serves the broader purpose of identifying dimensions which might be worth monitoring for Adult learning. The paper identifies 6 broad policy areas broken down into 18 policy objectives, which are considered to remain stable over time in OECD countries.

The six broad policy areas are the following:

- updating knowledge and raising skills / facilitating social and economic participation;
- reducing inequalities;
- ensuring adequate investment in adult learning;

- providing information and guidance/ facilitating Access;
- usability/ certification of skills; and
- quality of learning/ organisation of education delivery.

It is interesting to notice that these policy areas fall perfectly in line with the broad objectives of one of the key policy initiative of the EU in the field of Human capital development, notably the New Skills Agenda for Europe, which are, in particular:

- Improving quality and relevance of skills formation, which can be linked to:
 - Updating knowledge and raising skills / Facilitate social and economic participation
 - Reducing inequalities
 - Ensure adequate investment in adult learning
 - Quality / Organisation of education delivery
- Making skills more visible and comparable, which can be linked to:
 - Usability/ Certification
- Better information for better career choices, which can be linked to:
 - Information and guidance/ Facilitating Access

The approach is mostly that of non-monetary, indicator-based measures of human capital, though some specific elements on usability of skills, information and guidance are included. It also includes a measure of cost of education.

The paper also defines a theoretical framework, which distinguishes between policies (and related indicators) which intervene at different stages of the human capital development intervention logic, as per the below:

- Inputs: resources for the adult learning system, be them in the form of number of teachers/trainers or more generally investment in education/training
- Process: factors which allow the transformation of input into outputs, thus VET and LLL systems, Active Labour Market Policies (ALMP) providers and procedures etc
- Output: improvement in human capital of participants, e.g., new skills acquired, typically through proxies like qualifications, completion rates etc.
- Outcomes: higher earnings or higher employability or better social inclusion, but there should be sound analytical evidence on what contributed to such outcomes (in the papers' words, "*the existence of a causal link between the system and the phenomenon measured outside of the system*" (Borowski, 2013: 12)

The paper then details a list of approx. 30 indicators which cover the whole spectrum of 6 policy areas and 18 policy objectives, especially from the EU adult survey, PIAAC (also on civic participation) but also the EU Continuing Vocational Training Survey. When assessing these sources however, one should also consider their timeliness (i.e. whether the time at which they are done allows for measuring change that occurs while the intervention observed is ongoing).

The paper concludes that there are areas (especially that of usability /certification) where indicators are scarce and further work would be necessary. Given the alignment with the objective of the human capital policy of the EU, this framework should be considered when designing its monitoring and evaluation systems.

1.3.2 Evaluation

According to recent literature on the evaluation of human capital (see for instance Escudero, 2014) one should consider the following factors when evaluating human capital, skills related and Active Labour Market Policies:

- dynamics typically appearing at the **micro-level** of the analysis, or in the **short term**:
 - **self-selection**: those participating to support might do so because they are inherently different, for instance in terms of motivation and willingness to improve their position, from any group used as comparator;
 - **deadweight effect**: hiring of participants that would have occurred also in absence of the programme;
 - **signalling effect**: higher certainty for employers about the skills of the potential employees due to the fact that their abilities have been screened during the ALMP policy which makes employing them easier;
 - **lock-in effects**: lower employment levels for beneficiaries of training actions, as, once activated, they might feel the need to continue in education before becoming available in the labour market;
- **general equilibrium dynamics** (changes to both the wage-setting structure of labour and labour demand, typically appearing in the **medium to long term**):
 - **displacement effects**: those treated by the intervention becoming relatively better employable than the non-treated;
 - **substitution effects**: in the short to medium term (when physical capital is considered rather fixed), comparatively more productive human capital may lead to lower employment levels, as less workers are needed for the same output. The ensuing increase in output then creates an expansion of labour demand too, so employment levels increase in the medium to long term;
 - **scale effect**: the labour demand might be increased by that same increase in the marginal productivity of labour. The net effect on the labour market actually depends on whether the scale effects dominates the substitution effect or the other way around;
 - **labour market segmentation**: which may create uneven levels of employment, wages, elasticities, lead to concentration etc.;
 - **positive externalities**: know-how and competences may spread across industries, economic districts etc. Proximity matters;
 - **lower cost of vacancies due to improved matching**: this, in turns, means higher labour demand as the cost of “demanding” for labour diminishes;

- **lower wages**, given the improved matching lower need for employers to raise salaries to attract workers;
- **other benefits at the individual and societal level** from human capital, skills related and ALMP policies:
 - **on skills:** Lock in effects: training may generate positive effects even in absence of immediate employment as people may be prone to continue learning (e.g. towards a given qualification level now that they are back in educational activities);
 - **on health:** Improved health or hygiene conditions;
 - **on citizenry:** better participation to social life, civic education, social cohesion etc.;
 - **on crime prevention:** lower risk of people committing crimes to “*meet end needs*”;
 - **on work quality:** it contributes to decent work; and
 - **other soft outcomes:** Better confidence and self-esteem, contributing to engagement in family life and society (CEDEFOP, 2011).

In addition to the literature on micro and macro level dynamics on human capital development, a range of lessons learned reported in guidance documents on M&E of several international institutions should be also taken into account when devising an M&E monitoring system for human capital:

- to take into account **longer gestation periods for the appearance of impacts in the case of education policies** (African Development Fund, 1996);
- to **estimate qualitatively the rate of return in some projects** and only look at cost to identify cost effective ways (ILO, 2006);
- **institutional outcomes** that improve education provision should not be considered per se but as part of a given educational outcome (process) (ABD, 2015);
- to always look at impacts on **equitable access** to education and improves in **gender equality** which are linked to education;
- to always consider in the efficiency evaluation whether there has been **cream-skimming** (only best pupils treated);
- the inclusion of **disability dimension** is most urgent in the social, health, education and employment sectors;
- to consider **cluster evaluation as a possible method to group different HC development projects**, especially for complex programmes, as interventions may vary to a large extent, and so their intervention logics, input output and outcomes etc. (Leeuw & Vaessen, 2009);
- to recall **self-selection** issues in participating to programmes;
- to consider the **different levels at which impact can occur**, and for whom (direct, indirect, etc.); and
- to consider the **adverse impact** of a training programme which does not fully reach to the most disadvantaged and favours those relatively better placed.

1.4 Concluding remarks and monitoring of HC checklist

This chapter has shed light on a few key concepts and standards linked to the monitoring and evaluation of HC policies, including:

- the purpose and evolution of M&E in section 1.1, stressing the increasing emphasis on the concepts of results and effects, the distinct but complementary roles of monitoring and evaluation and the key principles thereof. It also highlighted some ambiguities in the way in which some guidance documents at the EU level deal with the difference between *results measured after* the policy and *effects generated by* the policies;
- the multifaceted concept of human capital in section 1.2, with its many angles and approaches to measure it, stressing the importance of more comprehensive indexes and proxies which can capture human capital in a holistic manner. Even more importantly, the section made it clear how different perspectives on the evaluation of human capital can shift the focus on different aspects, measures and determinants of human capital itself ultimately affecting the way in which effects of the policies are discussed. This means that one should be aware and explicitly discuss the framework that is being used, together with any implications; and
- a summary of guiding principles and key elements to be taken into account when defining a monitoring and evaluation system on human capital policies, including the importance of adapting the monitoring and evaluation strategy on the level at which a policy is implemented (micro, meso, macro) with related implications in terms of data to be measured and evaluation techniques which can(not) be applied.

In this last section, first, these elements are translated into an assessment grid for the analysis of monitoring systems of HC policies and, second, some conclusions are drawn as to how fit for purpose are the EU standards in this domain and what should be done to improve upon them.

1.4.1 Assessment grid for the monitoring systems of Human Capital Policies

Based on the overview of the literature, of the concept of human capital and on the principles on which a monitoring system of an initiative that aims to promote human capital should be based, a checklist is developed and presented below that will be used as analytical tool for the assessment of the monitoring system of some notable programmes (i.e., the ESF/ YEI and the FEAD) in chapter 2.

The checklist is the result of a combined reading of the better regulation guidelines, of the guidance documents of several other international institutions as well as of the debate on the measurement of human capital described in section 1.2.

It adds on the existing literature in that it aims at providing comprehensive lenses through which reflecting on whether the monitoring system of a given policy supporting human capital is fit for purpose, and notably:

- it allows managing agents to monitor progress towards the goals originally foreseen; and
- it collects data on the ground which is then suitable for in-depth evaluation with the appropriate tools.

This is somewhat a novelty in the EU scenario as no organic assessment on the issue of human capital development as a whole and its M&E arrangements in the EU exists, to the best of the author’s knowledge.

Although the checklist looks essentially at the more “monitoring-related” aspects of M&E, it is assessed whether the monitoring system is not only good in itself but also complementary with evaluation methods which will need to rely on monitoring data. Much emphasis is being placed upon a more encompassing definition of human capital, as per the most recent orientations described in section 1.2.

It should be recalled here that, as mentioned in section 1.1, the starting point is that of the definition of the intervention logic of the policy. Only there can one assess the monitoring arrangements through answering the questions defined in the checklist below.

Table 1 - Checklist for the monitoring of Human Capital Policies

Criteria	Specific question
Comprehensiveness	Does the monitoring system cover all objectives of the programme – as defined in its intervention logic – and especially the operational and specific ones with new arrangements? (European Commission, 2015b)
	Does it do this through surveys as well as continuous monitoring? (European Commission, 2015b)
	Do the indicators measure progress that can be linked to the different impacts on the different target groups, both direct and indirect? (Zall Kusek and Rist, 2004, adapted to take into account Martini’s (2009) critique on “impact indicators”)
	Are baseline and target values identified to track progress? (European Commission, 2015b)
	Are transversal objectives such as gender equality, non-discrimination, inclusion of people with disabilities included in the monitoring system? (European Commission, 2015b)
	Does it look at factors contributing to or impeding achievement of the outcomes? (UNDP, 2002)
	Is monitoring data suitable to be used for assessing the impact of the funds through counterfactual or other econometric approaches? (author’s own elaboration based on chapter 2 and 3 and the “early definition of data needs” principle in the BRG, European Commission, 2015b)
Participation	Is feedback from partner and beneficiaries foreseen in the monitoring? If so, how? (e.g. steering committees etc) (UNDP, 2002)
	Are experts and other stakeholders involved in the definition of monitoring arrangements? (Zall Kusek and Rist, 2004)
Proportionate and no overlaps	Does it treat the monitoring in a cost-effective manner (including electronic systems for data collection and exchange)? (European Commission, 2015b)

Criteria	Specific question
	Does it only prescribe new monitoring arrangements where needed and at a reasonable cost? (European Commission, 2015b)
	Does it maintain indicators constant over time unless new major needs emerge? (Zall Kusek and Rist, 2004)
Timeliness	Does it collect evidence in a timely fashion, i.e., not necessarily from the outset but only and every time this is needed? (European Commission, 2015b)
	Are milestones foreseen? (European Commission, 2015b)
Accessibility	Is data collected accessible for the general public? (European Commission, 2015b)
Relevance	Are only the relevant indicators included in the monitoring list or there's just too many of them? (DG Regio, 2013)
	Are indicators updated and changed where necessary (and only then)? (author's elaboration based on Zall Kusek and Rist, 2004)
	Are indicators set at a level where decisions are taken either on the policy or on the programmes? (Zall Kusek and Rist, 2004)
Suitability to measure human capital development	Does it take into account the main dimensions of human capital (capacity, deployment, development, know-how and creativity, soft or non-cognitive skills, health, culture and tolerance) as relevant to the objectives of the initiative? (WEF, WB, Lim et al. 2018)
	What is the underlying approach to measuring human capital, i.e., monetary or non-monetary? (Zall Kusek and Rist, 2004) Is it general human capital or firm specific human capital? (Becker, 1964 and 1976)
	Does it take into account relevant direct (first order) impacts as well as indirect (second order) effects? (Zall Kusek and Rist, 2004, but also Leeuw & Vaessen, 2009)
	Do indicators measure input, process, output or outcome-related factors of human capital development? (Borkowsky, 2013)
	Does it measure intermediate output/outcomes which can be linked to the intervention (enablers, soft outcomes) (European Commission, 2015b)

Source: Author's elaboration initially based on the better regulation guidance but then vastly expanded based on the literature reviewed in the chapter.

1.4.2 Concluding remarks

As it can be seen from Table 1 above, the EU better regulation guidelines, which represent the cornerstone of the EU practice in the field of monitoring and evaluation, are not fully suitable to instruct on a monitoring system which is fully fit for purpose, especially in the field of HC development. They effectively had to be complemented with other sources to develop a comprehensive assessment list for the monitoring of human capital. This has repercussions on the whole monitoring and evaluation practice of the EU.

The analysis carried out shows that this issue is also very much overlooked in the literature, as advice on sector-specific M&E requirements is scarce across the board, except maybe in the case of the OECD. But the arrangements for monitoring can differ radically depending on the sector to which they are applied, especially so in the case of such a complex concept as human capital development. Although information on how to evaluate the effects of human capital policies can be found more easily – as understanding what works is the ultimate aim of the M&E systems taken as a whole – proper evaluation cannot happen if the right data is not collected in advance, so this remains an area in which more attention is needed.

Amongst the main issues identified in the EU standards for monitoring and evaluation of human capital policies are:

- in terms of comprehensiveness of the monitoring system, they do not emphasise the importance of taking into account impacts on the different target groups of the policy, which is very important as it will be highlighted in chapter 2 and 3. More importantly, they fail to stress the importance of a collection of data in a form (and level of disaggregation) which is then suitable to carry out impact evaluation, both at the micro as well as macro level. Stronger links between the monitoring and evaluation practice are needed here, as it will be further stressed in chapter 2 and especially chapter 3 of this research;
- among the key governing principles for the monitoring system of the better regulation guidance, there is insufficient emphasis on the need to have monitoring systems which are participated, both in the design phase (by experts in the field) as well as in the continuous monitoring process;
- crucially, there needs to be some indication on sector-specific monitoring requirements. This is particularly important in the case of human capital development policies. Issues such as what dimension of human capital development is being monitored, from which perspective (individual, process of accumulation, societal) through which indexes and under what framework are paramount to make sure that the monitoring arrangements are suitable to inform managing actors on the progress of the implementation as well as provide information to those in charge of measuring the effects of support. Experience shows that monitoring certain aspects of the policy implementation affects the extent to which implementation strives to contribute to those aspects, especially from an accountability perspective. So if the monitoring arrangements do not focus on, say, both intermediate and final outcomes of human capital formation, pay heed to its determinants and embrace a more holistic approach to its measurement it is likely that the support itself which is offered by the policy will fail to touch upon key aspects of HC development. Similarly, if the focus is too much on the individual perspective (know-how) but not on the productivity one (or just the contrary), this has implications for the observable and unobservable effects of a

given policy. This is well showcased in chapter two, where an the concrete analysis of two monitoring systems of notable EU initiatives to boost Human Capital is provided.

- about the relationship between monitoring and evaluation, in line with Martini (2009), it is paramount that the difference between results and impacts be furthered in all guidance documents. For instance, as clarified in section 1.1.3, the guidance documents for monitoring and evaluation of the European Social Fund keep mentioning that “*Result indicators capture the expected effects on participants or entities brought about by an operation*” (DG EMPL, 2018: 7). This is remarkable also in the light of the updated principles included in the EVALSED as well as better regulation guidelines documentation. But in absence of a sharp distinction between a result measured after a programme and an effect generated by the programme there might continue to be confusion among both stakeholders and those dealing with the management of the programmes.
- Last but not least, in designing M&E arrangements more heed should be paid to the different forms of bias which might occur at the different levels of policy implementation, especially in the field of human capital development and active labour market policies. There are strong differences between M&E of a project, programme or the entire policy (thus typically micro, meso and macro level) and appropriate guidance should be produced as to how to treat such studies adequately across the board. Although the EVALSED guidelines give some brief indication to concepts such as deadweight, displacement and substitution effects, this is not tackled consistently in the EU M&E standards.

1.5 Main Bibliography

Abadie, A., and Gardeazabal, J. (2003), "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review*, 93 (1), 112–132. [493,494,496,497,501]

African Development Fund, 1996, Guidelines and Methodologies of Evaluation, Board of Directors, OPEV, ADF/BD/IF/96/214, 21 November. Available at: <https://www.oecd.org/derec/afdb/35142362.pdf>.

Alan, K. M. A., Altman, Y., & Roussel, J. (2008). Employee Training Needs and Perceived Value of Training in the Pearl River Delta of China: A Human Capital Development Approach. *Journal of European Industrial Training*, 32(1), 19-31.

Angrist, J. D. and Krueger., A. B. (2001) "Instrumental Variables And The Search For Identification: From Supply And Demand To Natural Experiments," *Journal of Economic Perspectives*, v15(4,Fall), 69-85. Available at: <https://www.nber.org/papers/w8456.pdf>.

Asian Development Bank (ADB) (2015), 2015 Guidelines for the Preparation of Country Assistance Program Evaluations and Country Partnership Strategy Final Review Validations, Available at: <https://www.adb.org/sites/default/files/institutional-document/32518/cape-guidelines-2015.pdf>.

Barro, R. and Lee, J-W. (1993). International comparisons of educational attainment. *Journal of Monetary Economics*, *Elvesier*, vol 32(3), p. 363–94, Available at: <https://ideas.repec.org/a/eee/moneco/v32y1993i3p363-394.html>.

Barro, R. and Lee, J-W. (2000). International data on educational attainment: updates and implications. NBER Working Paper 7911, Cambridge, MA. Available at: <https://www.nber.org/papers/w7911.pdf><https://ideas.repec.org/a/eee/moneco/v32y1993i3p363-394.html>.

Barro, R. and Lee, J-W. (2013), A New Data Set of Educational Attainment in the World, 1950-2010. *Journal of Development Economics*, vol 104, pp.184-198. Available at: <https://www.nber.org/papers/w15902.pdf>.

Beach, M. J. (2009). A Critique of Human Capital Formation in the U.S. and the Economic Returns to Sub-Baccalaureate Credentials. *Educational Studies: A Journal of the American Educational Studies*, 45(1), 24–38. Available at: <https://www.tandfonline.com/doi/full/10.1080/00131940802562313?scroll=top&needAccess=true>.

Becker, G. (1962) 'Investment in Human Capital: A Theoretical Analysis.' *Journal of Political Economy*, Vol. 70, No. 5, Part 2: Investment in Human Beings, pp. 9-49; Becker, G. (1964) *Human Capital*. New York: Columbia University Press. Available at: <https://www.journals.uchicago.edu/doi/pdfplus/10.1086/258724>

Becker, G. (1964) *Human Capital*. New York: Columbia University Press.

Becker, G. S. (1976). *The Economic Approach to Human Behavior*. Chicago: University of Chicago Press.

Behrman, Jere and Birdsall, Nancy, (1983), The Quality of Schooling: Quantity Alone is Misleading, *American Economic Review*, 73, issue 5, p. 928-46. Available at: <https://www.jstor.org/stable/1814663?seq=1>

Berry, C. R. & Glaeser, E. L. (2005). "The divergence of human capital levels across cities*," *Papers in Regional Science*, vol 84(3), pages 407-444. Available at: <https://rsaiconnect.onlinelibrary.wiley.com/doi/abs/10.1111/j.1435-5957.2005.00047.x>

Bjerke, L. (2012). Knowledge flows across space and firms. Ph.D. Dissertation. JIBS Dissertation Series No. 078. Jönköping: Jönköping International Business School, Jönköping University. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.453.9462&rep=rep1&type=pdf>

Blakey, T., Lochner, K., & Kawachi, I. (2002). Metropolitan Area Income Inequality and Self-Rated Health: A Multi-Level Study. *Social Science and Medicine*, 54, 65–77. Available at: <https://www.sciencedirect.com/science/article/abs/pii/S0277953601000077?via%3Dihub>

Bleakley, H. (2010). Health, Human Capital, and Development. *Annual Review of Economics* Vol. 2: 283-310. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3800109/pdf/nihms514075.pdf>

Bloom, D. E, Canning, D., Sevilla J. (2001). The Effect of Health on Economic Growth: Theory and Evidence. NBER Working Paper No 8587. Available at: <https://www.nber.org/papers/w8587.pdf>

Borkowsky, A. (2013), "Monitoring Adult Learning Policies: A Theoretical Framework and Indicators", OECD Education Working Papers, No. 88, OECD Publishing. Available at: <http://dx.doi.org/10.1787/5k4c0vxjlkzt-en>

CEDEFOP (2011), Vocational education and training is good for you. The social benefits of VET for individuals, Research Paper No 17, 2011, Luxembourg: Publications Office of the European Union, 2011 Available at: https://www.cedefop.europa.eu/files/5517_en.pdf

Cox D.R. (1958), *Planning of Experiments*, New York: Wiley.

Crawford, R. (1991). *In the Era of Human Capital*. New York: Harpercollins

De La Fuente Á, Domenéch A. (2000), A Human capital in growth regressions: how much difference does data quality make? *Economic Department Working Paper No. 262*, Paris: OECD, 2000 (ECO / WKP (2000) 35). Available at: <http://www.oecd.org/innovation/research/1825500.pdf>

De La Fuente Á, Domenéch A. (2006), Human capital in growth regressions: how much difference does data quality make?, *Journal of the European Economic Association*, March 2006, 4(1), pp.1-36. Available at: <http://www.iae.csic.es/investigadoresMaterial/a9167121855archivoPdf2774.pdf>

Department for International Development (DFID) (2005), *Guidance on Evaluation and Review for DFID Staff*, Evaluation Department July 2005. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/67851/guidance-evaluation.pdf

Department for International Development (DFID) (2009), Building the evidence to reduce poverty. The UK's policy on evaluation for international development., Available at: <https://www.oecd.org/development/evaluation/dcdndep/43183649.pdf>

DEReC (DAC Evaluation Resource Centre) <https://www.oecd.org/derec/guidelines.htm>

DG EMPL (2018), Monitoring and Evaluation of European Cohesion Policy European Social Fund Guidance document, August 2018. Available at: <https://ec.europa.eu/sfc/en/system/files/ged/ESF%20monitoring%20and%20evaluation%20guidance.pdf>

DG Regio (2013), EVALSED - The resource for the evaluation of Socio-Economic Development: Sourcebook – The Guide, European Commission. Available at: http://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/guide_evalsed.pdf

Draghicia, A & Harpana, I (2013), Debate on the multilevel model of the human capital measurement, Procedia - Social and Behavioral Sciences 124 (2014) 170 – 177. Available at: <https://core.ac.uk/download/pdf/82072284.pdf>

Edwards, R. (1979). Contested Terrain: The Transformation of the Workplace in the Twentieth Century. NY: Basic Books.

Escudero, V. (2014) Are active labour market policies effective in activating and integrating low-skilled individuals? An international comparison, MPRA Paper No. 55319, posted 16. April 2014 .Available at: <https://mpra.ub.uni-muenchen.de/55319/> European Commission (2015a) Better regulation guidance, SWD(2015) 111 final, Strasbourg; Available at: <https://ec.europa.eu/info/sites/info/files/better-regulation-guidelines.pdf>

European Commission (2015b), Better Regulation "Toolbox", available at: https://ec.europa.eu/info/sites/info/files/better-regulation-toolbox-2015_0.pdf

European Commission (2015c), Better regulation toolbox #41. Available at: https://ec.europa.eu/info/sites/info/files/file_import/better-regulation-toolbox-41_en_0.pdf

Florida, R. (2002a). The Rise of the Creative Class. New York: Basic Books.

Florida, R. (2002b). The economic geography of talent. Annals of the Association of American Geographers, 92(4): 743–755. Available at: https://www.creativeclass.com/rfcgdb/articles/5%20The_Economic_Geography_of_Talent.pdf

Florida, R. (2002c). Bohemia and economic geography. Journal of Economic Geography, 2: 55–71. Available at: http://creativeclass.com/rfcgdb/articles/6%20Bohemia_and_Economic_Geography.pdf

Florida, R., Mellander, C., Stolarick, K. (2007). Inside the Black Box of Regional Development - human capital, the creative class and tolerance. Working Paper Series in Economics and Institutions of Innovation 88, Royal Institute of Technology, CESIS - Centre of Excellence for Science and Innovation Studies. Available at: <https://static.sys.kth.se/itm/wp/cesis/cesiswp88.pdf>

- Frank, R. H., & Bernanke, B. S. (2007). *Principles of Microeconomics* (3rd ed.). New York: McGraw-Hill/Irwin.
- Franzini, M. and Raitano, M. (2011), “Differenziali salariali e capitale umano: alcune evidenze in cerca di spiegazione”. Available at: <http://www.siecon.org/online/wp-content/uploads/2012/08/Franzini-Raitano1.pdf>
- Garavan, T. N., et al. (2001). Human Capital Accumulation: The Role of Human Resource Development. *Journal of European Industrial Training*, 25(2), 48-68.
- Glaeser, E. L., Kolko, J., and Saiz, A. (2001). Consumer city. *Journal of Economic Geography*, 1(1), 27-50. Available at: <https://doi.org/10.1093/jeg/1.1.27>
- Glaeser, E. L., Maré, D. C. (2001). Cities and skills. *Journal of Labor Economics*, 19(2), 316-342. Available at: <https://dx.doi.org/10.1086/319563>
- Glaeser, E., Scheinkman, J. and Shleifer, A., (1995), *Economic Growth in a Cross-Section of Cities*, No. 5013, NBER Working Papers, National Bureau of Economic Research, Inc. Available at: <http://www.nber.org/papers/w5013.pdf>
- Goldin, C. (2014) *Human Capital*, *Handbook of Cliometrics*. Springer Verlag. Heidelberg, Germany. Available at: https://scholar.harvard.edu/files/goldin/files/goldin_human_capital.pdf
- Golsteyn, B. H.H. and Stenberg, A. (2017) *Earnings over the Life Course: General versus Vocational Education*, Discussion Paper Series, IZA DP No. 10593. Available at: <http://ftp.iza.org/dp10593.pdf>
- Gu, W. and Wong, A. (2010), *Estimates of Human Capital in Canada: The Lifetime Income Approach* (November 19, 2010). Available at: <https://ssrn.com/abstract=1711935> or <http://dx.doi.org/10.2139/ssrn.1711935>
- Idrees, A. S. and Siddiqi, M. W. (2013). Does Public Education Expenditure Cause Economic Growth? Comparison of Developed and Developing Countries. *Pakistan Journal of Commerce and Social Sciences*, Vol.7(1):174-183. Available at: <http://www.jespk.net/publications/116.pdf>
- ILO (2006) *Decent work in the Americas: An agenda for the Hemisphere 2006-2015*. Report of the Director-General. Sixteenth American Regional Meeting, Brasilia, 2006. Available at: <https://www.ilo.org/public/english/standards/relm/rgmeet/16amr/dwork.pdf>
- Kusek, J. Z. and Rist R. C. (2004), *Ten Steps Ten Steps to a Results Based Monitoring and Evaluation System - A Handbook for Development Practitioners*, The International Bank for Reconstruction and Development. Available at: https://www.oecd.org/dac/peer-reviews/World%20bank%202004%2010_Steps_to_a_Results_Based_ME_System.pdf
- Judson, R. (2002). Measuring human capital like physical capital: What does it tell us? *Bulletin of Economic Research* 54(3): 209-231. Available at: <https://doi.org/10.1111/1467-8586.00150>
- Jung, H.-S. and Thorbecke, E. (2003). The impact of public education expenditure on human capital, growth, and poverty in Tanzania and Zambia: a general equilibrium approach. *Journal of Policy Modeling* 25: 701–725. Available at: <https://www.sciencedirect.com/science/article/pii/S0161893803000607>

- Kiker, B. F. (1966) The Historical Roots of the Concept of Human Capital, *Journal of Political Economy*, Vol. 74, No. 5 (Oct., 1966), pp. 481-499. Available at: <https://www.jstor.org/stable/1829595>
- Kraay A. (2018), Methodology for a World Bank Human Capital Index, Background Paper to the 2019 World Development Report - Policy Research Working Paper 8593. Available at: <http://documents.worldbank.org/curated/en/300071537907028892/pdf/WPS8593.pdf>
- Kwon, D. B. (2009), Human capital and its measurement. The 3rd OECD World Forum on “Statistics, Knowledge and Policy” Charting Progress, Building Visions, Improving Life Busan, Korea - 27-30 October 2009, Available at: <http://www.oecdworldforum2009.org>
- Kyriacou, G. A. (1991), Level and Growth Effects of Human Capital: A Cross-Country Study of the Convergence Hypothesis. Working Papers 91-26, C.V. Starr Center for Applied Economics, New York University. Available at: <https://ideas.repec.org/p/cvs/starer/91-26.html>
- Lechner, M. (2010) The Estimation of Causal Effects by Difference-in-Difference Methods, *Foundations and TrendsR in Econometrics* Vol. 4, No. 3 (2010) 165–224 c 2011 M. Lechner DOI: 10.1561/08000000014 Available at https://michael-lechner.eu/ml_pdf/journals/2011_Lechner_DiD_2011_ECO%200403%20Lechner_darf%20aufs%20Netz.pdf
- Leeuw, F., and Vaessen, J. (2009), Impact Evaluations And Development: NONIE Guidance on Impact Evaluation, Draft Version for Discussion at the Cairo conference March-April, 2009. Available at: https://www.researchgate.net/publication/264869403_Impact_Evaluations_and_Development_NONIE_Guidance_on_Impact_Evaluation
- Lepak, D. and Snell, S. (1999). The Human Resource Architecture: Toward a Theory of Human Capital Allocation and Development. *Academy of Management Review*, 24, 31-48. Available at: <https://www.jstor.org/stable/259035?seq=1>
- Lim, S. S. et al. (2018) Measuring human capital: a systematic analysis of 195 countries and territories, 1990–2016, Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA, *Lancet* 2018; 392: 1217–34. Available at: <https://www.ncbi.nlm.nih.gov/pubmed/30266414>
- Lucas, R. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22(1), 3-42. Available at: <https://www.parisschoolofeconomics.eu/docs/darcillon-thibault/lucasmecanicseconomicgrowth.pdf>
- Mankiw N. G., Romer D., Weil D.N. (1992), Contribution to the Empirics of Economic Growth, *The Quarterly Journal of Economics*, Vol. 107, No.2 (May, 1992), pp. 407-437. Available at: https://eml.berkeley.edu/~dromer/papers/MRW_QJE1992.pdf
- Martini A. (2009), How Counterfactuals Got Lost on the Way to Brussels, Prepared for the Symposium “Policy and programme evaluation in Europe: cultures and prospects” Strasbourg, July 3-4, 2008, available at: https://www.researchgate.net/publication/237456114_HOW_COUNTERFACTUALS_GOT_LOST_ON_THE_WAY_TO_BRUSSELS

- Mellander, C. and Florida, R. (2007). Human Capital or the Creative Class – Explaining Regional Development in Sweden. KTH/CESIS Working Paper Series in Economics and Institutions of Innovation. Available at: <https://static.sys.kth.se/itm/wp/cesis/cesiswp79.pdf>
- Mincer, J. (1958) “Investment in Human Capital and Personal Income Distribution” *Journal of Political Economy*, 66(4), 281-302. Available at: <https://www.jstor.org/stable/1827422?seq=1>
- Ministry for Foreign Affairs Department for Development Policy (DDP), Guidelines for Programme Design, Monitoring and Evaluation, Katajanokanlaituri 3 PL 176, 00161 Helsinki, Finland available at: <https://www.oecd.org/derec/finland/38141776.pdf>
- Moretti, E. (2004). Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data. *Journal of Econometrics*, Vol. 121, No. 1-2, p. 175-212. Available at: <https://ideas.repec.org/a/eee/econom/v121y2004i1-2p175-212.html>
- Moretti, E. (2004). Human Capital Externalities in Cities, *Handbook of Urban and Regional Economics*, North Holland-Elsevier. Available at: <https://www.sciencedirect.com/science/article/pii/S1574008004800087>
- Mulligan, C., and Sala-i-Martin, X. (1995). Measuring Aggregate Human Capital. NBER Working Paper No. 5016. Cambridge, MA: National Bureau of Economic Research. Available at: <https://econ-papers.upf.edu/papers/112.pdf>
- OECD (1999), *OECD Science, Technology and Industry Scoreboard 1999: Benchmarking Knowledge-based Economies*, OECD Publishing, Paris, Available at: https://doi.org/10.1787/sti_scoreboard-1999-en.
- OECD (2001), “The Well-being of Nations: The Role of Human and Social Capital.” Paris: OECD,
- OECD DAC (2002), *Glossary of Key Terms in Evaluation and Results Based Management*, DAC Working Party on Aid Evaluation (WP-EV), available at: <https://www.oecd.org/derec/dacnetwork/35336188.pdf>
- Psacharopoulos, G. (1985). Returns to Education: A Further International Update and Implications, *The Journal of Human Resources* 20 (4): 583- 604. available at: <https://www.jstor.org/stable/145686?origin=crossref&seq=1>
- Psacharopoulos, G. (1994) Returns to Investment in Education: A Global Update, *World Development*, 22 (9): 1325–1343. Available at: <https://www.sciencedirect.com/science/article/abs/pii/0305750X94900078>
- Psacharopoulos, G. and Patrinos, H.A. (2018) Returns to investment in education: a decennial review of the global literature, *Education Economics*, 26:5, 445-458, Available at: <http://datatopics.worldbank.org/education/files/GlobalAchievement>ReturnsInteractive.pdf>
- Qian, H., Acs, Z., and Stough, R. (2012), Regional systems of entrepreneurship: The nexus of human capital, knowledge and new firm formation. *Journal of Economic Geography*. 13. Available at: <http://hdl.handle.net/10.1093/jeg/lbs009>

- Rauch, J. (1993), Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities, *Journal of Urban Economics*, 34, issue 3, p. 380-400. Available at: <https://www.sciencedirect.com/science/article/abs/pii/S0094119083710429>
- Riley G. (2012), *Economic Growth - The Role of Human & Social Capital, Competition & Innovation*.
- Roberts, A. R., Yeager, K. (2004) *Evidence-Based Practice Manual: Research and Outcome Measures in Health and Human Services*, Oxford University Press. Available at: <https://global.oup.com/academic/product/evidence-based-practice-manual-9780195165005?cc=it&lang=en&>
- Rodriguez, P. J., & Loomis, R. S. (2007). A New View of Institutions, Human Capital, and Market Standardization. *Education, Knowledge & Economy*, 1(1), 93–105. Available at: <https://www.tandfonline.com/doi/abs/10.1080/17496890601128357>
- Rogers, P. (2014). *Theory of Change, Methodological Briefs: Impact Evaluation 2*, UNICEF Office of Research, Florence, Available at: <https://www.unicef-irc.org/publications/747-theory-of-change-methodological-briefs-impact-evaluation-no-2.html>
- Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5), 71–102. Available at: <https://www.jstor.org/stable/2937632?seq=1>
- Rosen, H. S. (1999). *Public Finance*. New York: McGraw-Hill
- Rubin, D.B. (2007) The design *versus* the analysis of observational studies for causal effects: Parallels with the design of randomized trials. *Statistics in Medicine*. 2007; 26:20–36 Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.2739>
- Schultz, T. W. (1961). Investment in Human Capital. *American Economic Review*, 51, 1-17. Available at: <https://www.jstor.org/stable/1818907?seq=1>
- Sheffrin, M. S. (2003). *Economics: Principles in Action*. New Jersey: Pearson Prentice Hall
- Smith, A. (1776), “An Inquiry into the Nature and Causes of the Wealth of Nations”, W. Strahan and T. Cadell, London, Available at: <https://quod.lib.umich.edu/e/ecco/004861571.0001.001?rgn=main;view=fulltext>
- Solow, R. M. (1956), A Contribution to the Theory of Economic Growth, *Quarterly Journal of Economics*, LXX (1956), 65-94. Available at: <https://academic.oup.com/qje/article-abstract/70/1/65/1903777?redirectedFrom=fulltext>
- Stiglitz, J., Sen, A., and Fitoussi, J. (2009). *Report of the Commission on the Measurement of Economic Performance and Social Progress (CMEPSP)*. Available at: <https://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report>
- Thistlethwaite, D.L. and Campbell, D.T (1960) “Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment.” *Journal of Educational Psychology*, 51(6): 309–17.
- UNDP (2002), *Guidelines for Outcome Evaluators: Monitoring and Evaluation Companion Series, #1*, Evaluation Office United Nations Development Programme, New York. Available

at: <http://web.undp.org/evaluation/documents/HandBook/OC-guidelines/Guidelines-for-OutcomeEvaluators-2002.pdf>

UNECE Task Force on Measuring Human Capital (2016), Guide on Measuring Human Capital, UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE - ECE/CES/STAT/2016/6, United Nations, 2016. Available at: https://www.unece.org/fileadmin/DAM/stats/publications/2016/ECECESSTAT20166_E.pdf

UNSD (2008), System of National Accounts 2008, United Nations, New York. Available at: <https://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>

Veenstra, G. (2001). Social Capital and Health. Canadian Journal of Policy Research, Spring, 72–81. Available at: https://www.researchgate.net/publication/287114218_Social_capital_and_health

Veenstra, G., et al. (2005). Who You Know, Where You Live: Social Capital, Neighbourhood and Health. Social Science and Medicine, 60, 2799–2828. Available at: <https://www.ncbi.nlm.nih.gov/pubmed/15820588>

Victora, C. G., Adair, L., Fall, C., Hallal, P. C., Martorell, R., Richter, L., and Sachdev, H. S. (2008). Maternal and child undernutrition: consequences for adult health and human capital. Lancet, 371(9609), 340-357. Available at: <https://www.ncbi.nlm.nih.gov/pubmed/18206223>

World Economic Forum (WEF) (2017). – The Global Human Capital Report 2017, available at: http://www3.weforum.org/docs/WEF_Global_Human_Capital_Report_2017.pdf

Wilson, K., et al. (2004). Linking perceptions of neighbourhood to health in Hamilton, Canada. Journal of Epidemiology and Community Health, 58, 192–198. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1732692/pdf/v058p00192.pdf>

Youndt, M. A., et al. (2004). Intellectual Capital Profiles: An Examination of Investments and Returns. Journal of Management Studies, 41(2). 335-361. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-6486.2004.00435.x>

Zall Kusek, J., Rist, R. C. (2004). Ten Steps to a Results-Based Monitoring and Evaluation System : A Handbook for Development Practitioners. Washington, DC: World Bank. © World Bank. Available at: <https://openknowledge.worldbank.org/handle/10986/14926>

2 Chapter 2: Case study analysis on the monitoring and evaluation systems of selected EU programmes for human capital development

Abstract

Reconciling the need for exhaustive monitoring and evaluation (M&E) of EU funded programmes in the field of Human Capital (HC) development and that of M&E systems which are (a) comprehensive and relevant, (b) timely, (c) proportionated to their aims and not burdensome whilst (d) respecting the dignity of end-beneficiaries and privacy issues, is not a trivial job. Yet, the discussion on M&E systems doesn't echo much in the scientific literature and it is typically left to be discussed among practitioners concerned with either the design or evaluation/ audit of the funds (and the institutions they represent). This means that streamlined analytical tools are seldom used in understanding whether the M&E system is ultimately fit for purpose from the perspective of a modern concept of human capital development. This research aims to fill this gap and contribute to the debate on the M&E systems of three pivotal EU HC funds of the 14-20 programming period, namely the European Social Fund (ESF), the Youth Employment Initiative (YEI) and the Fund for European Aid to the Most Deprived (FEAD), with a view to their revision in the post-2020 programming period. It also goes beyond those to check, through content analysis, whether the ESF/YEI monitoring system can help track progress towards the objectives of the New Skills Agenda for Europe, of which is the main financial instrument. The methodological tools used for the case studies include the reconstruction of the intervention logic and the analysis of the M&E systems vis-à-vis the checklist for the M&E of HC policies devised in chapter 1 on the basis of the related literature review. Findings point to frequent gaps in the EU M&E systems which pass mostly unnoticed also due to the lack of dedicated literature.

Introduction

Motivation for the research and gaps identified in the literature

Based on the findings from chapter one a checklist was developed to aid the analysis of some specific EU programmes which aim at developing human capital. This is considered key as the literature on the M&E of the ESF/YEI and FEAD is either scarce or absent, respectively. In the case of the former, a study commissioned by DG Employment to Applica et al (2018) on the M&E of the ESF exists, but it lacks the necessary depth and an analytical framework to understand whether the system is fit for purpose. To name but a few issues, it does not assess the M&E system against the governing principles of a M&E system enumerated by the better regulation guidelines, nor it assesses its suitability to measure human capital based on the literature on the theme. The focus lies mainly on concrete issues experienced by stakeholders in dealing with it and on how to contain the administrative burden thereof whilst improving accuracy. In addition, it does not fully take into account the needs of evaluating the impacts of the ESF from a macroeconomic perspective. The analysis is thus very informative but needs to be extended and deepened given the aims of this research project. With respect to the FEAD, no dedicated study was ever commissioned or carried out that dealt with the assessment of its M&E system. This was indirectly assessed in the framework of the FEAD mid-term evaluation

and discussed in a dedicated report from the European Court of Auditors, as it will be discussed in section 2.3.

Structure of the chapter

After having identified the scope for EU action in the field of human capital and its main initiatives to this end (in annex I for conciseness), the assessment grid developed in chapter 1 is used throughout this chapter to zoom in on three of the biggest EU's funds in the field of human capital development, notably the European Social Fund (ESF), the Youth Employment Initiative (YEI) the Fund for European Aid to the Most Deprived (FEAD). The goal is to understand whether these systems are fit for purpose, especially from the perspective of a modern understanding of human capital development, and to provide recommendations on their future design.

In particular, section 2.1 is concerned with the identification of the scope of the analysis, that is, to the determination of the most relevant funds and initiatives aimed at human capital development in the EU. Section 2.2 concerns the assessment of the ESF/YEI monitoring system (first case study) and section 2.3 the assessment of the FEAD monitoring system (second case study). Finally, section 2.4 draws the key conclusion across the two case studies.

As to the structure of each case study in section 2.2 and 2.3, the starting point is a description of the goals and structure of each programme. Then, the focus shifts on the monitoring arrangements. Because, as highlighted in chapter 1, the key element for the creation of a monitoring system is the reconstruction of the programmes' intervention logic, this is done immediately afterwards. Thus, all the elements are present for an assessment of the monitoring arrangements also based on the available studies as well as the author's field experience. Once the key issues are clarified, the monitoring systems can be assessed against the assessment grid/checklist developed in chapter 1. Because the 14-20 programming period is about to end, the two case studies go on to look into changes which are being discussed to the fund's future monitoring arrangements to ascertain whether and to what extent these go in the recommended direction. Lastly, policy recommendation are provided.

In addition, an annex is included that is linked to this chapter, notably Annex II. This concerns the analysis of the monitoring system of the ESF/YEI and its potential to track progress towards the aim of the New Skills Agenda for Europe, including through content analysis (NVivo). Its findings are subsumed in the first case study, section 2.2.4.

Specificities of this chapter

The analysis presented below is corroborated by the work experience accrued with Fondazione Giacomo Brodolini throughout these years that included working on the monitoring and evaluation of the following funds:

- the Mid-term evaluation of the REC and JUST programmes;
- the Mid-term evaluation of the FEAD;
- the ESF monitoring and performance reports 2015-2018;

- the Study supporting the Impact Assessment on Human Capital Investments (ESF, YEI, FEAD, EaSI, EGF);
- the Mid Term Evaluation ESF's TO8 investments; and
- the Mid Term Evaluation of ESF's Youth Employment investments (ESF and YEI IP 8.ii).

This work experience has greatly aided the understanding of the many elements which concur to define policy making in practice, including the issues related to (i) data collection, storing and reporting, (ii) the needs of the different stakeholders and (iii) the difficulties which lie in real-world evaluation of such large and complex programmes. Although no dedicated field research was carried out within this strand of the research project, the author has contributed to the Mid Term Evaluations (commissioned by DG EMPL) of the funds discussed in this chapter and has thus participated in a range of meetings, focus groups, steering committees, conferences and debates related to the M&E of the programmes. He has also taken care of the data analysis as well as contributed to the drafting of the conclusions and recommendations of the above-mentioned reports, which was a key opportunity to reflect on the strengths and weaknesses of the monitoring and evaluation systems under examination.

What follows is thus a personal research effort that builds upon the experience accrued dealing with the funds' M&E systems.

Importantly, the level of detail of this chapter is rather high, also due to the direct "on field" experience working with the monitoring systems assessed. However, only at this level of detail it becomes possible to effectively unearth the issues which hamper the collection of evidence on the programmes. A too broad-cut approach would overlook some key aspects, and this research argues that this is what have happened in the few recent studies on the subject-matter. This is why this level of detailed is deemed as necessary for the analysis.

2.1 HC in the EU and selection of case studies

The starting point of the overview, presented in Annex I, was that an up-to-date definition of human capital should look beyond "mere" monetary elements of human capital development, to embrace more all-encompassing frameworks of measurement such as those developed by the WEF and the World Bank, also in combination with analyses which place greater emphasis on health (especially functional health, which has a bearing on productivity), non-cognitive skills, citizenry and the likes.

Based on this definition, the first finding is that Human Capital Development is one of the main aims of the EU and it is reflected in its framework strategy Europe 2020 as well as in significant thematic strategies thereof.

As a consequence, human capital is a recurring theme in a number of European policy priorities and initiatives, and, in turn, a substantial part of the EU budget is dedicated to human capital development activities.

As detailed in Annex I, 3 out of 5 headline targets of the Europe 2020 strategy are directly affected by human capital development, namely those on employment, education as well as poverty prevention. Part of these targets are further specified in the Education and Training 2020 targets, which is the current reference for upskilling policies in the EU.

Human capital is prominent also in 4 of the 7 flagship initiatives linked to Europe 2020, and is one key element of the more recent Juncker’s political guidelines for the EC 2014-2020. As part of those guidelines, a new flagship initiative was launched in 2016, namely a “*New Skills Agenda for Europe*” (COM, 2016) which fully focuses on human capital development in the EU.

Renovated impetus on developing human capital can be found in the adoption in 2017 of the European Pillar of Social Rights, whose foundation element is indeed human capital in its different guises. The social scoreboard is a relevant tool to monitor human capital, as it encompasses some of its key dimensions also in a broad sense, and especially unmet health needs, whilst remaining slightly less comprehensive on the costs and quality of human capital development and on soft, non-cognitive skills.

A range of more detailed data sources exist for the EU that take into account human capital-related trends and data, such as the European Skills Index (Cedefop) and a range of indicators from CEDEFOP which further extend the scope of the monitoring on skills quality, obsolescence, future trends, non-cognitive skills, skills utilisation etc. as well as Eurofound (see, e.g., the Tasks within occupation data).¹¹ Finally, also the OECD data can be leveraged upon for 20 MS to have data on the quality of education (PISA and PIAAC).

The overview of the all-important ‘financial lever’ of the EU, notably its programmes or “funding instruments”, described in detail in Annex I made it clear that the main and more all-encompassing instrument for the development of human capital is the European Social Fund (ESF), accompanied by funds such as Erasmus +, the Youth Employment Initiative (YEI), the EaSI – to name but a few.

However, a wide array of other funds and programmes are to be considered relevant within this project, as they may potentially cover different aspects of human capital, there including, e.g. enabling factors for social inclusion (i.e. FEAD), health (i.e. Health for growth), scientific excellence (Horizon 2020), human rights (i.e. the rights, equality and consumer protection programme).

The selection of case studies for this chapter is based both on a criterion of relevance with respect to the purpose of developing human capital as well as the fact that the author gained experience, during his internship at Fondazione Giacomo Brodolini, with the actual monitoring and evaluation of the initiative. Additional criteria have been to ensure that a certain variety of approaches was ensured as well as the fact that the funds are expected to support the New Skills Agenda. The reason for the latter is that such initiatives don’t typically come with dedicated budget attached but rather aim at steering resources from existing funds towards their aims.

Thus the ESF/YEI was chosen as a case study as it is the most relevant funding instrument for the support of human capital in the EU and also the a key financial lever for the implementation of the New Skills Agenda for Europe. In addition, the FEAD was chosen as it is a funding initiative which has an “enabling” nature, in the sense that should act to remove obstacles preventing people from getting into social inclusion, education, employment and so forth. So it

¹¹ See for more at: <https://skillspanorama.cedefop.europa.eu/en/indicators/tasks-within-occupation>

lies at the root of human capital formation and embodies the broader understanding of the concept of human capital as discussed in chapter 1.

2.2 The ESF/ YEI monitoring system

2.2.1 Introduction and key info on the structure and intervention logic of the ESF/ YEI

The **European Social Fund (ESF)** is one of the five European Structural and Investment Funds (ESIF), managing the investment in Europe's human capital. ESIF are structured across 11 thematic objectives (TOs),¹² which are further broken down into Investment Priorities (IPs) which are, in turn, detailed into Specific Objectives (SOs). While the articulation and wording of TOs and IPs is set out by the ESF regulation, the wording of each SOs is left to the Managing Authorities (MAs). The Managing Authorities are the implementing bodies of the 187 Operational Programmes, and, amongst other, are also in charge of collecting, storing and making available data from the monitoring system (further references on the monitoring system are provided in a dedicated section below).

Among the 11 thematic objectives of these funds, ESF concentrates on number 8 (promoting sustainable and quality employment and supporting labour mobility), 9 (promoting social inclusion, combating poverty and any discrimination), 10 (investing in education, training and vocational training for skills and lifelong learning), and 11 (enhancing institutional capacity of public authorities and stakeholders and efficient public administration).¹³

Every thematic objective is comprised of a number of investment priorities, further elaborating on the issues listed in thematic objectives names. For instance, for TO8, it ranges from access to employment to equal opportunities and active aging; for TO9, from active inclusion to provision of high quality service; for TO10 from prevention of early school leaving to promoting life-long learning; and so forth.

The budgets of the thematic objectives, as well as the expected results, are defined at the investment priority level (and further broken down by category of region), and differ from country to country, according to their needs (more specifically, depending on the countries regional needs and allocations). For instance, 9% of ESF budget is allocated to investment priority 10.i, but this investment priority represents more than 30% of ESF funded activities in Austria, while not being present in a number of countries (i.e. Croatia, Cyprus, Denmark, Finland, Ireland, Luxembourg, Slovenia, and Sweden).

Investment priorities are then further defined through specific objectives, which are set at national or regional level according to the Member States. Every investment priority has at least

¹² 1. Strengthening research, technological development and innovation; 2. Enhancing access to, and use and quality of information and communication technologies (ICT); 3. Enhancing the competitiveness of small and medium-sized enterprises (SMEs); 4. Supporting the shift towards a low-carbon economy in all sectors; 5. Promoting climate change adaptation, risk prevention and management; 6. Preserving and protecting the environment and promoting resource efficiency; 7. Promoting sustainable transport and removing bottlenecks in key network infrastructures; 8. Promoting sustainable and quality employment and supporting labour mobility; 9. Promoting social inclusion, combating poverty and any discrimination; 10. Investing in education, training and vocational training for skills and lifelong learning; 11. Enhancing institutional capacity of public authorities and stakeholders and efficient public administration

¹³ http://ec.europa.eu/regional_policy/en/policy/what/glossary/t/thematic-objectives

one specific objective. Given that budget is decided at investment priority level, specific objectives are useful to define strategic aims, rather than include hard caps to the policies. Finally, specific objectives are articulated in a number of actions /typologies of support offered, which further define the activities that will be implemented.

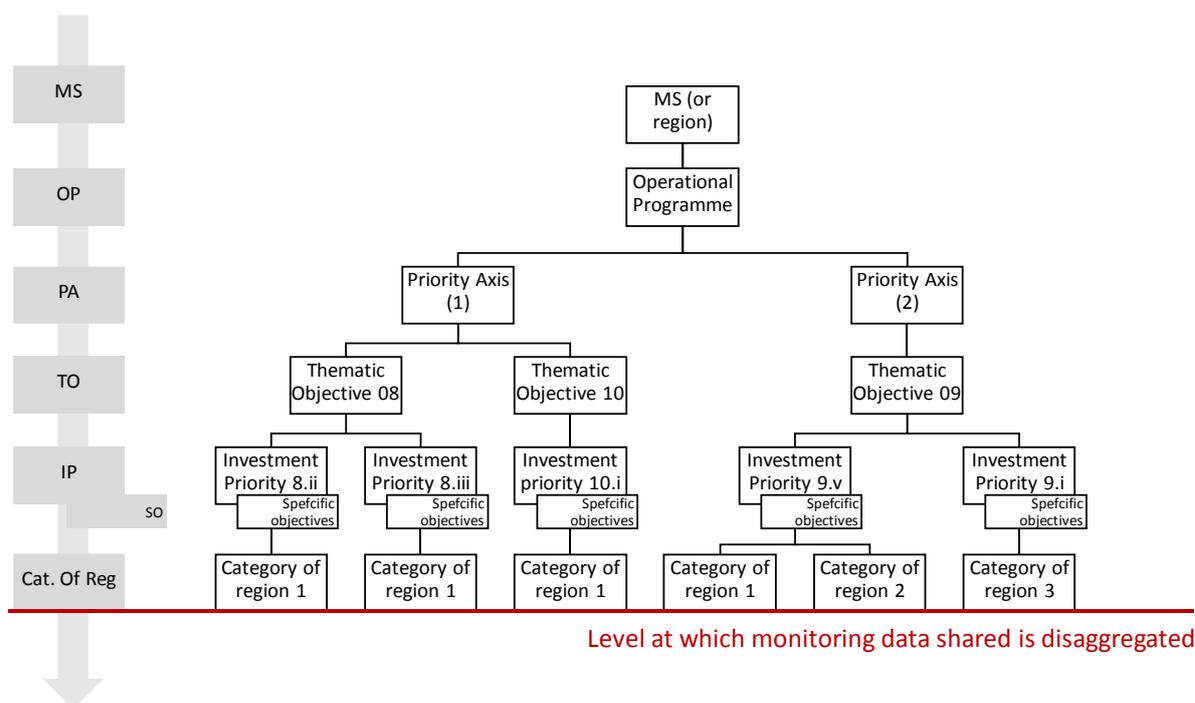
After having gone through the ESF structure from EU perspective, it is important to see the point of view of Member States. These organise European funds through so-called Operational programmes, that can be at national or at regional level. Operational programmes receive funds from several European budget lines, and the managing authorities at national or regional level tend to have an internal way of structuring the foreseen interventions, clustering funds according to the policies they have in mind through the priority axes structure.

Alongside the European Social Fund, another funding instrument, namely the **Youth Employment Initiative**, was established by the Common Provision Regulation in 2013 that, whilst having a dedicated set of eligibility criteria and dedicated resources, contributes to achieving the aims of Thematic Objective 8, more specifically within a single investment priority, IP 8.ii – “Sustainable integration into the labour market of young people (YEI), in particular those not in employment, education or training, including young people at risk of social exclusion and young people from marginalised communities, including through the implementation of the Youth Guarantee”.

Initially foreseen for the 2014-2015 period (with implementation up to 2017) YEI funding has been renovated for those regions still lagging behind in terms of youth unemployment rate in 2015 and will now last until 2023

Both funds share a common structure, despite some minor differences, which is illustrated in the figure below.

Figure 5- Structure of the ESF/YEI, from a monitoring perspective



Source: author's elaboration based on the fund regulations

This structure displays the standard structure of the ESF/YEI from a monitoring data perspective, although some slight difference exists between the programmes and also depending on whether these are defined at the national or regional level.

For instance, investment from the YEI is entirely channelled into one investment priority, namely investment priority 8.ii within Thematic Objective 8, to which it contributes together with ESF money. In addition, investment in YEI is not disaggregated by category of region.

Lastly, although specific objectives are defined that break down investment priorities, indicators are not linked to their disaggregation. Hence, specific objective represent a breakdown of the investment priorities which is however not reflected in the disaggregation of monitoring indicators, nor of financial resources, but only serves to better displays the programme's intervention logic.

Lastly, micro-data might be collected at a lower level than the IP level, notably at that of the single operation funded. But, as it will be said below, micro-data is only theoretically in the hands of MAs, but due to dimensionality as well as privacy issues it has never been aggregated across-OPs and MS to have a fully blown EU picture.

2.2.2 ESF /YEI monitoring system: main features and indicators

As anticipated, Managing Authorities are in charge of the Monitoring and Evaluation of each Operational Programme.

While ex-ante, in itinere and ex-post evaluations form by all means part of the monitoring and evaluation system of an Operational Programme, these fall outside the scope of this analysis, which focuses on the monitoring of the operational programmes.

The ESF/YEI monitoring is defined in detail in in the Common Provision Regulation 1304/2014 and ESF regulation 1303/2014. Information is included concerns a number of issues, such as coverage, quality, disaggregation, transmission and storage of data, but also reporting and dissemination requirements (Applica et al, 2018). Indicators are the main instrument of monitoring.

In general, one can say that there are:

- 3 main typologies of indicators, i.e:
 - Financial indicators (they measure financial allocations, the cost of operations selected for support and the eligible expenditure declared by beneficiaries to managing authorities)
 - Output indicators: they measure the number of participations of either individuals or entities to the supported activities
 - Result indicators: they measure either immediate or longer term (6-monthths) results which could be put in relation to the supported activities
- 3 main reporting tools
 - List of operations and beneficiaries: this is the only publicly available source of information which contains data at the level of the operations, thus at the sub-

Investment Priority level. The information included only concerns financial indicators, together with a brief description of the operation/project supported as well as name of the beneficiary.

- Annual Implementation Reports: they are drafted by MAs at the for each OP and include both quantitative data from the above indicators as well as qualitative information (e.g. what activities have been supported at the level of each priority axis, main issues affecting performance, achievements, a summary of the evaluations and evaluation plan, support to transversal objectives – gender equality, poverty, social innovation etc)
- Progress reports: they are drafted at the MS level and present a summary of the countries’ progress and achievements, including by looking at contextual factors, recommendations etc.

Information (except list of operations and beneficiaries) is typically made available either through the Managing Authorities websites or through the System for Fund Management in the European Union SFC2014. The latter is the electronic exchange system of the European Commission where data from different reports and monitoring activities is collated. It serves a range of purposes but when it comes to ESF/YEI monitoring data in particular it:

- collates financial data from Managing Authorities three times per year;
- includes progress reports, annual implementation reports and other sources of information.

When it comes to the result-orientation of the ESF/YEI, clearly output and result indicators are the focus of the analysis. However, relevant info should be drawn also at the level of input, that is, not only what has been achieved but through which activities and where.

Indicators are proxies to capture outputs and results. Their limitations are acknowledged, in that they can only to a limited extent capture the complexity of the programmes and their effects.

According to the “Guidance document on Monitoring and Evaluation (ESF and the European Regional Development Fund), Concepts and Recommendation” (DG EMPL, 2018a), which summarises extant knowledge and practices in the field of monitoring and evaluation of public policy and, particularly, ESI funds, they are nevertheless useful to serve one of the two main objectives of monitoring and evaluation, namely to help deliver a programme in an efficient manner (management) while they can tell very little on whether a programme has produced the desired effects.

As for the monitoring system of the ESF/YEI 2014-2020, this is based on a number of indicators that are defined ex-ante partially at the EU level and, for the other part, for each Operational Programme.

The two main families of indicators are:

- i. Common indicators. These were defined at EU level by the official regulation published in December 2013,¹⁴ and they are 33;

¹⁴ <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32013R1304> ; see annex 1.

- ii. Specific indicators. These are selected by Member States themselves, according to the type of programme they have designed. Among the recommendations of the Common Provision Regulation¹⁵ there was limiting their number and making them as adherent as possible to the objectives of the programmes. Yet the national managing authorities came up with a large number of specific indicators, over 3000 thousand (FGB et al, 2016).

Both common and specific indicators can be distinguished between output and result indicators. Common indicators have further internal classification, as follows:

- Common output indicators
 - On participants (individuals)
 - On entities
- Common result indicators
 - Immediate (upon leaving support)
 - Longer term (six months after leaving support)

In general, output indicators are aimed at capturing the number of beneficiaries treated (e.g. long term unemployed, above 54 years old, migrants, a given type of entity), while the result ones deal with effects after the treatment (e.g. participant in employment, including self-employment).¹⁶

Every indicator should mention the breakdown by gender.

Managing Authorities are required to select from their indicators at least one indicator per specific objective¹⁷ upon which to set a target. This also mean that not all indicators for which data is collected will have a target and, in turn, that aggregating targets would not reflect the overall volume of activities carried out under the ESF (and related results achieved).

The monitoring system of the YEI works along the same lines of that of the ESFs, although there are some differences in the result indicators and reporting requirements, which will be highlighted further below.

Crucially, these indicators are registered (and reported in the Annual Implementation report by Managing Authorities with a lower level of disaggregation than their actual collection. In fact, *“the Regulations require Member States to collect and store detailed data about each participant (i.e. each individual that benefits directly from ESF support), although no specific*

¹⁵ <https://ec.europa.eu/digital-single-market/en/news/eu-regulation-common-provision-regulation-cpr> .

¹⁶ Results indicators for participants have a further division, time based, between immediate and longer term results. Therefore we will have as immediate indicator we will have “participants in employment, including self-employment, upon leaving”, while as longer term indicator we will have “participants in employment, including self-employment, six months after leaving”.

¹⁷ Managing authorities are also required to define at least one specific objective per investment priority. This means that each investment priority will have at least one specific objective and one (related) indicator (be that common or programme-specific) selected.

requirements are set in terms of how, or at which level, this information should be maintained” (Alphametrics and Applica, 2016:6).

This means that only aggregated data are available through SFC for the Commission and the broader audience of relevant stakeholders. This is why background information on participants that may have had some result thanks to the ESF may hardly be retrieved. This was clearly mentioned also in the frame of the study “Pilot and feasibility study on the sustainability and effectiveness of results for European Social Fund participants using Counterfactual impact evaluations”, recently concluded for DG EMPL by Ismeri Europa, Ecorys and Ies, that further reported the answers of some Managing Authorities when such micro-data was requested “*In Spain, it is specified that external agencies cannot have access to data; in Malta it is stated that ‘microdata cannot be provided, but only aggregate figures can be released’*” (Ismeri et. Al, 2019:80).

2.2.2.1 Common vs specific indicators

In general, **common indicators**, as they report in an aggregated fashion microdata on participants, be them output, immediate or longer-term results, take into account only a single socio-economic feature of a participant (age, skills level, social condition etc.) and do not provide information on the type of support granted to the participant.

Example (common output indicator): number of participants in the interventions who are unemployed.

Conversely, *‘programme-specific indicators can be very specific in order to highlight certain aspects of the ESF support which are of particular importance for the Member States/regions. A programme-specific indicator consisting of more than one dimension also allows the managing authority and the monitoring committee to monitor more specifically than the common indicators selected specific aspects of programme implementation’*. (DG EMPL, 2018a)

Example (programme-specific output indicator): number of participants who are unemployed, below 25 y.o. and have received training for basic skills.

On the flipside, indicators that are designed to report detailed information on the level of progress towards a MS/regional specific objective, can hardly be aggregated. Multi-dimensionality is per se an issue when it comes to aggregating. In addition, attempts at categorising Programme-specific indicators (see for instance FGB, 2018) have shown that due to inherent differences in how indicators are designed or measured (e.g. unit of measure, it could be people involved but also hours of training granted) any aggregations would not ensure sufficient reliability.

Based on the common and specific indicators, MAs set their targets. However, while targets are set only for indicators that are “selected”, MAs will nevertheless have to report progress on all 33 common indicators, as well as on the common and specific indicators they have chosen.

From a screening of common and specific indicators (FGB et al, 2016) it is possible to notice that the vast majority of the indicators selected are programme-specific (68% of the total). As for the heterogeneity across Member States, this is relatively high. Poland, the greatest

beneficiary of ESF, uses common indicators for 10%, while Italy, Finland, and Spain (the three countries having done broader use of common indicators) score 75, 68, and 57% respectively.

2.2.2.2 Common output indicators¹⁸

Common output indicators measure, in principle, the number of participants that have received support through the ESF/YEI, irrespective of the results this might have produced. They are also intended to report background information on ESF/YEI participants, across the following classes:

- **Employment status** (unemployed, including long term-unemployed – of which, long-term unemployed; inactive - of which, inactive not in education or training; employed, including self-employed)
- **Age** (below 25 y.o., above 54 y.o. – of which, unemployed or inactive)
- **Educational attainment** (ISCED1-2, ISCED3-4, ISCED 5-8. Indirectly also the number of “no isced” can be reconstructed)
- **Disadvantaged participants** (who live in jobless households (of which, with dependent children); migrants, people with a foreign background, minorities; with disabilities; other disadvantaged; homeless; from rural areas)

It may be that in some cases information which is sensitive cannot be collected. This explains why there may be a difference between the “grand total of participants”, and the total of other classes.

Importantly, the ESF/YEI monitoring guidance clarifies it that “*Values for all common indicators have to be reported for all investment priorities (IP) chosen. Zero values may be reported in the case that no relevant value was recorded under that indicator in the IP concerned.*” (Alphametrics and Applica, 2016:15). This means that each participant will necessarily be recorded in the monitoring system of the ESF/YEI through a common indicator. The same applies to results indicators, which are described in the section below: if any of its participations have produced one of the results considered by common result indicators, these will necessarily be traced (even where a different programme-specific indicator had been selected with a target).

2.2.2.3 Focus: ESF common result indicators

“*Result indicators, both immediate and longer-term, aim to identify a change in the situation of participants compared to that on entry to the operation. In other words, to show the number of people for whom the operation might have had some result*” (Alphametrics and Applica, 2016:57).

Thus, they seek to provide some rough measures of changes to which the funds have contributed.

DG EMPL’s ESF monitoring guidance is actually a little more optimistic about it, and it alludes to the fact that effects could be somewhat measured through result indicators, with the following phrasing “*Result indicators capture the expected effects on participants or entities brought*

¹⁸ The full list of common output indicators is reported in Annex I

about by an operation. Result indicators shall correspond to the specific objectives set out for each investment priority selected. Result indicators go beyond output indicators in so far as they capture a change in the situation, in most cases related to supported entities or participants, e.g. in their employment situation. In order to minimise external factors influencing the value reported under the result indicators, it is advisable to set indicators which are as close as possible to the activities conducted under the respective investment priority. This implies that the effects on participants or supported entities are captured and not the overall effects on a certain group of society or categories of entities which might include people or entities who did not benefit from ESF support. Results can be immediate or longer-term. Data for result indicators are collected through various means and in varying intervals.” (DG EMPL, 2018a:7). So a slightly different emphasis on the causal nexus emerges between the Practical Guidance, written by an external contractor and DG EMPL’s general guidance, with the latter being a lot more optimistic about the “counterfactual” relevance of monitoring indicators.¹⁹

The main difference between immediate and longer-term indicators is the time at which the educational or employment status of the participant is recorded:

- **immediate result indicators** record the situation in which the participants find themselves within 4 weeks of the end the support. This applies also in case of the participant drops early (e.g. before the end of the training module); and
- **longer term result indicators** record the situation six months after the end of the support.

Importantly, while information on immediate result indicator is collected on each participant, sampling approaches are allowed for longer-term result indicators, as it would be too burdensome to trace back all participants at that time.

In principle, common result indicators are mutually exclusive, except in the case of the disadvantaged - who can be recorded twice. This means that if one inactive participant has gained employment after a training the result indicator “inactive participants engaged in job searching upon leaving” would not be relevant although they have, inevitably, started to seek employment before finding it. However, one participant may have both an educational as well as an occupational result. In that case, both are recorded.

Common result indicators under the ESF are defined as per below.

Table 2 – ESF Common Immediate result indicator: detailed description

Indicator		Description	Reference population
Code	Name		
CR01	inactive participants engaged in job	change in the employment status upon leaving (within 4 weeks), compared to the situation when entering the ESF operation (with the participant being inactive, not	inactive participants

¹⁹ In general, it is useful to report what it is included in the Practical note on the monitoring system of the ESF “It should be clear, however, that result indicators are simply observations of the situation and should not necessarily be interpreted as evidence of a successful (or not) operation. If the operation involves regular training offering a qualification then the indicator on participants gaining a qualification can be interpreted as reasonably clear indicator of success (or not) but in general the impact of an operation can only be assessed through careful evaluation.”

Indicator		Description	Reference population
Code	Name		
	searching upon leaving	engaged in job searching for work, when entering the ESF operation)	
CR02	participants in education/training upon leaving	Persons who have received ESF support and who are newly engaged in education (lifelong learning, formal education) or training activities (off-the-job/in-the-job training, vocational training, etc.) immediately upon leaving the ESF operation (within 4 weeks). The source of funding of the training is not relevant	all participants, except participants who were engaged in education/training upon entering
CR03	participants gaining a qualification upon leaving	Persons who have received ESF support and who gained a qualification upon leaving the ESF operation (within 4 weeks) Qualification means a formal outcome of an assessment and validation process which is obtained when a competent body determines that an individual has achieved learning outcomes to given standards. Only qualifications which have been achieved as a result of an ESF intervention should be reported. They should be reported only once per participant/operation.	all participants
CR04	participants in employment, including self-employment, upon leaving	This indicator is to be understood as a change in the employment status upon leaving, compared to the situation when entering the ESF operation (with the participant being unemployed or inactive when entering the ESF operation).	- unemployed - inactive participants
CR05	disadvantaged participants engaged in job searching, education/training, gaining a qualification, or in employment, including self-employment, upon leaving	This looks at all results CR01-02-03-04, but is limited to participants that should be considered “disadvantaged”, as per the list to the right.	- participants who live in jobless households* - participants who live in a single adult household with dependent children* - migrants, people with a foreign background, minorities (including marginalised communities such as the Roma)** - participants with disabilities** - other disadvantaged**

Table 3 – ESF common longer-term result indicators: detailed description

Indicator		Description	Reference population	Mandatory reporting
Code	Name			
CR06	participants in employment, including self-employment, six months after leaving	change in the employment status, including self-employment, 6 months after leaving compared to the situation when entering the ESF operation (with the participant being unemployed or inactive when entering the ESF operation)..	- unemployed - inactive participants	2019 and 2025
CR07	participants with an improved labour market	Persons who are employed when entering ESF support and who, following the support, transited from precarious to stable employment,	- employed participants	2019 and 2025

Indicator		Description	Reference population	Mandatory reporting
Code	Name			
	situation six months after leaving	and/or from underemployment to full employment, and/or have moved to a job requiring higher competences/skills/qualifications, entailing more responsibilities, and/or received a promotion 6 months after leaving the ESF operation.		
CR08	participants above 54 years of age in employment, including self-employment, six months after leaving	Same as CR06, but limited to those above 54 years of age	- unemployed - inactive participants above 54 years of age	2019 and 2025
CR09	disadvantaged participants in employment, including self-employment, six months after leaving	Same as CR06, but limited to those who have any of the disadvantages listed in the cell to the right.	- unemployed - inactive participants with any of the following disadvantage(s): in jobless households; in a single adult household with dep. children; migrants, with a foreign background, minorities - p. with disabilities** - other disadvantaged**	2019 and 2025

So, if we look at result indicators, the following chart clarifies what results are mapped for given categories of participants with respect to their labour market position.

Table 4 – Results mapped by the ESF monitoring system by labour market situation of participants

LM situation of participants		Immediate result indicators				Longer-term result indicators	
		Eng. In job seeking (CR01)	In E&T (CR02)	Gaining a qualification (CR03)	In employment (CR04)	In employment 6 months (CR06)	Improved LM condition 6 months (CR07)
Inactive	NEETs						
	In E&T						
Unemployed							
Employed							

Source: author's elaboration based on the ESF regulation and related monitoring guidance

As it can be seen, the standard ILO/Eurostat taxonomy is applied with respect to the labour market situation of participants, who are divided into 3 main categories (inactive, unemployed and employed) with inactive being further divided into those that are in education and training and those who are NEET (not in employment, education or training).

With respect to Immediate result indicators, all results are mapped for NEETs, which could be broadly seen in macroeconomic terms as shocks such as:

- entering the labour force (CR01);
- increasing their productivity through upskilling (CR02-03); and
- increasing employment rates (CR04).

A much similar reasoning applies to “Inactive who are in education and training”, with the only difference that an increase in their productivity is measured only where a qualification has been gained, which is likely more meaningful in macroeconomic terms.

In as far as the unemployed are concerned, there are two possible results measured:

- increasing their productivity through upskilling (CR03); and
- increasing employment rates (CR04).

Finally, as regards the employed, the only measurable result for them is linked to increases in productivity following gaining a qualification (CR03).

When it comes to longer-term results, only two results are measured, and namely:

- for the inactive and unemployed: increasing employment rates; and
- for the employed: improved labour market condition.

It is already apparent that there is no clear definition of the educational results. Especially when it comes to the result indicator “gaining a qualification”, the only known information is that the title obtained needs to be considered a qualification according to a competent body. But this leaves the door open to a very great deal of variation among MS and OPs.

In addition, the way result indicators are measured means that it is impossible to infer conditional distributions of results by some background features of the participants, including their labour market status (people moving into employment could be either inactive or unemployed at the entry to the operation) and educational attainment. As discussed further below, this has implications has to the suitability of the monitoring system to support evaluation of the fund.

2.2.2.4 YEI common result indicators

The monitoring system of the YEI does not depart from that of the ESF except for some dedicated indicators which aim to measure the “result” of the participation to the programme. Much like the ESF common result indicators, these are divided into immediate and longer term. As to the former, the following conditions upon exit from the operation are tracked:

- having completed the YEI supported intervention;
- having received an offer of employment, continued education, apprenticeship or traineeship upon leaving; and

- being in education/training, gaining a qualification or in employment, including self-employment, upon leaving.

These results are somewhat incremental, in the sense that it is first measured whether the support was completed, then whether it led to an offer of continued education/ work and finally whether employment was gained or the participant actually continued in education/ gained a qualification. Yet, it might be that participants find employment before completing the YEI intervention, or that become self-employed without, clearly, having received an offer for that. So they are typically treated separately.

The key difference with the ESF common result indicators is that in the case of the YEI the labour market status of the participant at the entry of the operation is known.

In addition to immediate results, also longer-term ones are measured (based on sampling approaches similarly to the case of ESF longer term result indicators). The longer term YEI result indicators measure the number of YEI participants:

- in continued education, training programmes leading to a qualification, an apprenticeship or a traineeship six months after leaving;
- in employment, in employment, including self-employment, six months after leaving; and
- in self-employment, six months after leaving.

In fact, the indicator “in employment, including self-employment” is the same as the ESF longer term, which creates uncertainties to the Managing Authorities. In addition, the labour market status of the participants at the entry of the operation is no longer traceable for longer term result indicators.

2.2.2.5 The Performance Framework

The performance framework constitutes a marked change of the 2014-2020 programming period in comparison to the former ones. The performance framework oversees all the Managing Authorities so that they set milestones values that selected indicators should achieve according to specific rules at the level of priority axes and category of regions.²⁰ The milestones are set for both financial and physical indicators. The milestones can be considered as mid-term goals in view of the ultimate targets set for 2023. The milestone check is set for 2018, and at that moment every programme is scrutinized to check whether the indicators are in line with what foreseen at the beginning of the programming period.

Based on the outcome of the milestones, European Institutions may engage with Managing Authorities to provide further support and, in case, alter the financial allocation for the following years. In particular, failure to achieve the performance framework may:

- if there is a severe failure (indicators at below 65% of their milestone values together with other conditions) a procedure which may lead to the interruption of payments is initiated had the MAs not taken proper remedy action; and

²⁰ In general, the Performance Framework is applied at the level of PAX. However, for the so-called “complex priorities”, that is, priorities which cover more than one category of region, specific target values should be selected for each category of region, and the rules determining the access to the performance reserve are applied at the category of region level (within each PAX).

- if there the target is not reached but values are above 65%, then approx. 6% of the allocation, the so-called “performance reserve”, needs to be shifted to better performing priorities through a revision of the OP.

Just under 30% of the output indicators have been included in the performance framework. From the financial point of view, Member States will on average have spent one fourth of their overall budget by then. As of October 2019 the performance review is currently being finalised.

As discussed further below, some requirements on the target setting for the performance framework have meant a proliferation of programme-specific indicators, which have however proved to be of little use when it comes to the comparative analysis of the ESF. This has happened especially as MAs were required to identify output indicators to be selected for the performance framework in relation to at least 50% of the operations funded within the PAX. But as operations may have very specific targets, this has meant that in order to meaningfully measure their progress, MAs have developed a great variety of programme specific indicators.

2.2.2.6 Availability of data breakdowns

As the table that follows shows, simply identifying what data is collected may tell quite little as to how detailed that information is and, in turn, which purposes it can actually serve. This is particularly true when it comes to understanding the impact on different target groups or on some specific regions. It is also quite problematic to disentangle the contextual factors that might have played a role in the implementation, as these are often intervention, target group or region specific and often data needs to be assigned to a given region rather arbitrarily or based on an educated guess. In the context of an evaluation, there is typically a need to understand what works better, where and for whom. This was the case, for instance, of the mid-term evaluation of the ESF thematic objectives, which are being carried out as of October 2019 and to which the author is contributing in two cases (youth employment and employment and mobility). And in order to do so, not only data on the programme under examination, but also information on environmental factors which may affect implementation is necessary. But the monitoring system of a programme is arguably the key source of information for data on its implementation, whereas the analysis of contextual features is typically left to evaluators. Below, an overview table is provided that summarises the availability of monitoring data within the frame of the ESF/YEI, from the broad “Member State” level, to the very detailed “information by typology of intervention AND target group, in NUTS2 regions”. It is worth mentioning that, for instance, information on the effects of active labour market policies is available in the literature at the level of the typology of intervention, by target group (see for instance Card and Kluge, 2018). Similarly, information for the macro-equilibrium model RHOMOLO employed in the ex-post evaluation of the ESF and in the impact assessment of Human Capital Funds, needs to be broken down at the NUTS2 level, ideally by target group and typology of intervention. It should be recalled, as clarified in 2.2.2.3, that very detailed micro-data is collected within the framework of the ESF, but not reported by MAs to other stakeholders, if not through ad-hoc requests and not in all MS. This means that micro data de facto not available to evaluators having to address the evaluation questions – including what works, where and for whom.

So the key issue here would be that of the availability of micro-data. One should consider that micro-data is collected for 187 programmes, by hundreds of thousands of beneficiaries all over Europe, each of which using dedicated exchange systems which have been one of the key issues of implementation according to the MAs in recent years (see for example FGB 2017, 2018). The volume of data is indeed quite an obstacle, if one considers that only in Italy by the end of 2018 approx. 5 million people have participated in ESF/YEI supported activities, across 29 operational programmes and potentially hundreds of thousands operations. In the context of the pilot study on the feasibility of Counterfactual Impact Evaluation under the ESF (Ismeri et. Al, 2019), the issue was explored in detail by the consortium in charge of the study and it was found that not all Managing Authorities are in a position to share micro-data with third parties or the Commission, due to either technical difficulties or privacy issues. For example, no micro-data is available for Spain (Ismeri et. Al, 2019). During an evaluation in Italy (Evaluation of effectiveness of the National Operational Programme on Legality), the same happened to FGB. Similar issues are being faced by FGB and our partners in the frame of the studies supporting the mid-term evaluations of the youth employment (Metis et. Al, ongoing) and of employment and mobility (FGB et. Al., ongoing). These issues might be solved if there's a focused effort from the Commission at consolidating all micro-data DBs across the EU-28. Such an attempt should not only consider the problem of bringing together an extremely large volume of data from fragmented data sources, but also the issue of its classification. In fact, there is not common definition of what a typology of intervention is, so to group together operations is not a trivial task.

Table 5 – availability of data from the ESF/YEI monitoring system, by typology of indicator and level of disaggregation

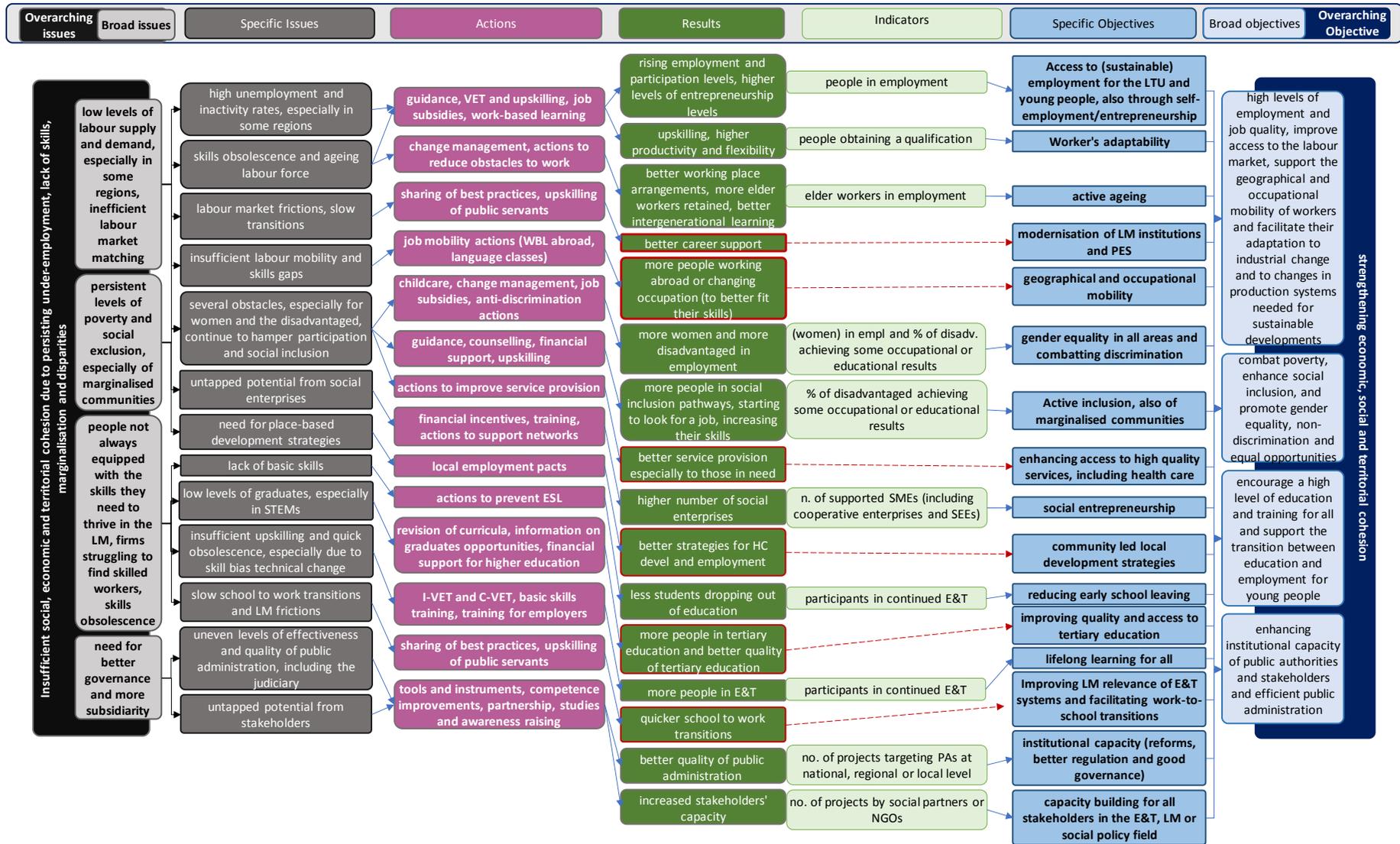
Level of aggregation	Inputs (allocation, expenditure)		Outputs (participants)		results (educational and occupational)		effects	
	EU-28 (SFC 2014)	Ad-hoc (micro-data)	EU-28 (SFC 2014)	Ad-hoc (micro-data)	EU-28 (SFC 2014)	Ad-hoc (micro-data)	EU-28 (SFC 2014)	Ad-hoc (micro-data)
MS	Yes		Yes (although it is the number of individual participations to a given operation that is actually measured)		Yes but only Common result indicators can be aggregated. Longer-term result indicators available only in AIR 2019 and 2025		Currently only through RHOMOLO simulations	In some cases depends on the availability of CIEs, the possibility to aggregate them and the level of breakdown at which these report their findings
OP	Yes		Yes (as above)		Yes but same as above			
NUTS2	In some cases, as not all operations are reported at the NUTS2 level		No (can only be assigned pro quota based on costs, which are also not always available)		Yes but same as above + they can only be assigned pro quota based on costs			
Investment priority	Yes		Yes		Yes but same as MS and OP level			
Target group	No	In principle yes, but the fact not available to evaluators	No, only marginal distributions are known, which do not suffice	In principle yes, but the fact not available to evaluators	No	In principle yes, but the fact not available to evaluators	No (only through dedicated evaluations, not yet produced)	
Intervention	No, except if available in qualitative sections of AIRs or when an IP has a single intervention		No, except if available in qualitative sections of AIRs or an IP has a single intervention.		No, except if available in qualitative sections of AIRs			
Intervention and target group	No, except if available in qualitative sections of AIRs or when an IP has a single intervention and single target group		No, except if available in qualitative sections of AIRs or if an IP has a single intervention and target group		No, only where available in qualitative sections of AIRs or an IP has a single intervention and target group			

Source: author's elaboration

As shown by the Table above, data on input, output and short term results, is mostly available in an aggregated form which typically means complications when it comes to the analysis of the Fund's achievements.

In the next section, the intervention logic of the fund is reconstructed so that it will be possible to appraise also which of the expected fund's results are covered given the information presented above. As mentioned in chapter 1, the intervention logic is the starting point for an assessment of the monitoring system, as it provides the backdrop against which monitoring arrangements should be evaluated in terms of the system being comprehensive, proportionate, relevant, cost-effective and so forth.

2.2.3 Intervention logic of the ESF and related monitoring arrangements



Source: author's elaboration based on the ESF regulation, monitoring guidance and own experience

As it is apparent from the ESF intervention logic, the extent and variety of the issues addressed by the ESF is vast. Yet, there's only a few indicators which can give some information on the activities carried out, to whom these are directed and with what measured result. Results highlighted in red in the figure are those which are left entirely unattended by the monitoring system, but this is not to say that for the rest indicators can respond to the information need related to the specific objective pursued. The need is particularly apparent in TO9 (social inclusion) with no intermediate or social inclusion outcome measured, just some indication of the share of those that are disadvantaged among the participants. But this applies to TO10 too, where information on early school leaving or access to tertiary education is missing. And, even in TO8, for which the currently collected employment-related indicators seem to provide a better match, some areas remain uncovered, such as, for instance, that of mobility. This has led the European Court of Auditors to say that, in fact, although it is known that TO8 supports mobility, the extent, content and results of this support are not known (ECA, 2018).

2.2.4 Focus: the ESF and the support it can provide to measure progress towards the goals of the New Skills Agenda for Europe (NSA)

Within the frame of this research project, an analysis was carried out that investigated how and whether the ESF/YEI monitoring system could support the monitoring of progress towards the aim of the New Skills Agenda for Europe (NSA). Its detailed development is included in Annex II for the sake of conciseness, but here a summary is presented as it serves to better understand also some of the limits of the ESF/YEI monitoring system in practice, thus feeding into the analysis.

In light of the importance placed upon human capital development in the EU2020 strategy and bearing in mind the increasing and widespread emphasis on monitoring and evaluation of public policies and expenditure, the main goal of the analysis has been to check to what extent the NSA, arguably the most significant and recent initiative for skills development in the EU, is currently monitored (or could be monitored in the future). The starting point of the analysis was that the NSA comes with no dedicated monitoring system attached, in that it is not endowed with a dedicated budget. However, the NSA does aim at steering funding from other EU instruments towards the achievement of its objectives. Thus, the analysis went on to check whether NSA's most relevant supporting funds, the ESF/YEI, may allow tracking such progress through its own monitoring system.

Analyzing the contents and main actions of the NSA recommendation, it was considered that not all NSA actions need a monitoring system at an individualized level, in that progress of more 'systemic' actions, such as the revision of the EQF or of the Europass Framework – to name but a few – is suitable to be observed on an ad-hoc basis. Thus, this analysis focused on two main rather individualized actions of the NSA, namely:

- 1- the Upskilling Pathways, which aims to provide individualized and modular basic skill training to adults who have up to ISCED 3 educational attainment
- 2- 'making VET a first choice', which aims to strengthen the quality and relevance of VET pathways.

Another two actions were considered partly relevant for the monitoring of individualized support, notably the ‘sharing of best practices on brain drain’ as well as the ‘skills profile tool for third country national’. However, these two are much more limited in scope and could be considered only to some extent, as they will also rely on other EU funding (and not just the ESF/YEI).

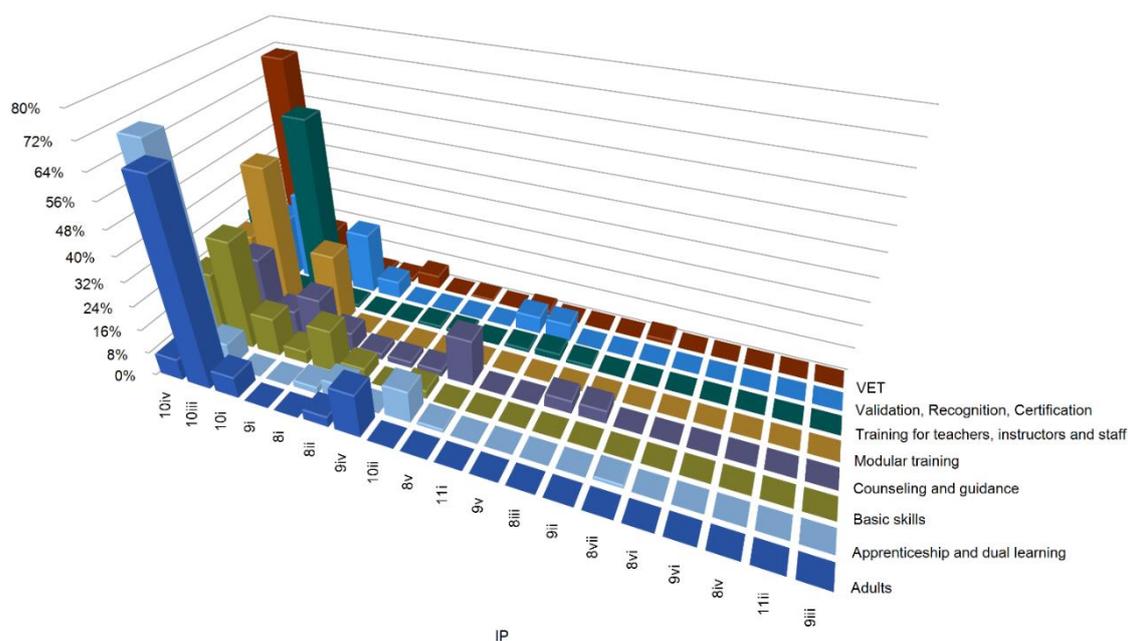
The ensuing analysis on ESF Common output indicators drew attention to their very broad formulation, which fails to produce figures on ESF interventions at a level of breakdown that can be reliably deemed relevant for the NSA and aligned to its priorities.

Only one Common Output Indicator was found to be relevant, to some degree, to give an indication on how interventions targeted at the low-skilled are unfolding. This is notably the indicator on the educational attainment of participants. However, it can only provide a broad idea of how interventions relevant for the Upskilling Pathways, the flagship initiative of the NSA, are progressing towards their targets, because, for instance, it does not exclude isced4 participants (which are not eligible for the Upskilling Pathways) nor it can tell how many of these have: a) received basic skills support (let alone modular and individualized support) b) were adults, which are the key distinguishing features of the Upskilling Pathways.

Because, as anticipated in section 2.4.2.1, MS have made vast use of Programme-specific indicators, the analysis then sought to ascertain whether their more specific formulation could provide better support for monitoring relevant NSA actions. Although it is acknowledged that programme specific indicators can hardly be aggregated on a vast scale, more focused approaches could however benefit from that information (including looking at target achievements).

The resulting textual analysis, carried out through NVivo on the phrasing of the over 3000 programme specific indicators through a number of keywords and lemmas, clarified that some scope exists for tracking progress towards the aims of the NSA action on VET, as shown by the Figure below.

Figure 6 – percentage of relevant programme-specific indicators by theme (node) and IPs



Source: author's elaboration through NVIVO based on the text of the programme specific indicators as reported in FGB et al, 2016

As it can be seen, there is a sizable share of indicators which have been categorised, according to textual analysis, under themes (nodes, as they are called in NVivo) that are relevant to the purpose of 'making VET a first choice' as well as the Upskilling Pathways. These include VET, validation recognition and certification of skills, modular training, etc. Such an high level analysis already showed that nothing can be learned for the remaining two NSA actions considered relevant for a monitoring at individualised level.

Nevertheless, the analysis pointed quite clearly to the lack of appropriate coverage of the ESF monitoring system based on Programme-specific indicators, first of all because some countries make a much wider use of these and their formulation seems quite aligned to the priorities of the NSA, but for many others that is not the case.

This is particularly true for the 'Upskilling Pathways', whose scope is rather articulated. When checking for programme-specific indicators taking into account all the dimensions of the UP (age, level of skills, type of support) only a few indicators were found.

The picture is relatively more favourable when it comes to VET interventions, probably due to the fact that 'make VET a first choice' is a slightly less innovative measure than the Upskilling Pathways, and as such its main aims were already reflected through the ESF 2014-2020. In any event, programme-specific indicators could not serve the purpose of providing an overall picture of how the NSA is unfolding, given their acknowledged limitations when it comes to aggregability.

On a side but relevant note, the analysis shed also light on the strong interlinkages between VET and the UP, which is an element to bear in mind also when considering the monitoring for these two actions, as many of the activities they encourage are mutually beneficial to the aims of the two initiatives.

All in all, the analysis provided some evidence that, despite the themes of the NSA being matched, at large, by the monitoring system of the ESF/YEI, no organic monitoring system that can track progress towards the objectives of the NSA exists. Thus, some efforts should be fostered at MS level to set up appropriate frameworks to collect relevant data.

The analysis particularly highlighted how pivotal would be to have micro-data that can disaggregate the volume of activity supported, and related results, by typology of intervention and target group. Only there could progress of the main actions of the NSA be measured through the ESF/YEI. But obtaining micro-data of the EU is, as already discussed, not a viable option at the moment.

2.2.5 Main issues with the ESF/ YEI monitoring system

Over recent years, EC's officials and stakeholders have been engaged, both directly or through the work of contractors, in a reflection on the result-orientation of the ESF/YEI monitoring system. The emphasis on these aspects was already apparent at the end of the 2007-2013 programming period and arguably significant efforts have been carried out to make the system more result-oriented. Amongst others, a list of common indicators whose collection is mandatory has been defined, and the number of participations now mirrors more closely that of the participants to the fund.²¹ It is also worth mentioning that the presence of a compliant monitoring systems was made an Ex-ante conditionality within the framework of the 2014-2020 programming period. However, several aspect continue to appear as problematic. Below the main issues that have emerged either in the literature or within the studies in which the author took part are reported.

- **lack of data on typology of interventions supported and target groups addressed:** as anticipated, the ESF/YEI monitoring systems requires Managing Authorities to report on the financial progress of their programmes, according to a certain breakdown (i.e., by PAX, IP, Category of Region and NUTS level). However:
 - **there is no exact indication on the typologies of intervention supported:** the description of “what is being done” is left to the qualitative sections of the Annual Implementation reports whose minimum contents, however, are not defined by the regulation. So one could somehow imagine that in order to pursue a given investment priority, a certain broad typology of operation was carried out. But in practice activities may vary to a large extent (e.g., within the frame of TO8, from hiring subsidies, to training, guidance, support to employment mobility, work-life balance arrangements, child care support and the likes). In principle, one could look at project/operation lists to check summaries of all operations supported. However, there is no clear definition of what an operation is,²² thus operation could be also very disaggregated projects and can happen to be several thousands of them per OP, which make their analysis in

²¹ In the 2007-2013 phase, in many instances the indirect beneficiaries of the ESF were considered participants, see for example the case of a school whose teachers are trained. In the previous programming period, the entire population of that school was considered as participants to ESF support.

²² According to the Common Provision Regulation n 1303/2014, art. 9(2) operation “*means a project, contract, action or group of projects selected by the managing authorities of the programmes concerned, or under their responsibility, that contributes to the objectives of a priority or priorities ...*”

practice very problematic. Thus, proxies need to be used instead (see for example the Mid Term evaluations of TO8 investments and of YE investments carried out by Metis et. Al (ongoing) and FGB et. Al (ongoing) for DG Employment. Their reliability and usability is however rather low, for instance for quantitative approaches;

- **there is no exact indication on the target groups addressed:** even where the AIRs tell something about investments being supported through the OP, quality detail on the target groups address is often missing. Just like in the above case, one could go and look at project summaries to get some insight into which groups were targeted but: (i) it is not mandatory that in the project summaries a clear indication of target groups is provided, nor that their weight is indicated vis-à-vis the whole range of those potentially mentioned for that same project (ii) it is unrealistic within the design of current EU-28 studies to go through potentially hundreds of thousands project descriptions to categorise them;
- **participations vs. participants:** according to the monitoring guidance, despite the attempt to shift to a system in which each participant was recorded only once, this is only the case of participants entering and leaving the same operations several times. If a participant is then supported through a different operation, it has to be counted again. So aggregated values do not directly reflect the number of individual participants being supported by the ESF/YEI.
- **complex monitoring requirements for Managing Authorities and Beneficiaries and lack of adequate resources to deal with these:** the complexity of setting up procedures for data collection, validation and storing for both MAs and beneficiaries was confirmed by two thirds of the participants to the survey carried out by Applica & al. (2018). This finding is confirmed also by the screening of AIRs done within the framework of the studies carried out by FGB (2017, 2018) in the context of the performance of the ESF/YEI. Underreporting or just unreliable values are flagged up by MAs themselves due to issue with the monitoring system and in some countries manual verification of the monitoring sheets was necessary also given the time lag between the commencement of the programmes and the time at which M&E were actually up & running. Difficulties were also encountered in the collection of sensitive data from disadvantaged participants and because of troubles applying consistently the definitions. Often times, MAs had to hire new resources to deal with the setting up or maintenance of the monitoring system, especially with IT issues encountered in several MS (FGB 2017, 2018);
- **lack of relevance:** this is particularly apparent if one thinks of the intervention logic of the ESF/YEI. For some thematic objectives and Investment Priorities more than others, there are simply no results which can be meaningfully linked to the objectives. Take for example TO09 (social inclusion), TO11 (institutional capacity) but also TO10 (education and lifelong learning). There are no soft results on intermediate outcomes (i.e., before entering the labour market, such as increased skills, self-confidence, information) no results which can really match some specific investment priorities (e.g. IP 8.vi “active ageing”, nothing on improved work-life balance nor flexible working arrangements). Also, no results on administrative capacity, higher level education, and so forth. This has also meant in many cases a lack of data for the evaluations to be carried out at the national level. In the survey carried out by Applica & Al. (2018) nearly two thirds of the MAs have highlighted this

issue. The same study suggests that in the evaluation plans, there was a general tendency of not identifying precise data sources which could be used to actually answer the evaluation questions foreseen. This is arguably the result of the more general issue of data collection;

- **issues with the target setting:** the 14-20 monitoring has seen the introduction of the performance framework and of the need for MAs to identify at least one indicator per specific objective which had a baseline and a target value. For a selection of them (those selected for the performance framework) also a milestone value had to be identified. Nevertheless, the requirements for the performance framework (the indicator selected should cover operations which represent at least 50% of a given PAX and category of region – DG EMPL, 2018b)²³ meant that in many cases MAs had to resort to programme-specific indicators instead of common ones. The lack of benchmarks has meant that in several cases targets identified used wrong calculation methods or assumptions (FGB 2017, 2018; Applica & al 2018)
- **data granularity (breakdowns and disaggregation)** although micro-data is mandatorily collected at the participant level, this not readily available in most countries and it cannot be made available in others, according to Ismeri et. al (2019). This means that background information on participants and their results are not known, as one can only appraise how many, say, low skilled received support in a programme but not how many of these were young, male, achieved some result etc. This also means it is not possible to know with precision;
- **data for a reasoning on unit costs (per output and result) and success rates:** an issue which appeared as particularly relevant within the context of the Mid-Term Evaluations of the ESF/YEI is that of a lack of comparable data on unit costs and success rates. This is due to a range of issues, many of which have to do with the monitoring system. These are listed below:
 - **time lag between the reporting costs, output and results:** in general, unit costs would be an interesting measure of ongoing performance of a programme. However, it is very hard to calculate them also because of the process of ESF/YEI implementation and the time points at which info is collected. Typically the process goes along the following lines:
 - an operation is selected, and its cost registered;
 - the operation starts and participants are registered at the entry;
 - the operation ends and immediate results are tracked;
 - the eligible expenditure related to the operation is declared by the beneficiaries to the MAs and there is a certification process; and
 - information is collated by MAs and reported in AIRs.

This is further exacerbated by the fact that MAs can report alternatively on partially implemented or fully implemented operations. Despite the fact that the monitoring guidance requires MAs to state which of the two options they are using in the overview section of Annual Implementation Reports, evidence shows that this is seldom the case (FGB 2017, 2018). Thus, comparison of unit costs and success rates becomes

²³ "In order to capture the essential information on the progress of a priority, the selected output indicators have to relate to operations representing the majority of resources allocated to a priority".

particularly difficult, as a low level of reporting on results might well be due to the time lag. Also, it is difficult to understand which costs should be used as a reference (e.g. does it have to be cost of selected operations or eligible expenditure?). This issue is dealt with in the simulation exercise carried out in chapter 3, including through econometric techniques;

- **participations/participants:** as per the explanation above, it may be that one participant entering several times the same operation has a result only at the end of the latter operation. His success rate would be, however, considered as 100% despite the many tries. On the other hand, when calculating costs, 4 participations will be considered just one. So in interventions with many participants who re-enter the operation the unit cost would look a lot higher than they actually are; and
- **information on target groups and typologies of interventions:** as anticipated, the information on the typologies of interventions and target group is very limited and can only be reconstructed through proxies. Different target groups and typologies of interventions, however, come with significantly different costs attached. So any comparison of efficiency needs to take this into account. Data on output can say something on target groups, including through econometric estimation (see chapter 3). Nevertheless, the lack of reporting to the EC of conditional distributions, i.e., not just the number of low skilled but how many of these are inactive or young, means that only rough estimations can be retrieved;
- **issues on data validation/drawing up of representative samples:** the requirements on longer term result indicators as well as output indicators is that data should not be collected at the individual participant level but in a representative way. However, this mean that the specific procedures need to be established and the process be transparent. This has cause difficulties for managing authorities. Also ensuring the statistical validity of the figures collected at the beneficiary level was considered as difficult (Applica et Al, 2018).

2.2.6 How does the ESF/YEI monitoring system fare vis-à-vis the Monitoring of HC checklists

After having identified some of the key issues that were reported in the literature and/ or have emerged within the studies carried out with FGB on the ESF monitoring and evaluation system, it remains to assess how the ESF/YEI system scores when it comes to the checklist developed based on the work done in Chapter 1. This shift in focus should allow some new angles to be brought to the fore, especially when it comes to the broader issue of human capital development and its measurement. The table below presents the findings of this case study with respect to the questions included in the checklist for the monitoring of human capital development initiatives.

Table 6 - assessment of the ESF/YEI monitoring system vis-à-vis the monitoring checklist

Criteria	Specific question	Evaluation
Comprehensiveness	Does the monitoring system cover all objectives of the programme – as defined in its intervention logic – and especially the operational and specific ones with new arrangements?	To a limited extent. It only slightly does so, but the situation differs according to which thematic objective is being considered as well as the level at which information needs to be collected. More specifically: if one considers the aggregated monitoring (AIRs and SFC) labour market transitions are monitored in all cases, but there are no indications on soft outcomes or these can be retrieved only through programme specific indicators, which however can hardly be aggregated (see FGB 2017, 2018). Labour market transitions are relevant especially for the aims of TO8, but these cannot really suit the specificity of TO9, TO10 and TO11. Programme specific indicators cannot be aggregated, thus they can hardly say something at the EU level. It is particularly apparent that the questions typically formulated by the EC on the effectiveness of their funds can hardly be answered with the data gathered by the ESF/YEI monitoring system.
	Does it do this through surveys as well as continuous monitoring?	To some extent. There are no standardised survey procedures set out in the Regulations, but these are foreseen as standard practice in the frame of studies which are commissioned by DG EMPL and carried out by external contractors.
	Do the indicators measure progress that can be linked to the different impacts on the different target groups, both direct and indirect?	No , this is left to the evaluations, which need to resort to other indicators. But controlling for measures of indirect impacts in such an heterogeneous fund would overcomplicate the system. However, there is very little data on results by target group, if not in micro-data which is hard to access.
	Are baseline and target values identified to track progress?	Yes , although the methodology for selecting baselines and targets has caused issues over recent years.
	Are transversal objectives such as gender equality, non-discrimination, inclusion of people with disabilities included in the monitoring system?	To a good extent. Data is broken down by gender in all cases and originally different forms of disadvantages were tracked. Nevertheless, issues with these indicators have meant that the requirements on collecting data on the disadvantaged have been relaxed by the Omnibus regulation. In the AIRs there are in addition qualitative sections which shed light on the advances of the OPs with respect to several transversal objectives (including integrated territorial approaches, rural areas, non-discrimination social innovation, gender mainstreaming etc.)
	Does it look at factors contributing to or impeding achievement of the outcomes?	To a good extent. Although not directly presenting a categorisation of issues affecting performance (or measure them) the AIR template is very much focused on issues affecting performance. These are then screened and categories within the frame of the ESF performance synthesis studies and used in the MTE

Criteria	Specific question	Evaluation
		too. Yet some factors are difficult to be used in evaluation as the monitoring is not at the regional level (where context may affect performance etc.).
	Is monitoring data suitable to be used for assessing the impact of the funds through counterfactual or other econometric approaches?	Only partly. In principle, micro-data can be used for counterfactual impact evaluation. However the fact the system is not very comprehensive in terms of results measured and especially not very relevant to many of the objectives of the supported interventions means that many evaluation studies have faced severe obstacles in getting the data they needed. Evaluation plans only seldom include detailed sources of data for the evaluations, and these cannot be retrieved ex-post, especially when it comes to information on the information on the groups of non-treated (counterfactual group). Recommendations have been drafted within the studies that MAs should facilitate access of contractors to administrative registries and actually assess ex-ante (i.e. in their evaluation plans) whether the right data is available for the evaluation questions they draft.
Participation	Is feedback from partner and beneficiaries foreseen in the monitoring? If so, how? (e.g. steering committees etc)	To a good extent. Monitoring committees are established for each programme that involve sufficiently broad partnerships. The AIRs are approved by the monitoring committees. There are a range of inter-institutional steering committees also for the studies which put together the evidence from AIRs, or carry out evaluations on the ESF/ YEI.
	Are experts and other stakeholders involved in the definition of monitoring arrangements?	To some extent. Specific studies are foreseen to this end and meetings with stakeholders in the frame of different studies or research activities. However the most comprehensive study done on the M&E system cannot answer many of the questions here described, so perhaps it wasn't very comprehensive. In particular, it doesn't a clear benchmark/ evaluation grid against which to assess the M&E arrangements of the ESF/ YEI.
Proportionate and no overlaps	Does it treat the monitoring in a cost-effective manner (including electronic systems for data collection and exchange)?	Partly. This issue is hugely debated. Evidence from MAs and beneficiaries points to the fact that the requirements are particularly complex, including measures which were originally aimed at making monitoring easy (e.g. sampling approaches for longer term results). At the same time, some monitoring arrangements (e.g. programme-specific indicators) are of little use for monitoring, but, as explained earlier, their widespread use is due especially to the performance framework requirements, as well as to the fact that common indicators cannot capture the intended outcomes of much of the support carried out under TO9, TO10 and TO11.
	Does it maintain indicators constant over time unless new major needs emerge?	Yes, disproportionately. Great heed is paid to keeping monitoring costs and burdens at bay as much as possible, including doing as little changes as possible to existing monitoring requirements, which have changed from 2007-2013 but are expected to remain rather stable from 2014-2020 to 2021-2027 (see

Criteria	Specific question	Evaluation
		next section on the future of the ESF). This may be counterproductive in some cases. The feedback gathered from stakeholders however points to the need of very stable monitoring arrangements.
	Does it only prescribe new monitoring arrangements where needed and at a reasonable cost?	To a good extent. The use of administrative registers is strongly encouraged and the array of result indicator is extremely limited. One could argue that there are too few new monitoring arrangements.
Timeliness	Does it collect evidence in a timely fashion, i.e., not necessarily from the outset but only and every time this is needed?	To some extent. Although data is collected yearly and should be suitable for the analysis, issues with the time lag and partially vs fully implemented operations mean that data for the analysis of efficiency is still not very sound 5 years into the implementation of the programme and notably for the mid-term evaluations.
	Are milestones foreseen?	Yes , in the performance framework and there is also the so called N+3 rule, which entails that MAs have to certify expenditure of the budget allocated yearly within 3 years of its allocation. However, whilst the performance framework has helped MAs focus on the progress of their key interventions, it has also meant the proliferation of programme-specific indicators due to the requirements of the PF and the fact that the list of common indicators is not comprehensive.
Accessibility	Is data collected accessible for the general public?	To a good extent. There are datasets made available in the open cohesion portal and in the open data portal of the EU. In addition, Annual Implementation Reports are available on the MAs web pages, together with the list of operations selected and
Relevance	Are only the relevant indicators included in the monitoring list or there's just too many of them?	Too few indicators. The number of indicators is kept low especially for administrative burden purposes. But the fact that there's too few of them means that some are just not relevant for the interventions under M&E.
	Are indicators updated and changed where necessary (and only then)?	Less than necessary. The overriding concern on keeping the indicator system stable (which is fully legitimate considering the costs) has pushed DG EMPL to formulate changes for the next programming periods which do not address many of the issues encountered but only aim at keeping things constant to avoid administrative burden.
	Are indicators set at a level where decisions are taken either on the policy or on the programmes?	To a good extent. Although data is collected at the OP level (and even below) so that each managing authority is responsible for that, some issues with, e.g., the NUTS2 breakdown of the data mean that it is difficult to disentangle contextual factors from implementation issues. This, in turn, results in lower accountability, because it is very hard to know whether the context (and some decisions taken at the regional level) have affected the implementation of the programmes (see section 2.4.7).

Criteria	Specific question	Evaluation
Suitability to measure human capital development	Does it take into account the main dimensions of human capital (capacity, deployment, development, know-how and creativity, non-cognitive skills , health , culture and tolerance) as relevant to the objectives of the initiative?	To a very small extent. The approach is mostly based on the deployment element (labour market transitions and moving unemployed or inactive into/ closer to employment). There is very little in terms of educational attainment (there is an indicator on “gaining a qualification” but this is not linked to any ISCED level) and very little also on the quality of job and education. For the quality of education in particular, test of adult competences (e.g. pre-post PIAAC longitudinal tests) could be foreseen. Health, culture and tolerance related indicators are not used, although there is some emphasis on inclusion of minorities and disabled, which however have caused difficulties. Overall, much is left to specific evaluation exercises, which however risk having little data to underpin their findings.
	What is the underlying approach to measuring human capital, i.e., monetary or non-monetary?	The underlying approach is non-monetary and based on a small array of output indicators, with the limitations described above. Monetary measures are not considered, although financial data is collected and thus one could approach the evaluation also from the perspective of the cost of support to education. This is actually the approach that was used, although in a rather specific form, in the context of the Impact Assessment through the RHOMOLO model. More details on this in Chapter 3, as this is its focus.
	Does it take into account relevant direct (first order) impacts as well as indirect (second order) effects?	No. Monitoring focuses on individuals and there are no results collected on institutional changes etc, except in programme specific indicators. So it only looks at what individuals can achieve upon exiting the operations. Second order (indirect) effects could be assessed through the general equilibrium macro-economic model RHOMOLO. However, monitoring data is not directly suitable to serve as an input for the RHOMOLO simulation, which means that currently the methodology needs to make strong and unrealistic assumptions to calculate the impacts of the ESF. This is the focus of the third chapter of this research project.
	Do indicators measure input, process, output or outcome-related factors of human capital development?	To some extent. In general, inputs are only financial in the ESF / YEI whereas in the framework offered by Borowski (2013) for the measurement of human capital the number of say, teacher trained would be an input. This is however considered as an output within the frame of the ESF. Output indicators in the framework developed by Borowski correspond partly to result indicators in the ESF /YEI M&E system (people in ET or gaining a qualification) and yet there is no indication on the level attained. Finally, result indicators in the Borowski’s framework are partly covered by ESF result indicators but only when it comes to employment outcomes, no social inclusion. Increased earnings are not measured directly but included in one general longer-term indicator which looks at those already employed at the entry of the operation which generally states improved LM conditions (and increased earnings is among the list of situations which contribute to the indicator).

Criteria	Specific question	Evaluation
	Does it measure intermediate output/outcomes which can be linked to the intervention (enablers, soft outcomes)	To a very limited extent. Programme specific indicators tend to serve this purpose but they have different unit measures, combine different characteristics of the individuals addressed and the related interventions and therefore they have not so far proved suitable to measuring aggregated output/outcomes. They might be used however to check monitoring arrangements in some specific areas/programmes or for some specific purposes.

2.2.7 The future of the ESF monitoring system

2.2.7.1 A new landscape for HC development funds

Over recent months, much work has been carried out to prepare the regulations for the next programming period, to start in 2021.

The Impact Assessments carried out by the DGs in charge of the funds, and namely DG EMPL for the ESF, also building upon supporting studies and mid-term evaluations on the current implementation of EU funds carried out by FGB and to which the author has contributed, have pointed particularly in the field of cohesion policy to:

- focusing more on EU added value
- streamlining the funding landscape
- reducing red tape and administrative costs
- increasing flexibility of the EU budget and related initiatives
- improving result-orientation

This has led inter alia to the proposal of some incremental but significant changes to the overall balance of the provisional post-2020 budget, including amongst other:

- a significant increase of the budget dedicated to Horizon actions, under the new Horizon Europe programme
- a doubling of the funding to Erasmus+ to boost learners' mobility
- a significant streamlining of the funds falling under the new heading "II. COHESION & VALUES" and sub-heading "7 Investing in People, Social Cohesion & Values", managed under the remit of DG EMPL, thus including the ESF, the YEI, the FEAD, the EaSI and the EGF

As to the latter, which is the most significant in the context of this study, the EC proposed to merge 4 funds under DG EMPL remit, namely the ESF, the YEI, the FEAD and the EaSI and two further instruments (Integration of Migrants and Health programme) within a single umbrella fund, the ESF+. For the coming 2021-2027 period, the Commission proposed to allocate 101.2 billion euros (€89.7 billion in 2018 constant prices) to this new fund, therefore representing a share of around 27% of the cohesion policy (and roughly 8% of the whole MFF). The 101.2 billion euros include 100 billion for the ESF+ under shared management and 1.2 billion euros under direct management.

In January 2019, however, the European Parliament's Employment Committee approved a decision to increase the ESF+ and make youth and children the main beneficiaries, which was then approved by plenary. This was validated on 4 April 2019 with the Parliament proposing to increase ESF+ funding by 19% from approximately €89,6 billion to approximately €106,8 billion in 2018 prices of which 105,7 billion euros to be managed jointly by the EU and Member States. As of October 2019, the matter is being discussed in interinstitutional negotiations. It is then yet to be seen if this increase will end up in the final text. A range of other changes have been suggested by the Parliament, including with respect to the monitoring arrangements, with new output indicators (on children, on Roma people) and especially new longer term result indicators (employment rates at 12 months of leaving the operations). Such changes are however being discussed and not yet enshrined into a regulation.

The ESF+ will comprise a mix of these shared management and direct management funds mainly with a view to:

- simplifying the funding landscape for beneficiaries thus cutting on administrative burden
- ensuring better upscaling of social innovation
- increasing flexibility of EU action and synergies between the funds
- increasing visibility of the EU action

Among the main concrete changes that this would entail one should consider, based on the current EC proposal for the Common provision regulation as well as for the ESF+ regulation:

- 5 policy priorities replacing the 11 thematic objectives, of which 1 for the ESF+, namely “A more social Europe - implementing the European Pillar of Social Rights” – which further stresses the link with the EPSR
- the provision of 11 priorities for the ESF, in contrast to the previous 19 across 4 thematic objectives
- a simplification of the monitoring system, together with the roll-over of existing structures to the extent possible and more widespread use of Simplified Cost Options
- better integration between different funds
- A closer link with the European Semester, including a mid-term review of the financial allocation

The current ESF+ proposal envisage earmarking for both youth employment and measures aimed at eradicating poverty and supporting social inclusion. It also foresees, according to the most recent parliament proposal, money earmarked to a Child Guarantee (currently being discussed).

2.2.7.2 The new ESF+ monitoring system

In particular, the study on the monitoring system of the ESF, the study supporting the impact assessment on Human capital investments and the Impact assessment on human capital investment have all dealt with the issue of result-orientation of the ESF. In addition to reflecting on the issues highlighted in the previous sections, they have proposed some indication on the way forwards, which has led – although with some changes – to the drafting of the proposal for ESF+ regulation. Amongst the main changes sit:

- differentiated monitoring arrangements also within the shared management funds, notably for the FEAD (this is subject of the other case study in section 2.5);
- the target setting will be at the programme level and not at the PAX/IPs/Category of region level. This a major reduction in target setting;
- No direct collection of sensitive data from participants, but just informed estimates or data through access to administrative registers, which should be simplified too;
- collection of data at the NUTS2 level, to ensure that macro-economic impacts can be suitably assessed; and
- Common results which are not relevant from the outset (e.g. finding employment for pupils) will no longer be collected.

In addition to these changes in the monitoring requirements, some other developments concern directly the existing common indicators. Amongst other, the number of common output

indicators would go from 19 to 15 for individuals and from 3 to 2 for entities, of common result indicators from 5 to 4, and of longer term results from 4 to 2, with generally lower emphasis on the elderly and disadvantaged participants.

So the changes to the monitoring system are minor and all intended to reduce administrative burden. The Impact Assessment document actually clarifies that “*only changes that reduce the burden of data collection and processing while increasing the relevance for management purposes are envisaged*” (COM, 2018b) ²⁴

Interestingly, the Regulatory Scrutiny Board had commented to DG EMPL that the Impact Assessment “does not sufficiently demonstrate that the monitoring and evaluation system is robust and will deliver the necessary information in a timely way” (COM, 2018b:9). This suggestions only led to a paragraph being written in the revised proposal, on the periodicity of report (6 times a year) and level of target setting (programme level).

So overall, whilst it can be agreed that the goal of reducing administrative burden is a relevant one, this does not address the majority of the issues highlighted above as well as the concern from the RSB on the robustness of the M&E framework.

To some extent, the revisions being proposed by the European Parliament in its 4 April 2019 amended proposed regulation²⁵ goes back to a slightly increased ambition of the monitoring system, including:

- additional common output indicators on NEETs, young people, participants from the Roma community, participants from areas with high levels of poverty and social exclusion, participants transitioning from institutional to family and community based care;
- additional seven common output indicators for ESF+ support to the social inclusion of the most deprived people (intended at the former FEAD type II Operational Programmes, as clarified in the next case study); and
- the timeline for the common longer-term result indicators (illustrating the situation of the participants both 6 and 12 months after the exit).

Whichever the orientation, it looks like some of the key questions which hamper the understanding of “what works and for whom”, notably disaggregated data on participants and typologies of intervention, will remain unsolved.

2.2.8 Conclusions and recommendations

In the light of the analysis presented above, and the direction taken by the European Commission in their proposal for the ESF+ in the post-2020 period as well as the related amendments by the European Parliament, the following recommendations are formulated:

- **to further discuss whether to include information on the typologies of interventions supported and their results.** This sort of data breakdown would be most suitable for understanding what works and especially for cost-effectiveness approaches as well as

²⁴ COM(2018) 382 final, Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the European Social Fund Plus (ESF+) {SEC(2018) 273 final} - {SWD(2018) 289 final}, available at https://ec.europa.eu/commission/sites/beta-political/files/budget-may2018-european-social-fund-plus-regulation_en.pdf

²⁵ European Parliament legislative resolution of 4 April 2019 on the proposal for a regulation of the European Parliament and of the Council on the European Social Fund Plus (ESF+) (COM(2018)0382 – C8-0232/2018 – 2018/0206(COD))

macroeconomic approaches, but obviously that comes at a cost; one easy way out would be to just ask managing authorities to tag the projects included in the project's lists to a pre-defined list of typology of intervention. Inspiration could be drawn from the current Labour Market Policies (LMP) database.²⁶ It should be mentioned that, for example, in the context of the evaluation of Youth Employment, clear issues appear as to the demarcation of the study. This is because, despite the fact that there is one dedicated investment priority to youth employment, MAs have often provided youth employment support also beyond that. The commission, aware of such practice, has requested the Contractors to look into other investment priorities and thematic objectives, but there is not data available on that, which hampers the analysis and obviously question the suitability of the monitoring system in place;

- **to discuss whether it would be possible to have more TO-specific or priority-specific indicators.** The overriding emphasis on labour market transitions risks overshadowing much of the benefits from ESF+ support. This is not just true of TO9, 10 and 11 but also for TO8, which supports “sustainable and employment and mobility” but, just to name one, there is no indicator on labour mobility foreseen. The best way to do this is to standardise and streamline the current programme specific indicators, and move to a list of indicators which is TO specific, but from which MAs can select only those that are relevant to their interventions;
- **to employ a more encompassing definition of human capital development,** and include results which are suitable for some specific objectives which do aim at direct employment. Amongst others, upskilling of workers in the creativity industries, improvement of health-related conditions of the most deprived, increases in non-cognitive skills. Also the lack of quality indicators for the level of skills accrued seems problematic if one wants to consider the effect of upskilling, as it was demonstrated how quality of education explains a significant part of growth that is due to human capital (see for example Hanushek and Woessmann (2016)). In this train of thoughts, also the issue of “key competences” which is *“the knowledge, skills and competences all individuals need, at any stage of their lives, for personal fulfilment and development, employment, social inclusion and active citizenship”* and include: literacy; multilingual; mathematics, science, technology and engineering; digital; personal, social and learning to learn; citizenship; entrepreneurship; cultural awareness and expression should be considered for monitoring purposes;
- **to have more disaggregated data available in SFC:** one particularly significant issue is that of micro-data and, more broadly, that of data disaggregation. In the Applica & Al. (2018) study on the ESF monitoring system it was suggested that the current requirements for micro-data were maintained as they can provide a vital source of information for evaluation activities whilst avoiding to make full datasets of micro-data available to the EC and other contractors if not ad-hoc, in order to mitigate any impacts on administrative burden. However, as clarified in section 2.4.2, there are strong reasons to make available slightly more disaggregated data than the ones available now in SFC, especially conditional distributions. This is particularly important for some specific target groups (e.g. the low skilled, young or old people) and generally for cost-effectiveness purposes or to understand

²⁶ See <https://ec.europa.eu/social/main.jsp?catId=1143&intPageId=3227&langId=en>

whether performance of the programmes is correlated to the features of the participants. The overall administrative burden should not increase substantially as this data is already collected by MAs, it only needs to be aggregated in a different form for reporting. The level at which evaluation operates is the intervention/ target group level, and this should be somehow taken into account in the design of the monitoring system. It would be also key to allow tracking progress towards some specific policy initiatives (see section 2.4.6) such as the all-important New Skills Agenda;

- **to reduce the number of programme specific indicators (if not to repeal them, as per the recommendation above)**, and have them standardised as much as possible into categories. Without the possibility to aggregate them and make comparison they are of little use for monitoring. The argument of programme specific indicators being necessary to identify given typologies of interventions and/or given target groups can be addressed by:
 - **having input data (on operations) structured by typology of intervention** (broad categories, it would be just about creating a list and having MAs categorise interventions according to given criteria. One should allow for some discretionary decisions, but would still get a much better understanding on the operations on the ground);
 - **using eligibility criteria to check the progress of a given target group** (there is no risk that different target group appear in the monitoring of a common indicators if only one target group is eligible, so no need to use a specific formulation of the indicator itself); and
- **to carry over the rest of the current monitoring system**, to keep costs proportionate, without though losing on accountability. Thus a more comprehensive system (e.g. replacing some programme specific indicators with common priority or thematic objective specific indicators or to report more disaggregated data through SFC should be considered).

2.3 Monitoring of the Fund for the Fund for European Aid to the Most Deprived (FEAD)

2.3.1 Main goals and structure of the fund

The FEAD is a the Fund for European Aid to the Most Deprived, funded in 2014-2020 within the Multi Annual Financial Framework with just about 3.5 billion euros of dedicated budget (thus approx. 4% of the ESF/YEI budget).

It is not an entirely new initiative as for its most part it closely follows the “Food distribution Programme for the Most Deprived Persons”, created in 1987 to tackle the issue of agricultural waste and surpluses which followed the early implementation of European Agricultural policy (CAP) and related quotas on agricultural products. The rationale was that food surpluses hence generated could serve as a means to alleviate the worst forms of food deprivation. But the more the CAP became a balanced policy with lower surpluses, the less the MDP could fulfil the persisting food need across the EU. Thus, a complementary fund was created that could back the EU cohesion policy in its quest to promote social inclusion in all countries and regions by removing obstacles to social inclusion (“enabling” function).

The FEAD is based on the article 175 (3) of the TFEU as a corollary of the EU cohesion policy, i.e. a specific action which is taken beyond the structural funds that are foreseen in Title XVIII and that is promoted through ordinary legislative procedure.

It is worth noting that within the TFEU poverty prevention “per se” is not a policy that falls under the EU mandate, if not within the framework of development and cooperation policies (union external actions). Nevertheless, social exclusion is by all means a field in which the EU has right to act, so reducing poverty as an obstacle to social inclusion is definitely within the EU powers. This slight distinction has however some non-negligible implications when it comes to the intervention logic of the FEAD and, in turn, to the requirements for its monitoring and evaluation systems, as it will be explained further below.

As to the structure of the FEAD, it remains an indirect management fund, so there are managing authorities which draft an operational programme that need to be approved by the EC and defines, according to a standard template and a set of common rules, criteria for the implementation of the operations under their remit in terms of the typologies of support offered, the objectives, the implementing actors as well as the monitoring arrangements. The FEAD goes also a long way into enacting the subsidiarity principle, as it strongly involves networks of partner organisations (POs) which are typically anti-poverty/charity networks, NGOs and other not for profit entities, and are responsible of the delivery of support on the ground.

When it comes to the fund’s intervention logic and structure, within the frame of the FEAD, unlike the rest of cohesion policy, there are no “thematic objectives” which need to be supported, but MS have been left the possibility to choose between two slightly different approaches to the definition of their operational programmes, namely:

- Operational Programme I: 'food and/or basic material assistance operational programme' which is concerned with the distribution of food and/or basic material assistance to the most deprived along with accompanying measures to be provided by partner organisations (mostly NGO) which distribute food or basic material assistance on the ground;
- Operational Programme II: 'social inclusion of the most deprived persons operational programme' which is not concerned with food or basic assistance support but rather aims to bridge the gap between social inclusion measures foreseen in the ESF or by national policies and some target groups which could not be otherwise covered through the provision of non-material assistance (especially guidance on available social integration opportunities or general counselling, basic skills trainings, civic education, health promotion etc.)

Either way, according to its implementing regulation²⁷ and in line with the remit of the Treaties, the FEAD’s ultimate goal is to “*promote social cohesion, enhance social inclusion and therefore ultimately contribute to the objective of eradicating poverty in the Union by contributing to achieving the poverty reduction target of at least 20 million of the number of persons at risk of poverty and social exclusion in accordance with the Europe 2020 strategy, whilst complementing the Structural Funds*”.

²⁷ Reg (EU) n. 223/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 11 March 2014 on the Fund for European Aid to the Most Deprived

Most MS (all 28 except the Netherlands, Sweden, Germany and Denmark) opted for operational programme I, thus of the approx 3.5 billion euros allocated for the EU as a whole, only just over 0.1 is went to OP II.

From this perspective, and because it is acknowledged that food or basic assistance cannot per se alleviate poverty, accompanying measures are also foreseen as a mandatory requirement for Partner Organisations distributing food or material support,²⁸ and the and this is why also accompanying measures can be funded (at a 5% flat rate) by the FEAD. In addition to this, the OP can fund at a maximum of 5% activities such as capacity building for Partner Organisations, in addition to monitoring and evaluation.

Finally, the FEAD Regulation also defines in its article 5 a range of horizontal principles, including trying to contribute to the reduction of food waste, ensuring that gender equality principles are mainstreamed in the operations and that the partnership principle is duly applied.

2.3.2 Monitoring and evaluation arrangements

In addition to defining the objectives and structure of the FEAD Operational Programmes, the FEAD Regulation contains specific provisions for the monitoring and reporting of the FEAD that seeks to ensure that the fund's contribution to its objectives can be assessed and, especially, made visible. Monitoring, indicators and reporting requirements are defined in the Commission Delegated Regulation (EU) No 1255/2014 and reporting requirements on indicators are further explained in the Guidance Fiche Monitoring Under FEAD (DG EMPL, 2015).

Monitoring and evaluation arrangements broadly include:

- for the Managing Authorities in the Member States:
 - the drafting and submission of an Annual Implementation Report;
 - for Operational Programmes I, the roll-over of a structured survey on end-recipients (the most deprived) and partners organisations, to be finalised by the end of 2017 and 2022 but **no mandatory evaluation required**;
 - for Operational Programmes II, at least one evaluation of the effectiveness, efficiency and impact of the operations supported;
- for the European Commission:
 - a mid-term evaluation of the fund, by the end of 2018; and
 - the possibility to evaluate OPs at their own initiative during the programming period.

The key monitoring tool is thus the Annual Implementation Report, which is accompanied by a Survey for Operational Programmes I. In the Annual Implementation Report (AIR), along with information on the operations implemented and challenges identified, quantitative monitoring data is collected, defined as per the table below.

²⁸ Unless the operations foresee only support to children, as per art. 7(4) of the FEAD regulation Reg (EU) n. 223/2014. It can be either the partners organisations distributing food that provide for accompanying measures or other organisations that work as additional partners.

Table 7 – list of indicators to monitor the FEAD

Type	Operational Programme I	Operational Programme II
Input indicators	<p>(1) Total eligible expenditure approved.</p> <p>(2) Total eligible expenditure incurred by beneficiaries and paid. Thereof, where relevant:</p> <ul style="list-style-type: none"> (a) [...] food support; (b) [...] basic material assistance. <p>(3) Total amount of eligible public expenditure declared to the Commission.</p>	
Output Indicators	<p>on food support</p> <p>(4) Quantity of fruits and vegetables and from (5 to 11) quantity by type</p> <p>(11) Total quantity of food support distributed. Thereof:</p> <ul style="list-style-type: none"> (a) % of food for which only transport, distribution and storage were paid for by the OP (in %); (b) proportion of FEAD co-financed food products in the total volume of food distributed by the partner organisations (in %). <p>(12) Total number of meals distributed partly or totally financed by the OP.</p> <p>(13) Total number of food packages distributed partly or totally financed by the OP</p> <p>on basic material assistance</p> <p>(15) Total monetary value of goods distributed, by Target Group:</p> <p>(16 – 18) list of most relevant types of goods distributed, by target group and type</p>	<p>on social inclusion assistance</p> <p>(20) Total number of persons receiving social inclusion assistance, by Target Group:</p> <ul style="list-style-type: none"> (a) children(<15 y.o.); (b) elderly >65 y.o.; (c) women; (d) migrants, participants with a foreign background, minorities [...]; (e) persons with disabilities; (f) homeless
Result Indicators	<p>on food support distributed</p> <p>(14) N° of persons receiving food support, by Target Group</p> <p>on basic material assistance</p> <p>(19) N° of persons receiving basic material assistance, by Target Group</p>	<p>No common result indicator defined, only programme specific</p>

Source: author's elaboration (synthesis) based on Reg (EU) No 1255/2014, Annex

Whilst input indicators refer to expenditure across all types of FEAD OPs, the first key point is that a distinction can be drawn between the approach of Operational Programme I and II, which is reflected in the way output and result indicators should be interpreted.

In Operational Programmes I, the amount/value of the food/goods distributed is considered as an output whereas the number of end-beneficiaries receiving food or basic assistance support is considered as a result of the programme. The rationale behind this is that OPs I are aimed at providing immediate relief to the most deprived, therefore the latter's mere participation can be considered as a result, according to the fund's logic. Conversely, in OPs II, the participation to FEAD activities is considered as an output, whereas participants' better inclusion as a result. Furthermore, here MAs are allowed to define specific output and result indicators, in the light of the broad range of target groups and needs that could be addressed by the related interventions.

Another key difference in the monitoring is that no targets are set for common output and result indicators. Only in the case of programme-specific indicators for OPs II, target values are set. This implies that for a major share of monitoring data, there exist no obvious and direct benchmarks which can help assess whether the implementation is on track in terms of outputs and results. Thus, to compare MS one needs to factor in a number of features both of the programme and of the social context, such as FEAD financial allocation, poverty reduction targets and/or the amount of persons at risk of poverty or social exclusion in the country - to name but a few. This is however a limitation that affects data analysis.

The issue of participation is also concerned by the way data is collected under FEAD. According to the FEAD Regulation, it is for the Partner Organisation to keep a record of their activities. However, with a view to limiting the associated administrative burden, they are allowed to make estimations rather than a thorough assessment on end beneficiaries. In particular, for result indicators²⁹ for OPs I these “[...] can be reported based on an informed estimate by the partner organisations. These estimations can be based on sampling approaches (e.g. counting of recipients at certain days/weeks of the year and extrapolated) or other methods. Managing authorities are required to document the methods used and accepted.” (DG EMPL, 2015:8) This element should be read also in the light of an additional monitoring requirement: participants should be counted only once irrespective of the number of meals/food packages they receive under FEAD support. The underlying rationale is that there would be high merit in understanding to what extent the potential audience of FEAD is addressed by its support and – equally – with what intensity (i.e., how many times per year, with how much food/what value of goods etc).

However, practice from previous years of FEAD implementation has showed (COM, 2017:4-5) that the respect of this principle, alongside the use of estimates on participants, can prove challenging. This issue however does not affect OPs II in that Article 32(d) of Regulation 223/2014 requires MAs to collect micro-data at the level of the individual participant.

Even more strikingly, understanding effectiveness towards social inclusion with result indicators that only count the number of people reached is obviously not possible, as will be further discussed below.

²⁹ Following the same reasoning, also the amount of food distributed can be based on an estimation. There are however expectations that this estimation is structurally more reliable than the counting of participants.

Aware of the limitation of monitoring data and likely in an attempt to fill the gaps, the FEAD Regulation foresees the roll-out of a structured survey. The main questions of the 2017 Survey are listed in the table below.

Table 8 – summary of questions in the structured survey 2017 of Operational Programmes I

Respondent	Main questions
Partner organisations	A1. Types of FEAD assistance, A2. types of Accompanying Measures A2-4 other forms of support provided
End-recipient	<p>B1-3. Background detail (age, gender, target group)</p> <p>B4-6 type of assistance received and by whom</p> <p>B7-9 whether it is the first time getting assistance, how often that happens and if the need is expected to continue in the future</p> <p>B10 main obstacles in receiving support</p> <p>B11 “has FEAD made a difference to you or other members of the household”, and if not, why</p> <p>B12-13 other forms of assistance received by the respondent or their family members</p> <p>B14-15 types of accompanying measures received and whether they have been useful</p> <p>B16-23 other background information on income, accommodation etc.</p>

Source: author’s summary based on Commission Implementing Regulation (EU) 2016/594

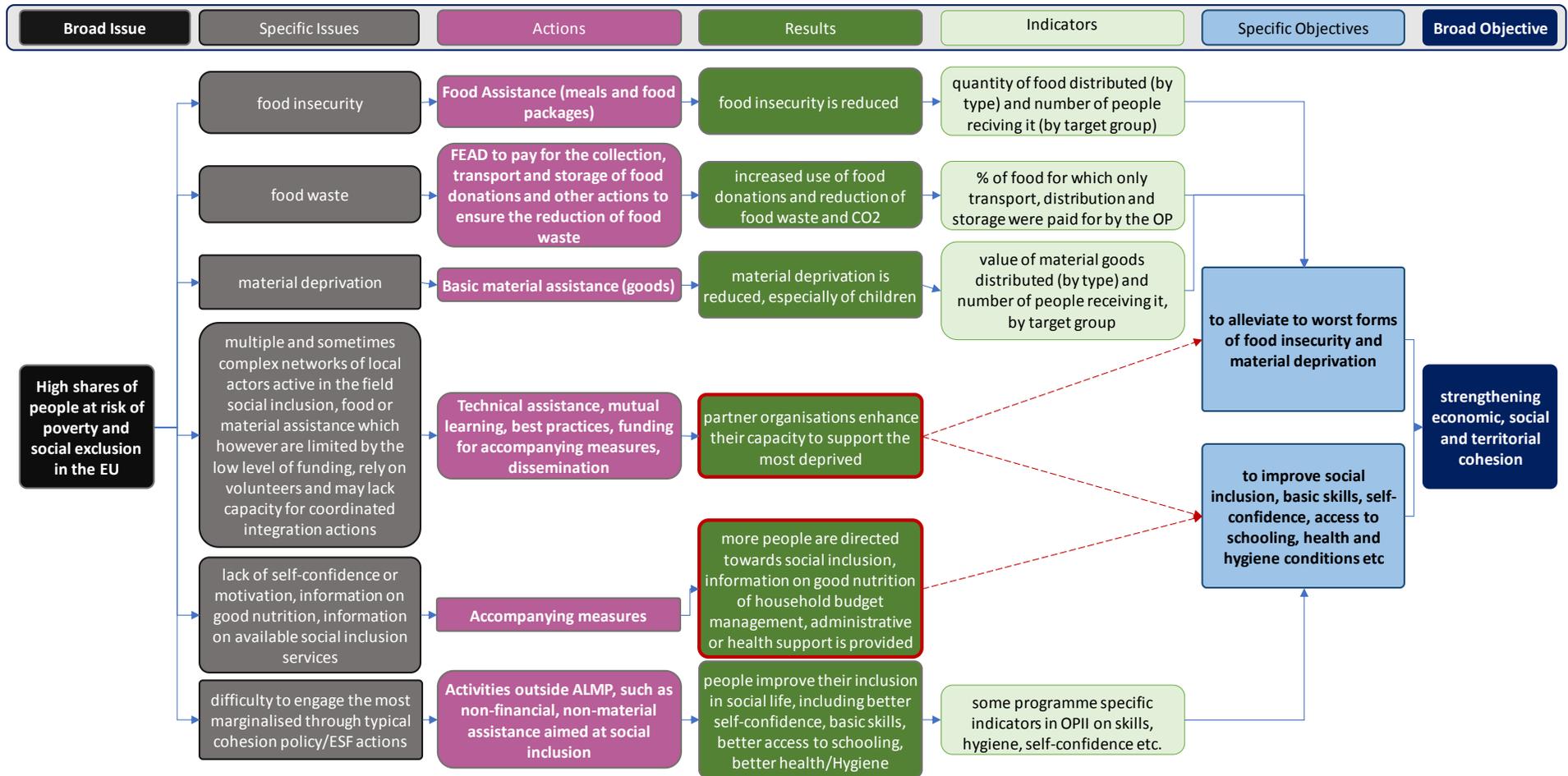
As it can be seen, some questions on the accompanying measures and effectiveness of FEAD support have been included by the Commission in the mandatory structured survey of partner organisations and end-beneficiaries.

In addition, in 2016 and 2021 and in the final implementation report, the MAs need to report on the contribution of the supported operations to achieving the objectives of the funds.

2.3.1 FEAD’s intervention logic

Based on the above description, an intervention logic of the fund is provided which seeks to highlight also if and how indicators defined within the FEAD M&E can help monitor progress towards the achievement of the expected results and objectives.

Figure 7 – FEAD’s intervention logic



Source: author’s elaboration based on the FEAD regulations and related act

The intervention logic above shows the rationale as well as relevance of the FEAD actions towards the end of mitigating the most extreme forms of poverty and material deprivation. However, it also highlights some key gaps in the measurement of the fund's effectiveness, which will be further discussed in the sections below. Preliminarily, it should be noted that, especially in lack of targets set based on a gap analysis ex-ante: (i) only knowing how many people have received food or basic assistance support doesn't allow to formulate a judgement on how much of the need was covered; (ii) only knowing how much of the FEAD money was used to leverage on food donation cannot tell what is the result in terms of reduced food waste; (iii) no monitoring information is available that deals with the frequency, quality and outcome of accompanying measures provided; and finally (iv) information on the results on improvement of partners organisations is not foreseen. In addition, some specific objectives of the aid are difficult to measure in relation to overall At Risk of Poverty or Social Exclusion (AROPE)³⁰ figures as FEAD is too small a fund to realistically expect a decisive contribution to reducing poverty and – even if this were the case – it is difficult to distinguish the effects of FEAD compared to other, mainly national, social protection benefits and the overall development of the economy (see for example Greiss et. Al, 2019).

Nevertheless, FEAD is expected to make a contribution to the Europe 2020 target of reducing the number of EU citizens living in poverty by 20 million. This includes a reduction in homelessness and in child poverty, in food waste, in CO2, improvements in the delivery mechanisms and in the coordination of the delivery of food and material goods as well as in the social inclusion and employment of people affected by poverty. FEAD is also expected to enhance the capacity of the partner organisations, decrease administration and contribute to the more efficient implementation of aid through the spread of good practice. Impacts can also be understood in terms of differentiated impacts such as higher registration in social inclusion support programmes, ESF measures, and other formal and non-formal training activities. Possible indicators for impacts would be, e.g. beneficiaries registering in social services programmes upon participation in FEAD interventions, beneficiaries entering ESF measures upon participation in FEAD interventions, or beneficiaries taking part in formal/non formal training, acquiring basic skills, searching for/getting a job upon participation in FEAD interventions. Yet, these results are not monitored.

More importantly, as it is known that monitoring indicators could not, in any event, say anything (if taken in isolation and outside of a counterfactual study) about effects, in lack of targets and precise figures on the volume of activity on the ground, there are doubts as to the capacity of the FEAD monitoring system to fulfil its first purpose, which is to inform managing actors on whether things are going according to plan, or deviate from it.

Targets are set in Operational Programmes II, but these only refer to programme specific indicators, which make aggregation as well as cross-programme comparison very hard indeed.

³⁰ As per the Eurostat website the AROPE “corresponds to the sum of persons who are either at risk of poverty, or severely materially deprived or living in a household with a very low work intensity. Persons are only counted once even if they are present in several sub-indicators. The AROPE rate, the share of the total population which is at risk of poverty or social exclusion, is the headline indicator to monitor the EU 2020 Strategy poverty target.” Available at: [https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:At_risk_of_poverty_or_social_exclusion_\(AROPE\)](https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:At_risk_of_poverty_or_social_exclusion_(AROPE))

2.3.2 Main issues with the FEAD Monitoring and Evaluation system as emerging from the literature and work in the MTE evaluation

Along the lines of the considerations reported above, which are based upon the reconstruction on the FEAD intervention logic and a brief analysis of its monitoring arrangements, a number of similar issues have been discussed over recent years about the M&E system of the FEAD. This happened especially with a view to its future merge with the ESF and other funds within the ESF+ and in the context of the Mid-Term Evaluation (Metis et. Al, 2018) and of the Impact Assessment of Human Capital funds (FGB et Al, 2018). Overall, the overriding concern of the EC and implementing actors (intermediate bodies and partner organisation) was on keeping the monitoring and evaluation system as light as possible given the “emergency” nature of the fund. This however means strong limitations when it comes to evaluating whether the FEAD works as expected and contributes to its statutory objectives. There’s also a risk that such a slim monitoring system fails to highlight the key benefits from the FEAD operation, potentially undermining support for its continuation. Below the main issues encountered, including by the author, within the frame of the FEAD MTE evaluation are summarised. These were also discussed in the context of the Impact Assessment on Human Capital Investments, and listed as follows:

- **lack of indicators on accompanying measures:** clearly, what is not covered by the monitoring system and the common FEAD indicators are the accompanying measures under OP I and the results of the social inclusion assistance under OP II. The survey on end-beneficiaries of type I OPs has however tackled this issue through three specific questions, notably:
 - To partner organisations: What types of accompanying measures have they received?
 - To end beneficiaries: Could you please tell me whether you found this advice or guidance useful or not?

It was also asked to beneficiaries whether they believed FEAD had made a difference to their lives and, if not, why so, including the need to receive different forms of support. Interestingly, the FEAD MTE study from Metis & Al. (2018) investigated the correlation between the level of satisfaction of FEAD support and the offer of accompanying measures, also controlling for the quantity of food pro-capita distributed, finding a statistically significant positive correlation between the two. This made the report clarify that accompanying measures seem to have played a role in the satisfaction of end recipients and that food alone was not enough to make a difference to their lives. However, the report also highlights issues with the provision of accompanying measures due to:

- insufficient funding (only 5% flat rate reimbursement possible);
- lack of competence especially due to the fact that staff of partner organisation is chiefly made of volunteers; and, importantly
- optional nature of accompanying measures and no dedicated monitoring on these.

Such a deficit in monitoring and provision of accompanying measures was highlighted also in the latest report of the EC to the EP, which further highlighted that this has hampered monitoring effectiveness towards social inclusion.³¹

Even more markedly, the 2019 report from the European Court of Auditors (ECA, 2019:5) suggests that “Only a few Member States monitor the results achieved by these measures. Consequently, their contribution towards alleviating the worst forms of poverty could not be established” and recommends “Improving the assessment of social inclusion of FEAD recipients”;

- **number of participants may entail double counting:** to avoid double counting, it was asked Partner Organisations to register participants so that those coming several times to the same place were not counted as two different participants. However, if a participant moves from one partner organisation to another the lack of a dedicated electronic monitoring means that it is impossible for POs to actually count them just once;
- **no target on the use of the FEAD to improve on food waste:** although the FEAD monitoring system devoted one indicator to the share of food distributed for which only the transport, distribution and storage costs were funded by the FEAD, this possibility was actually explored only in Luxembourg and, by the end of 2016, Slovakia. Data from the MTE shows that in Luxembourg the efficiency of the FEAD is a lot higher, arguably due to the fact that this option was used, so a much greater leverage effect was achieved. But the lack of targets might have discouraged MAs to make decisive efforts in this direction, opting for more traditional forms of support instead, which are however potentially less efficient and contribute less to reducing food waste. In 2018, the EC has taken action through the Omnibus regulation to facilitate the procedure for cost reimbursement for food donations,³² but changes to the monitoring system have not been mentioned;
- **some indicators on the typologies of food not very relevant in practice:** although there’s an interest in understanding not only how much food was provided but also what type of food (that can heavily affect efficiency) it is however highlighted (DG EMPL, 2018:43) that this division was not of much added value in the analysis, also given the granularity of the typologies of food described (e.g. sugar vs flour vs milk, not much difference in price);
- **for type II OPs, there is no requirement to have a breakdown by gender for programme-specific indicators:** FEAD delegated regulation (EU) No 1255/2014 supplements the FEAD main regulation by defining the list of all input, output and result indicators for type I and II OPs. However, for type II OPs only one output indicator is defined (number of people receiving social assistance) which breaks down the number of people in given target groups, including women. However, for programme-specific

³¹ “Accompanying measures are a key element of the fund as they aim to alleviate social exclusion of the most deprived. However, Member States’ reporting was limited, which did not allow for an in-depth examination. This was partially due to the above-mentioned implementation delays, but still posed a challenge in terms of the extent to which its contribution to the FEAD’s objective of enhancing social inclusion could be assessed.” COM (2018) 742 final

³² ECA, 2019 “The share of donated food integrated in the programme is still low. The Commission has addressed this concern by amending the FEAD Regulation through the “Omnibus Regulation” adopted by the European Parliament and the Council and which entered into force in August 2018. Member States will therefore be able to define flat rates, lump sums or unit costs (simplified cost options) and use them as the basis for payments to partner organisations that collect and distribute donated food.”

indicators (both output and results) for which MAs are required to set a baseline and target value, there is no indication of a mandatory breakdown by gender. Data should be collected by gender “where relevant”. However, no gender disaggregated data is reported in AIRs (and thus it was not given to the contractors who carried out the MTE, i.e. Metis & AI (2018));

- **no micro-data from the structured survey:** the contractors have not been granted access to the background features of the participants giving some specific answers. This did not allow testing whether answers/ levels of satisfaction with the support differed significantly depending on the target group;
- **no details on expected results (no targets nor baselines) for Operational Programmes I:** type I OPs do not have to provide details on expected results, and only estimations of the number of individuals addressed by their interventions. Hence, neither baseline nor target values are set, which makes it harder to gauge the level of progress towards the achievement of the OP objectives and more generally the implementation of the programmes. In the FEAD MTE, who is writing has developed an alternative methodology to measure progress and effectiveness, which is based on:
 - **number of participants reached vis-à-vis overall allocation;** and
 - **number of participants who received food support vis-à-vis a proxy of the population that is food insecure** (needs coverage).

However, both measures of progress had significant biases, and in particular:

- measuring progress as number of participants vis-à-vis the overall allocation can be considered a proxy of advancement (the more the programme unfolds, the higher the number of people per euro of allocation). However cross-country comparison can be affected by:
 - the form of support (food support vs material support);
 - the quality of the goods provided; and
 - differences in purchasing power.

This measure could however be used with somewhat higher reliability in a longitudinal fashion (year on year, provided that the forms of support remain somewhat stable over time);

- **needs coverage:** the advantage of this proxy is that it can directly compare the number of people supported with the population of those suffering from food insecurity. However:
 - issues with the data on participants may affect the estimates;
 - food insecurity rates differ depending on the target groups. The definition of target groups under FEAD do not exactly match the statistics which could be found with Eurostat. So it is hard to know exactly, say, youth below 15 y.o. are actually food insecure and have been reached through the FEAD;
 - food insecurity rates are just not collected for all target groups; and
 - the number of homeless is not known. Different sources were exploited to have a rough understanding of how many people are suffering from homelessness or housing exclusions in the different countries, but only rough figures could be found.

- **no indication on improved sanitary conditions/health related indicators:** despite the nature of the FEAD being that of a fund which aims at providing help to the most deprived, and health being a major component of human capital especially for them, this issue doesn't seem to be tackled at all by the fund.

2.3.3 The FEAD vis-à-vis the Monitoring of HC checklist

After having identified some of the key issues that were reported in the literature and/ or have emerged within the studies carried out with Metis and FGB on the FEAD monitoring system and impact assessment, let's see how the FEAD system scores when it comes to the checklist developed based on the work done in Chapter 1. As anticipated in the case of the ESF, the checklist should allow some new angles to be brought to the fore, especially when it comes to the broader issue of human capital development and its measurement. Table 8 below, presents the findings of this case study with respect to the questions included in the checklist for the monitoring of human capital development initiatives.

Table 9 – assessment of the FEAD monitoring system vis-à-vis the monitoring checklist

Criteria	Specific question	Evaluation
Comprehensiveness	Does it cover all objectives of the programme, and especially the operational and specific ones with new arrangements?	To a limited extent. The main objective of the programme is to lift 20 million people out of poverty. Whilst it is acknowledged that the FEAD cannot alone and given its volume produce a sizeable contribution to this objective, there are no indicators which measure intermediate outcomes, especially on people directed towards further social inclusion in type I OP. The result-orientation in type II OPs is somewhat stronger but there are no common result indicators there either, only some which are programme specific and cannot be aggregated. However, see next point on the role of surveys.
	Does it do this through surveys as well as continuous monitoring?	Yes. Both surveys and continuous monitoring are used to track progress and lower administrative burden. However, detailed data from the surveys was not made available to the contractors carrying out the mid-term evaluation of the FEAD.
	Do the indicators measure the different impacts on the different target groups, both direct and indirect?	To a limited extent, as there is no data on the actual results of the funds, only participations and some indication on results (but programme-specific) in the case of type II OPs. However, the monitoring does distinguish among the different target groups supported when it comes to results (participants reached). It does not do so in terms of quantity of food distributed or values of goods.
	Are the baselines and targets identified to track progress?	To a very limited extent. Baselines and targets are set only on programme-specific indicators for type II OPs. No milestones are however envisaged.
	Are transversal objective such as gender equality, non-discrimination, inclusion of people with disabilities included in the monitoring system?	To a good extent. Qualitative sections are foreseen (i.e. AIRs section 2.2) on the respect of horizontal principles (gender equality, food waste, complementarity with ESF, etc.) in both type I and type II OPs. In type I OPs data on participations (which is a “result” in the Intervention Logic of the fund) is broken down by gender and covers also other marginalised or disadvantaged target groups. Nevertheless, in the case of type II OPs, the breakdown by gender in programme specific output and result indicators is not reported in the AIRs although MAs are generically required to store this data.
	Does it look at factors contributing to or impeding achievement of the outcomes?	To some extent. In the template for AIRs there is no dedicated section (unlike the ESF /YEI) on issues affecting performance. Nevertheless, these factors are discussed within the frame of evaluations through field research and also thanks to the FEAD Network activities.
	Is monitoring data suitable to be used for assessing the impact of the funds	No. In the evaluations and EC documents it is argued that the emergency /first aid nature of the fund means that monitoring should be necessarily light and respect the dignity of the end-beneficiaries. This inevitably means that evaluations, especially counterfactual evaluations, are not standard practice in the FEAD environment. This is particularly evident as there is no obligation set in the regulation to evaluate

Criteria	Specific question	Evaluation
	through counterfactual or other econometric approaches?	type I OPs and only before 2022 in type II OPs. ³³ Also a discussion on macro-economic impacts seems somewhat beyond the possibilities of FEAD. However, such an initiative may have significant impacts that are at the moment not evaluated.
Participation	Is feedback from partner and beneficiaries foreseen in the monitoring? If so, how? (e.g. steering committees etc)	Yes. Feedback is included in the AIR (section 3) and collected through various research activities.
	Are experts and other stakeholders involved in the definition of monitoring arrangements?	To some extent. In the context of the MTE and the Impact Assessment of Human Capital Investments monitoring arrangements have been discussed with stakeholders but no holistic analysis was made. This is also where this research aims to contribute.
Proportionate and no overlaps	Does it treat the monitoring in a cost-effective manner?	Partly. This issue is hugely debated. Evidence from field research points to the fact that the requirements are complex especially for volunteers in Partner Organisations. However, it is also largely acknowledged that the monitoring system is light overall. It was highlighted however that particularly some information on the typology of food distributed was not very relevant to the analysis, so there's room for improvement there.
	Does it maintain indicators constant over time unless new major needs emerge?	Too little room for change. There's an overriding concern on not changing monitoring requirements, which is legitimate if one considers the nature of the FEAD but risks hampering result-orientation. The suggested changes to the monitoring framework (i.e. introduction of some monitoring requirements on accompanying measures and food waste, and reference values to be developed) were only partly considered. Section 2.3.5 also highlights how monitoring requirements on accompanying measures are mentioned in DG EMPL's impact assessment which goes together with the regulation but do not appear in the proposed legislative text.
	Does it only prescribe new monitoring arrangements where needed and at a reasonable cost?	Partly. As there is no use of administrative registers or other data (except for Eurostat in the context of needs assessment for the MTE) it is difficult to argue that there has been an effort to use existing monitoring arrangements. The system is however so light that a judgment here applies only partly.

³³ According to the report of the EC to the EP (COM (2018) 742 final) there are 4 MS that have carried out evaluations, but these are either management-related or customer-satisfaction type evaluations.

Criteria	Specific question	Evaluation
Timeliness	Does it collect evidence in a timely fashion, i.e., not necessarily from the outset but only and every time this is needed?	To a limited extent. Although data is collected yearly and should be suitable for the analysis, this is not very useful for an assessment on the effects of the FEAD. The structured survey for type I OPs is a valuable source to this end, but is done only once and arrived somewhat late for the evaluation (findings only integrated in the final report and they could not feed field research activities. In addition, not all countries sent the data on time and no disaggregation on respondents' data). Also the judgement from MAs on the effects of the FEAD are foreseen for the AIR 2019 and 2025, which means they haven't been used for the MTE and it is unlikely that the 2025 ones will be available for the final evaluation of the FEAD too.
	Are milestones foreseen?	No. In general, baselines and targets are not foreseen, except for programme specific indicators of type II OPs.
Accessibility	Is data collected accessible for the general public?	To some extent. There are datasets made available in the open cohesion portal (but only on payments) and the reporting is public (MTE, reporting of the EC to the parliament). ³⁴ In addition, Annual Implementation Reports are available on the MAs web pages. However, no complete dataset with AIR data is made available to the general public and survey data is not available either.
Relevance	Are only the relevant indicators included in the monitoring list or there's just too many of them?	Partly. There's a need to change the balance. Some indicators on typology of foodstuff are considered redundant, whilst some monitoring on key elements of result-orientation of the fund are missing.
	Are indicators updated and changed where necessary (and only then)?	Too little room for change. As above. The FEAD is a new fund so it is difficult to look at previous monitoring requirements, but it looks like in the future only marginal changes will be applied to the fund.
	Are indicators set at a level where decisions are taken either on the policy or on the programmes?	To some extent. Although data is reported at the National Level, poverty prevention policies are typically designed by national laws too (despite measures being often applied at the local level). The need for more disaggregated/ local data may however stem from the fact that social services traditionally see a strong role of municipalities. Better integration with local social services would also increase visibility of action/ strengthening the "poverty emersion" function of the FEAD.
	Does it take into account the main dimensions of human capital (capacity,	To a very small extent. The fact that the result orientation of the FEAD is quite modest due to the need to have a very light monitoring system means that very little is known about how much it contributes to

³⁴ COM(2018) 742 final REPORT FROM THE COMMISSION TO THE COUNCIL AND THE EUROPEAN PARLIAMENT Summary of the annual implementation reports for the operational programmes co-financed by the Fund for European Aid to the Most Deprived in 2016, available at <http://data.consilium.europa.eu/doc/document/ST-14699-2018-INIT/en/pdf>

Criteria	Specific question	Evaluation
Suitability to measure human capital development	deployment, development, know-how and creativity, non-cognitive skills , health , culture and tolerance) as relevant to the objectives of the initiative?	human capital development through social inclusion. However, its function is potentially very important to the end of developing human capital in its broad meaning (health, know-how, self-confidence) and many soft outcomes have been pointed at in evaluation studies on the role of the FEAD. Some of these are traced through programme specific indicators in type II OPs, but these cover only 4 countries and account for 2.44% of the FEAD budget (source Metis et. al, 2018). Amongst other, improved health conditions, improved access to early schooling, improved skills, stronger social network. But these are very seldom tracked.
	What is the underlying approach to measuring human capital, i.e., monetary or non-monetary?	None. Except for the cost of the fund, there's basically no approach to measuring the impact on human capital development in type I OPs, only number of people receiving foodstuff or basic goods. In type II OPs the approach is again non-monetary, indicator based.
	Does it take into account relevant direct (first order) impacts as well as indirect (second order) effects?	Partly. Monitoring focuses on individuals and there are no results collected on institutional changes etc. Qualitatively, the impact on Partner Organisations is assessed through focus groups or interviews and referred to in the studies.
	Do indicators measure input, process, output or outcome-related factors of human capital development?	Partly. The nature of the FEAD means that its intervention logic differs from standard interventions on human capital, especially for type I OPs (results are just people having received food or first necessity goods). There's nothing on process (e.g. on the capacity building for NGOs). In type II OPs the situation becomes closer to the Borowski framework although results are rather output (number of people supported). Some programme-specific result indicators of type II OPs look at actual results (improvements in conditions).
	Does it measure intermediate output/outcomes which can be linked to the intervention (enablers, soft outcomes)	To a very limited extent. Common output indicators for type II OPs look at people assisted towards social inclusion and programme specific indicators for type II OPs aim to look at intermediate outcomes (improved skills, improved health or hygiene conditions, improved social networks).

2.3.4 FEAD in the ESF+ and its foreseen monitoring

With the new programming period coming closer and the debate on its regulation being currently at the level of interinstitutional discussions, there are already some considerations that might be done on the current EC proposals, accompanying documents and European Parliament's amendments on the expected changes. The biggest of these changes is inevitably the FEAD being merged within the ESF+ with the ESF, YEI under shared management as well as with the Health and EaSI programmes (which remain however under direct management). Given the specific nature of the FEAD and the paramount need to respect the dignity of end-beneficiaries as well as reducing administrative costs for Partner Organisation (whose workforce is typically made up of volunteers and/ or people in retirement age) there has been strong support to keep monitoring requirements separate for the FEAD investment priority, namely *“addressing material deprivation through food and/or basic material assistance to the most deprived, including accompanying measures”*.

It is currently not entirely clear which direction will the formerly FEAD priorities take especially in terms of monitoring requirements as the EC's proposal and Parliament's amended text differ.

According to the former, type II OPs would cease to exist as a stand-alone priority or fund with dedicated rules, and be subject to the monitoring requirements which are common to other ESF+ priorities.

For type I OPs, *“Monitoring of performance will be improved by setting baselines, and adding a few indicators on access to services, to reflect the extension of scope of assistance. The potential additional burden of these new requirements will be offset by the simplification of common output indicators (such as food type provided). It is also envisaged to slightly expand the questions of the structured surveys for end-recipients with additional input from partner organizations involved, which is already the current practice in Member States.”* (DG EMPL, 2018). This would be coherent with the findings of the MTE and EC report to the EP on the implementation of the FEAD.

However neither in the text of the proposed regulation (Chapter III, article 21 “indicators and reporting”) nor in the related annex II, there is mention of indicators on access to services. Reference values are added. The text of the structured survey will be published in the form of an implementing regulation at a later stage.

According to the Annex II of the proposed regulation,³⁵ foreseen changes to the reporting indicators include:

- total value of food and goods distributed now disaggregated into three target groups (children, homeless and other target groups);
- no distinction as to the typology of food provided and the form of provision (food packages vs meals); and
- no gender disaggregation nor indicator focusing on the number of women addressed.

³⁵ Currently only available at <http://data.consilium.europa.eu/doc/document/ST-9573-2018-ADD-2/EN/pdf> in that the annexes to the proposed regulation are no longer accessible from the DG EMPL website as of July 14, 2019.

However, this proposal was amended by the EP's on 4 April 2019, with the main change concerning indicators for the former type II OPs, now having to measure the number of individuals directed towards social inclusion services. In addition, the group of Roma participants should be kept separate from that of other minorities and migrants.

Changes go almost entirely in the direction of making the monitoring system even lighter, which is a fair request from implementing actors but has consequences as to possibility of monitoring and evaluating the fund.

2.3.5 Key outstanding issues and recommendations

As the changes to the fund are still in the process of being discussed, below are the main recommendation that stem from the analysis and literature consulted, including taking into account the need for a more encompassing definition of human capital development.

- **Improve result orientation, including by:**
 - **Setting up monitoring arrangements for accompanying measures:** lack of result orientation especially on accompanying measures, which are key to the aim of social inclusion and more generally to make a difference to the lives of those concerned, was highlighted from various stakeholders and studies, including by the European Court of Auditors (ECA, 2019). If there's a legitimate interest in keeping burden as low as possible for Partner Organisations, one could nevertheless suggest that this is monitored by social services (i.e. asking participants whether they have been directed there after receiving FEAD food or material assistance). The documents accompanying the regulations state that this requirement will be implemented but it is not yet to be found in the proposals, or limitedly to the formerly OP II investment priority, whose result orientation was much stronger than just counting people directed towards social inclusion services.
 - **Making disaggregated data from surveys available, and more generally survey data available on time.** As surveys are key to understand whether FEAD has made a difference to the lives of the end-beneficiaries (and, if not, why), these valuable info that cannot be extracted from the standard continuous monitoring systems should be made available to researchers in due time so that it can inform the field analysis, discussion with stakeholders etc. This issue is not yet tackled and the timing of the surveys is at present the same of the last round (4 years into the implementation for the first round and 2 years after the end of the programming period).
 - Asking for **target setting on results** (as it is foreseen now in the text of the regulation for the ESF+. The methodology followed by the MTE study (needs coverage) can be used as a basis for a needs assessment and then information on costs from the current programming period used as a reference (after careful consideration) for the following programming period.
 - **Target setting on** the use of FEAD-type interventions to fund the transport, storage and delivery of **food donations**, which, together with the new Omnibus regulation should help boost the use of FEAD-type interventions in this key field. This is not currently foreseen in the ESF+ proposal

- **To report values that are broken down by gender also in type II OPs-like interventions.** It is worth noting that at present the gender breakdown will be indirectly included for type II OPs, as this typology will cease to exist being absorbed under the standard strand of the ESF+, but no gender disaggregated data are at the moment foreseen for type I OPs (now proposed priority ‘xi’ of the ESF+)
- To **improve coverage of target groups**, by including amount of food or goods distributed to each of them
- To **measure** (potentially through survey approaches to keep the administrative burden low) **the impact on improved institutional capacity for partner organisations** or administrations in charge of social services
- To make the M&E better suitable to track impacts in terms of human capital development including by:
 - Investigating, potentially through surveys, the ways in which FEAD-type intervention have made a difference to the lives of end beneficiaries – if any, including:
 - Better health, hygiene
 - Better self-confidence
 - Better skills
 - Better access to schooling (or increased attainment rates, reduction of drop-outs)
 - Better social networks and civic participation

This can also aid steering efforts towards some specific objectives that the MS wants to pursue. One should also look at increased **security** for citizens

- To reduce administrative burden, including by:
 - Getting rid of the excessive breakdown of typologies of foodstuff or goods distributed. One possible distinction from food is that of using main macro nutrients, e.g.
 - Carbs based foodstuff (bread, rice, pasta, sugar, chocolate etc.)
 - Protein based foodstuff (eggs, milk, meat, fish etc.)
 - Fats-rich foodstuff (oil, butter etc.)
 - Vitamins-rich foodstuff (fruit, vegetables etc.)
 - Training personnel of POs to deal with a number of requirements of the monitoring system, including eligibility criteria for participants
 - Improving electronic exchange systems

Last, and even though the target population of the FEAD is arguably not affected by strong spontaneous dynamics which bring them into social inclusion without external support, in order to understand “what works, where and for whom” the need for counterfactual analysis, ideally through randomised experiments, remains strong also within the FEAD context – and it is unmet. This is all the more important if the programme is to bridge people ever closer to social inclusion and not just address them with first-aid type support.

2.4 Conclusions

The goal of this research was to take a closer look at the monitoring and evaluation systems of some key funds in the field of human capital development in the EU to understand, from the perspective of human capital development in its whole-encompassing sense, whether they are fit for purpose.

It adds to the current debate in two main ways, notably (1) by offering insight into the limits of the monitoring systems from the perspective of researchers involved in carrying out studies on the performance and effectiveness of the funds and (2) by assessing the suitability of the monitoring system through the analytical lenses of the checklist/ assessment grid formulated in chapter one, which brings together recommendations from several international organisations and an updated definition of human capital.

In essence, the case study analysis shows that, despite commendable efforts being made by those in charge of the design of monitoring and evaluation systems to strike a balance between the relevance and informativeness of monitoring system and their cost for all stakeholders, current monitoring systems would need improvements. In particular, they seem to fall somewhat short of fulfilling the need for which they are created in the first place, notably (i) supporting managing actors with timely and comprehensive information on the implementation and (ii) providing input into the evaluation of the programmes.

In the case of the ESF/YEI, the M&E system can be considered a comparatively efficient tool for management, but fail to properly inform evaluation, especially with respect to some of the funds' objectives.

For the FEAD, the monitoring system, in lack of targets and monitoring on some important element of the programme, might even struggle to signal whether things are proceeding according to plan or some factors are affecting implementation. So it can neither fully answer the information need from a management perspective, nor from that of the evaluators.

Some of these limits have been pointed at in dedicated studies on the subject matter or within the context of the debate on the Multiannual Financial Framework for the post-2020, including by other institutions concerned (see, e.g., the European Court of Auditors on the monitoring of the FEAD). But for instance in the case of the FEAD no dedicated assessment of its monitoring system exist.

Effectively, a low interest from the academia on the analysis of the pure “monitoring and evaluation system” emerges. And, some of the limits which were flagged up in these case studies are just arguably below the radar of the few analyses carried out so far, potentially as they only become apparent to researchers being involved in the actual evaluation of the funds, or because no dedicated analytical framework was used in assessing them. This is for instance the case of the study on the monitoring and evaluation of the ESF carried out by Applica et. Al (2018) which does not consider key issues such as the insufficient disaggregation of monitoring information, the scarce relevance of programme specific indicators to provide input into programmes' evaluation and the lack of information on typologies of operations supported. All three of these elements, as it will be seen also in chapter 3, significantly hamper impact evaluation.

In addition, from the perspective of human capital as “*The knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic well-being*” too, the quest for simplicity and low administrative burden has meant very scarce data collection on issues related to improvement of basic skills, self-confidence, non-cognitive skills, citizenship, hygiene and sanitary conditions and so forth. This is also partly overlooked by existing studies and insufficiently reflected in policy recommendations thereof.

Finally, the continuing overriding concern of the European Commission on keeping administrative burden as low as possible – which might be justified in light of the findings from consultations they carry out with stakeholders and an actual issue to be addressed – means that also in the new proposal for the ESF+ the design of the monitoring system would suffer from many of the same problems, including:

- ESF/YEI investment priorities within the ESF+:
 - lack of common indicators which can cover the aims especially of TO9, and on soft-outcomes/intermediate results;
 - no micro-data available and no information on the typologies of interventions;
 - programme specific indicators still there;
- FEAD investment priorities within the ESF+:
 - lack of indicators on accompanying measures and social inclusion and on enhancement of the capacity of partners' organisations; and
 - lack of targets on reduction of food waste.

In terms of recommendations, in this research it is argued that whilst keeping much of the structure of the system intact, some marginal gains could be achieved in particular by:

- asking Managing Authorities to label the operations they support via a standardised categorisation of typologies of interventions;
- making an effort, i.e. at the Commission level, to collect and collate micro-data, for it to become available for research, or to share slightly more disaggregated data in through the System for Fund Management (i.e. with info on participants' background features);
- requiring more standardised approaches to the evaluations, including Counterfactual Impact Evaluations (e.g. definition of given outcomes to be measured, at certain points in time, to increase aggregability – this is further expanded in chapter 3);
- getting rid of programme specific indicators, which increase costs without providing fully usable information; and
- getting rid of some unnecessary breakdown of indicators in formerly FEAD operation and training personnel to reduce administrative burden.

A more radical, although much needed, change would entail:

- extending the list of common indicators;
- introducing monitoring and counterfactual on social inclusion measures for the formerly FEAD activities;
- producing data that is always available at the NUTS2 level; and
- investing in the measurement of a range of soft-outcomes (or intermediate results) that are not directly employment or social inclusion but key milestones towards them.

All in all, it is clear that Monitoring and Evaluation should not in any event become the focus of a programme nor disproportionately steer resources away from the implementation of the support. Nonetheless, it should also be noted that if M&E - which remains costly even its simplest form – ends up leaving too much room for uncertainty its added value becomes negligible and does not justify its costs – for however small or reduced they might appear.

2.5 Main Bibliography

Applica, WIIW, Tarki, (2018), Study on monitoring and evaluation systems of the ESF, European Commission, DG Employment. Available at: <https://op.europa.eu/en/publication-detail/-/publication/9d82b3ca-6eb3-11e8-9483-01aa75ed71a1/language-en>

Alphametrics and Applica (2016) Monitoring and Evaluation of Cohesion Policy, European Social Fund, Guidance document Annex D - Practical guidance on data collection and validation, ESF Support Centre, DG EMPL, available at: <https://ec.europa.eu/sfc/en/system/files/ged/Annex%20D%20-%20Practical%20guidance%20on%20data%20collection%20and%20validation.pdf>

Borkowsky, A. (2013), “Monitoring Adult Learning Policies: A Theoretical Framework and Indicators”, *OECD Education Working Papers*, No. 88, OECD Publishing, Paris, Available at: https://www.oecd-ilibrary.org/education/monitoring-adult-learning-policies_5k4c0vxjlkzt-en

Card, D., J. Kluve and A. Weber (2018), What Works? A Meta-Analysis of Recent Active Labor Market Program Evaluations, *Journal of the European Economic Association* 16(3). Available at: <https://www.nber.org/papers/w21431.pdf>

COM(2016) 381 final, COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS A NEW SKILLS AGENDA FOR EUROPE Working together to strengthen human capital, employability and competitiveness {SWD(2016) 195 final}, Available at: <https://ec.europa.eu/transparency/regdoc/rep/1/2016/EN/1-2016-381-EN-F1-1.PDF>

COM(2017) 404 final – REPORT FROM THE COMMISSION TO THE COUNCIL AND THE EUROPEAN PARLIAMENT – Summary of the annual implementation reports for the operational programmes co-financed by the Fund for European Aid to the Most Deprived in 2015 Brussels, 28.7.2017, Available at: <https://ec.europa.eu/transparency/regdoc/rep/1/2017/EN/COM-2017-404-F1-EN-MAIN-PART-1.PDF>

COM (2018) 742 final – REPORT FROM THE COMMISSION TO THE COUNCIL AND THE EUROPEAN PARLIAMENT Summary of the annual implementation reports for the operational programmes co-financed by the Fund for European Aid to the Most Deprived in 2016, available at: <http://data.consilium.europa.eu/doc/document/ST-14699-2018-INIT/en/pdf>

COM(2018b) 382 final, Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the European Social Fund Plus (ESF+) {SEC(2018) 273 final} - {SWD(2018) 289 final}, available at: https://ec.europa.eu/commission/sites/beta-political/files/budget-may2018-european-social-fund-plus-regulation_en.pdf

COMMISSION DELEGATED REGULATION (EU) No 1255/2014 of 17 July 2014, supplementing Regulation (EU) No 223/2014 of the European Parliament and of the Council on the Fund for European Aid to the Most Deprived by laying down the content of the annual and final implementation reports, including the list of common indicators. Available at:

<https://op.europa.eu/en/publication-detail/-/publication/55a01a0c-747c-11e4-b593-01aa75ed71a1/language-en>

DG EMPL (2015) – Guidance Fiche Monitoring under FEAD, Brussels, European Commission, EMPL A3/SLG/JM (2015). Available at: https://circabc.europa.eu/sd/a/4f06fb0e-4821-4c6a-a9ce-ff1c51293d3a/2015_05_12_FEAD_06_Guidance_note_on_monitoring_and_indicators_under_FEAD-final1.pdf

DG EMPL (2018a), Monitoring and Evaluation of European Cohesion Policy European Social Fund Guidance document, August 2018, available at: <https://ec.europa.eu/sfc/en/system/files/ged/ESF%20monitoring%20and%20evaluation%20guidance.pdf>

DG EMPL (2018b), Guidance for Member States on Performance Framework, Review and Reserve, EGESIF_18-0021-01, version 2.0, available at: https://ec.europa.eu/regional_policy/sources/docgener/informat/2014/gn_performance_framework_review_and_reserve.pdf

DG EMPL (2018), COMMISSION STAFF WORKING DOCUMENT IMPACT ASSESSMENT Accompanying the document Proposal for a Regulation of the European Parliament and the Council of the European Union on the European Social Fund Plus (ESF+) and Proposal for a regulation of the European Parliament and the Council of the European Union on the European Globalisation Adjustment Fund (EGF), European Commission, SWD(2018) 289 final. Available at: <https://ec.europa.eu/transparency/regdoc/rep/10102/2018/EN/SWD-2018-289-F1-EN-MAIN-PART-1.PDF>

European Court of Auditors - ECA (2018) Free Movement of Workers – the fundamental freedom ensured but better targeting of EU funds would aid worker mobility, Special Report N°06, available at: <http://publications.europa.eu/webpub/eca/special-reports/eu-labour-mobility-6-2018/en/>

European Court of Auditors - ECA (2019) FEAD-Fund for European Aid to the Most Deprived: Valuable support but its contribution to reducing poverty is not yet established, Special Report N°05, available at: https://www.eca.europa.eu/lists/ecadocuments/sr19_05/sr_fead_en.pdf

European Parliament legislative resolution of 4 April 2019 on the proposal for a regulation of the European Parliament and of the Council on the European Social Fund Plus (ESF+) (COM(2018)0382 – C8-0232/2018 – 2018/0206(COD)). Available at: [https://www.europarl.europa.eu/RegData/seance_pleniere/textes_adoptes/provisoire/2019/04-04/0350/P8_TA-PROV\(2019\)0350_EN.pdf](https://www.europarl.europa.eu/RegData/seance_pleniere/textes_adoptes/provisoire/2019/04-04/0350/P8_TA-PROV(2019)0350_EN.pdf)

FGB, Ceps and COWI (2016), The analysis of the outcome of the negotiations concerning the Partnership Agreements and ESF Operational Programmes, for the programming period 2014-2020, European Commission, DG Empl, Luxembourg: Publications Office of the European Union, 2016 ISBN: 978-92-79-62769-9 doi: 10.2767/90132. Available at: <https://op.europa.eu/en/publication-detail/-/publication/b2c01d15-ffe6-11e6-8a35-01aa75ed71a1/language-en>

FGB (2017), Synthesis report of ESF 2016 annual implementation reports, European Commission, DG employment, Luxembourg: Publications Office of the

European Union, 2017, ISBN 978-92-79-71474-0. Available at: <https://op.europa.eu/en/publication-detail/-/publication/4fd56da6-99bf-11e7-b92d-01aa75ed71a1/language-en>

FGB (2018a), Synthesis report of ESF 2017 annual implementation reports - Final report – Study, European Commission, DG employment, Luxembourg: Publications Office of the European Union, 2018, ISBN 978-92-79-81703-8. Available at: <http://ec.europa.eu/social/BlobServlet?docId=19279&langId=en>

FGB (2018b), ESF Performance and Thematic Reports - ESF Contribution to the New Skills Agenda Thematic Report, European Commission, DG EMPL, Publications Office of the European Union, 2018, ISBN: 978-92-79-73300-0. Available at: <http://ec.europa.eu/social/BlobServlet?docId=19224&langId=en>

FGB, Metis, Applica and Ockham (ongoing), Study supporting the evaluation of employment and mobility, European Commission, DG EMPL

Greiss, J, Cantillon, B, Marchal, S & Penne, T (2019) Europe as agent that fills the gaps? The case of FEAD, University of Antwerp HERMAN DELEECK CENTRE FOR SOCIAL POLICY, WORKING PAPER No.19.03 February 2019. Available at: <http://www.centrumvoorsociaalbeleid.be/sites/default/files/CSBWorkingPaper1903.pdf>

Hanushek, E. A., and Woessmann L. (2016). “Knowledge Capital, Growth, and the East Asian Miracle.” Science 351 (6271): 344–345. Available at: <https://science.sciencemag.org/content/351/6271/344/tab-e-letters>

Iseri Europa – IES – Ecorys (2019) Pilot and feasibility study on the sustainability and effectiveness of results for European Social Fund participants using Counterfactual impact evaluations – Final Report, European Commission, DG Employment, Luxembourg: Publications Office of the European Union, 2019. Available at: <https://op.europa.eu/en/publication-detail/-/publication/84cc9eb9-b33d-11e9-9d01-01aa75ed71a1>

Metis, FGB and Panteia (2018), FEAD Mid-Term Evaluation. Final Report, European Commission, DG EMPL. Available at: http://publications.europa.eu/resource/cellar/51421b36-54f8-11e9-a8ed-01aa75ed71a1.0001.01/DOC_1

Metis, FGB, Applica and Ockham (ongoing), Study supporting the evaluation of youth employment, European Commission, DG EMPL

Regulation (EU) No 1303/2013 of the European Parliament and of the Council of 17 December 2013 laying down common provisions on the European Regional Development Fund, the European Social Fund, the Cohesion Fund, the European Agricultural Fund for Rural Development and the European Maritime and Fisheries Fund and laying down general provisions on the European Regional Development Fund, the European Social Fund, the Cohesion Fund and the European Maritime and Fisheries Fund and repealing Council Regulation (EC) No 1083/2006, Available at: <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=celex%3A32013R1303>

Regulation (EU) n. 1304/2013 of the European Parliament and of the Council of 17 December 2013 on the European Social Fund and repealing Council Regulation (EC) No 1081/2006. Available at: <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32013R1304>

Regulation (EU) n. 223/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 11 March 2014 on the Fund for European Aid to the Most Deprived. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32014R0223>

3 Chapter 3: A new framework for the estimation of macroeconomic impacts of the European Social Fund and Youth Employment Initiative: strengths, weaknesses and room for improvements

Abstract

Within the frame of cohesion policy (CP), substantial resources are dedicated to monitoring and evaluation. Whilst much reporting is produced which helps shape current and future policies in the field, evidence on the overall effects (i.e. not only on participants, but of the entire policy from a macroeconomic perspective) remains limited. Through a brief review of the literature examining the effects of the CP and of the European Social Fund (ESF) and Youth Employment Initiative (YEI) this research identifies a knowledge gap which concerns their macroeconomic effects and related disaggregation by regions. In particular, this research highlights how the estimation of macroeconomic effects through the General Equilibrium Model RHOMOLO of the European Commission (EC) in the EC's ex-post, in itinere and ex-ante evaluation of the ESF/YEI is based on assumptions which cannot mirror the actual purpose and composition of the support provided, in lack of better data. Thus, this research proposes a new methodological approach, based on a review of ALMP's micro and macro effects. In particular, it suggests a way to use information on participants, results and effects of Youth Employment (YE) initiatives in the ESF/YEI 14-20 to serve as input for RHOMOLO. In essence, monitoring data on participants and occupational results is used together with the net increase of their probability of being in employment in the medium term - as per available CIEs on YE - to generate evidence on the partial equilibrium effects of the support, which are to be translated, in future research, into macroeconomic impacts through RHOMOLO. In doing so, several methodological and practical obstacles appear, including the need to (i) regionalise ESF/YEI data; (ii) reconstruct (econometrically) background features of the ESF/YEI participants (iii) identify a proxy for the funds' progress; and (iv) identify reliable coefficients on the net occupational effects on ESF/YEI's participants from the existing literature that can reflect their composition and the forms of support they receive. After a preliminary estimation of the regional partial equilibrium effects for the 23 MS (9 of which based on CIEs) benefitting from YE support, the strengths and weaknesses of the proposed method are discussed, together with some recommendations on how to improve data availability and increase accuracy of the estimates in the future.

Introduction and structure of the chapter

Genesis of the research

The idea for this research stemmed from the work being carried out as part of the research team in the frame of the Study supporting the Impact Assessment of Human Capital Investments (FGB et. Al, 2018).³⁶ It is concerned, at large, with the estimation of macroeconomic impacts of the ESF/YEI investments in youth employment support (to be extended also to other thematic objectives under the ESF in future research activity).

³⁶ Carried out for DG EMPL by FGB together with Ockham IPS, Metis and COWI

In particular, what is explored below is the possibility of adopting a novel approach for translating information on the ESF/YEI supported activities into “inputs” or “shocks” for the Joint Research Centre’s (JRC) dynamic spatial computable general equilibrium RHOMOLO model. More specifically, the aim is to showcase whether and how regionalised information on output and results (from the monitoring systems) as well as short-term/micro-economic impacts (from available CIEs studies) can be elaborated and combined so as to feed into the RHOMOLO model. In turn, the latter is meant to shed light on the broader macroeconomic implications of the EU funded support – although this part of the research falls outside the scope of the activities carried out in this PhD project and reported below.

The gaps identified in the literatureThe reason for this research is that although substantial literature exists that assesses the impacts of European Funds, cohesion policy and the European Social Fund, the various existing methodologies come with strengths and weaknesses which suggest that simulating impacts through general equilibrium models can help understand better what works, where and when. It is – to the best of the author’s knowledge – the only existing method that can provide regionalised estimates of impacts and how these are expected to change over time, while focusing on the channels through which the policy affects its target population and socio-economic context.

To name just a few alternative ways of estimating impacts of ESF/YEI interventions and key related issues (better discussed in section 3.1.1 below):

- (i) micro-level counterfactual impact evaluation (CIE) typically assume the absence of spillover effects and interactions among agents, which is not in line with the intervention logic of EU funds; their results are “partial equilibrium” in the sense that only focus on those receiving support ignoring both what happens to the context (non-recipients of the policies, see for instance Crepon et al, 2013) as well as alternative ways of using the intervention’s money. In addition, external validity is typically limited, which reduces aggregability;
- (ii) macro-econometric models tend to (see for instance Dall’Erba & Fang (2017)): (i) focus on cohesion policy as a whole, (ii) be based on growth model à la Solow (not very suitable for human capital development investments), (iii) ignore the differential contribution of specific typologies of interventions (or even specific funds) and (iv) focus on the impact on GDP only; in addition, (v) there is no explanation of the logical link between the intervention and its effects, no regionalised estimates and little room for effects heterogeneity; and
- (iii) recent macro-level counterfactual impact evaluations (based on RDD) face some external validity issues (it is a local average treatment effect for less developed regions close to the cut-off, see for instance) and, again, only focus on average GDP impacts and assume away effects’ heterogeneity (see for instance Pellegrini and Cerqua (2016)).

On the other hand, current initiatives that involve macroeconomic models such as QUEST III or RHOMOLO, which are based on semi-endogenous growth models, tend to overlook the specificity of ESF/YEI investments, also due to lack of suitable data. Just to name one, in the latest simulations carried out through these EU models, the ESF expenditure considered as the

“policy shock”³⁷ for the model was treated in the same way as investment in general education and government expenditure (see for instance Sakkas, 2018). But ESF/YEI support to, e.g., youth employment largely deviates from general education policies, as it chiefly comprises integrated pathways of guidance and counselling, work-based learning, Vocational and Educational training and financial incentives for employers and employees (see for instance Metis et. Al, ongoing or FGB et. Al 2016). The Labour Market effects of these activities are arguably not comparable to those of expenditure in general education. Nonetheless, several simulation exercises have been carried out in the context of the ex-post evaluation of the 2007-2013 ESF and already included in the related reporting (DG REGIO 2016a and 2016b, European Commission 2016,³⁸ DG EMPL 2018³⁹), based on these assumptions, also due to the lack of suitable macro-economic evaluations of the impacts of the ESF (FGB et. Al, 2018b:35).⁴⁰

Motives and key elements of the research

This brief overview suggests why new approaches to providing input into the RHOMOLO model to then mimic the macro-economic shock generated by ESF investment are advisable.

To this end, I developed a theoretical framework which takes stock of the various dynamics affecting access to employment policies identified in the literature (e.g., lock-in effects, deadweight effects, substitution effects, displacement effects and so forth), so that different sources of bias are duly considered when elaborating and combining different data sources.

As explained further below, through a dialogue with the JRC, different research avenues have been considered, ultimately opting for a first experiment that involves changing exogenously some parameters of the model (i.e. participation rates and employment rates) to mimic a labour supply shock generated by the ESF/YEI for youth employment support based on duly elaborated ESF/YEI monitoring data. This, in turn, will affect equilibrium levels and therefore the endogenous parameters of the model.

Such an approach, however, entailed a number of practical and methodological issues, because the ESF/YEI monitoring system does not per se offer suitable information to determine exact changes of exogenous parameters in RHOMOLO. More plainly, ESF/YEI M&E information

³⁷ In order to simulate the effects of a policy which deviates from the baseline scenario, macroeconomic models need some input on how to change some parameters so that they can replicate the “shock” produced on the equilibrium levels by that policy. It might be increased demand, increased prices, lower transport costs etc. This then affects the endogenous mechanisms of the model, thus developing a new sequence of equilibrium levels.

³⁸ European Commission (2016), Ex-post evaluation of the 2007-2013 ESF Programmes, SWD (2016) 452 Final, p. 35 “At macroeconomic level, simulations using the Quest III and RHOMOLO models, 64 on the impact of the HC investments show positive effects in terms of GDP (+0.25 %) and productivity. These effects are much stronger in the EU-12 (+1.5 %), but they are also positive for the EUR-15 (+0.2 %). The gains in GDP are apparent in the medium term, but they are significant and highly persistent in the long term, since they interact positively with the accumulation of physical capital and technology. It should be borne in mind that such broad macro-level developments take time to materialise and are influenced by many other macro-level developments.”

³⁹ the macroeconomic impact (multiplier) is calculated on the basis of the JRC’s RHOMOLO model, p.2

⁴⁰ “[...] the results of the available impact evaluations show positive impacts of ESF investments in certain fields while also highlighting weaknesses in others. The specificity of the programmes and target groups makes it very challenging to make general statements on impacts but the studies looking at the impacts of the ESF on macroeconomic indicators show positive results and a contribution of the ESF to education levels, employment creation, and social inclusion.”

cannot by itself tell us how big is the exogenous shift of the labour force (in terms of more active individuals and more employed individuals in the short/medium term) generated by the ESF/YEI in each region, year and skill level.

Effectively, the main issues have been: (i) regionalising monitoring data for output and results based on proxies; (ii) estimating conditional distributions of participants' labour market status and educational attainment; (iii) identifying a good proxy for the materialisation of the fund's output and outcomes over time; (iv) proxying net effects on participants (i.e. moving from data on results measured after the intervention as per the ESF/YEI system to data on effects generated because of the interventions); and, most importantly (v) identifying a suitable theoretical framework for the analysis.

In essence: the regionalisation (i) was done ascribing output and outcomes pro-quota regional (NUTS2) allocation; (ii) estimation of the distribution of participations and results over time and countries, as well as the prediction of conditional distributions of background features of participants based on marginal distributions, is done econometrically, through multi-level mixed effects models. (iii) the

Next, (iv) the deadweight is calculated either through a proxy from labour market transition data or by borrowing coefficients from the CIEs carried out on the YEI until 2018, notably on the change in probability of being in employment in the medium to long term for participants – by skill level where available. Finally, the estimation of “second order” /general equilibrium effects (i.e., displacement effect, substitution effects, spillovers etc.) is left to the mathematical structure of RHOMOLO (and the wage setting or Phipps curve, which can be used alternatively as explained in section 3.1.2 below) that uses input data on additional people seeking employment or in employment in the medium terms resulting from this research as an exogenous shock to its labour market parameters. Such a shock allows to then mimic the longer-term general equilibrium aggregated impacts.

Although part of the results (notably those in (i), (ii), (iii) and partly (iv)) have already been used by the JRC and DG EMPL in the context of the “Study supporting the 2019 evaluation of Youth Employment” (Metis et. Al, ongoing – henceforth also “youth employment report”), these are yet to be published (likely in late 2020). Hence, they cannot be discussed earlier than that, falling beyond the scope of this project. However, this is already an area where this research has provided a specific contribution to the current policy evaluation practice.

But more importantly, the results of a RHOMOLO simulation which builds upon the whole analytical framework suggested here will only become available after this is carried out by the JRC, and such possibility is currently being discussed as priority was given to the finalisation of the Youth Employment report based on more consolidated approaches given the time constraints.

Thus, the focus of this chapter is on the suitability and potential of the proposed approach, and particularly of the ESF/YEI monitoring data, to serve as an input to the RHOMOLO model and a discussion of the different ways to elaborate it, not on the actual results of the simulations, which cannot be included. Another reason to consider more the quality of the overall framework rather than the actual results from the simulation is that the data used is tentative, with Annual Implementation Reports 2019 being in the validation process through exchanges with the

European Commission. The research will be complemented when the results become available. However, it is not expected that values from the validated AIRs will largely deviate from the ones used below given that a number of checks is already applied to the dataset by the EC before mid-July. Hence, the findings of this research in terms of the suitability of the method remain fully relevant.

In what follows, the limits of the proposed approach will be brought to the fore and discussed individually alongside some solutions envisaged to address them. Much of the experimentation which is described in this section will deal with these issues.

Structure of the chapter

As explained above, the scope of this experimental part of the research is rather broad, and there are different elements which should be considered to contextualise the work.

Hence, the chapter starts by defining the background (section 3.1), both in terms of the literature on macroeconomic effects of cohesion policy and of the ESF/YEI in particular (section 3.1.1), as well as of the functioning of the RHOMOLO model and the way the labour market is described within it (section 3.1.2). This background section is instrumental in understanding what are the main gaps in the literature but also how can RHOMOLO be used in future research to fill such gaps – possibly by leveraging on the analytical framework presented further below. In particular, the discussion of the description of the labour market in RHOMOLO will clarify the structure and specificity of the data needed to simulate the change brought about by a human capital policy through RHOMOLO.

Next, in section 3.2, the framework for the experimentation is discussed. The section starts by describing some alternative ways in which HC investment can be modelled in RHOMOLO, and, through section 3.2.2, it is clarified that given the structure of ESF/YEI monitoring data only some of these are viable at the moment. Finally, in section 3.3.3 the actual framework for the simulation is described, that shows in detail how each of the problems identified in the literature is dealt with under the proposed framework, including limitations. Attention is hence placed to issues discussed in previous chapters hampering the estimation of effects both at the micro (deadweight, lock-in) as well as macro (displacement, substitution etc.) level.

Subsequently, section 3.3 describes the methodological steps carried out to implement such analytical framework, thus addressing the issues of the regional (NUTS2) distribution of ESF/YEI indicators as well as the distribution over time, the econometric estimation of background features of participants, but especially the identification of net effects under different scenarios. In particular, in section 3.3.6 a review of the literature on Counterfactual Impact Evaluations on YEI (or YEI related) programmes is carried out, to identify coefficients on the partial equilibrium effects of the programmes.⁴¹

Finally, the partial equilibrium data estimated is discussed in section 3.4, and the strengths, weaknesses of such data as well as of the overall analytical framework suggested are tackled in section 3.6. In the latter, policy recommendations to improve data collection and

⁴¹ This is, in essence, the change in the probability of being in (self)employment in the medium term (6 to 24 months) for those being treated through ESF/YEI support.

systematisation are included, in line with the approach used in the whole research project, also feeding back into the conclusions of chapter 2.

To keep the text readable, the most detailed or procedural parts of the work, including info on the functioning of RHOMOLO, have been collected in Annex III.

3.1 Background: review of the literature and description of the RHOMOLO model

3.1.1 Brief survey on the literature on the impacts of cohesion policy in general, and on the ESF/YEI and active labour market policies in particular

The literature on the effects of cohesion policy is rather rich. Nevertheless, evidence on its effects is at least mixed (see for example Barca (2009), Hagen and Mohl (2009), Marzinotto (2012), Pienkowski and Berkowitz (2015), Crescenzi and Giua (2017)) and, according to the surveys on its effect tends to depend on model specifications (Darvas, Mazza and Midoes (2019), Pellegrini et. al (2013)) and on the data used (regional vs national, expenditure or allocation data etc).

One could roughly characterise the different research streams according to the below criteria:

- micro (project) and meso (programme) level vs macro (region/country) level;
- entire cohesion policy vs human capital investment only (mostly but not only ESF – at the extreme also disaggregation by typology of intervention);
- effects on GDP vs effects on other variables too (especially employment); and
- aggregated vs regionally or target-group disaggregated results.

Each of these elements have a bearing on the policy relevance of the approaches. Different methodological approaches have been devised that produce suitable answers to the research questions investigated, which, bearing in mind the above criteria can be broadly grouped as follows:

- **Counterfactual impact evaluation on projects/programmes:**
 - **Stylised features:** spill-over effects and other general equilibrium effects are typically assumed away (SUTVA hypothesis, so no interaction between the participants that are treated by the intervention and the environment around them that affects equilibrium level. This assumption is reasonable for sufficiently small projects). Various methods applied, also in combination, Regression Discontinuity Design, Difference in Differences, Statistical Matching, Instrumental Variables etc.
 - **Categorisation criteria:** micro/meso level, focus on specific ESF interventions (also different typologies of intervention) and effects on occupational chances for participants, disaggregated results.
- **Macro-econometric evaluation of programmes:**
 - **Stylised features:** a Beta-convergence coefficient is typically estimated within a neoclassical framework à la Solow, to see if less developed regions effectively show convergence. This helps comparing different papers and aggregate them, but is

potentially affected by some issues, which might include parameter heterogeneity, outliers, omitted variables, model uncertainty, measurement error and especially endogeneity (sometimes the very reason for a policy is endogenous to its outcome variable) according to Rodrik (2012). It is also based on a framework that doesn't consider endogenous or semi endogenous growth. Various specifications have been applied that seek to treat endogeneity and look into omitted variables, as clarified also by Dall'Erba & Fang (2017). This has reportedly led to increasing reliability of the estimates, whilst this learning effect is typically associated with smaller convergence coefficients. At any rate, the convergence coefficient is a growth parameter, whilst employment levels are often overlooked. There's also a tendency to use lagged allocation values instead of actual regionalised payments. The different compositions of support, heterogeneity of effects are not considered either.

Some attempts have been carried out through a comparable macro-econometric approach to focus more specifically on employment effects (see e.g. Mohl and Hagen, 2011), but similar issues might apply there.

- **Categorisation criteria:** macro level, focus on cohesion policy as a whole, effects on GDP and employment levels in some cases, aggregated.
- **Macro-econometric “counterfactual” methods:**
 - **Stylised features:** that's the case of the RDD model, that was applied leveraging on the fact that eligibility criteria for cohesion funds act as a threshold around which rather similar regions might be distributed smoothly, so that in the proximity of the threshold a reliable counterfactual group is identified (local random assignment around the cut-off, as per Lee (2008)). This approach was used for instance in Becker et al. 2010; Becker, Egger, & Von Ehrlich, 2013; Pellegrini, Terribile, Tarola, Muccigrosso, & Busillo (2013) and more recently in Pellegrini and Cerqua (2016). Research in this strand typically controls for co-variates of regions in the control group (but it remains hard to account for the strong heterogeneity across regions) to estimate the average GDP effect, also taking into account the intensity of funding. Research in this strand also tends to assume away effect heterogeneity (or checks for that only through dummies such as EU-15 or and across programming periods) which is somewhat inconsistent with the idea of place-based policies and the traditional general equilibrium approaches. Another limitation is that such an approach does not look into the composition of the support, i.e. it does not distinguish among forms of intervention, which may clearly have very different effects. In addition, what is estimated is the local average treatment effect (LATE), thus internal validity is stronger but external validity might be reduced.
 - **Categorisation criteria:** macro level, focus on cohesion policy as a whole, effects on GDP levels, aggregated.
- **Macro-economic models:**
 - **Stylised features:** these are typically large and complex general equilibrium models, such as the QUEST model (Varga and in t'Veld 2009) and the RHOMOLO

model (Lecca et al., 2018). These models, the former a dynamic stochastic general equilibrium and the latter a computable general equilibrium, are based on semi endogenous growth approaches (thus the R&D element is key within the design and so is the share of people who are high skilled or employed in the research sector). Effects of the funds are estimated within a controlled environment in which calibration of the model parameters is done based on real data, there are transportation costs and spill-overs, a number of economic sectors, input-output tables that account for indirect effects and different agents with their utility functions which needs to be maximised simultaneously to bring the system in equilibrium. Agents can be either forward (QUEST) or backward looking (RHOMOLO), and their unit of observation is the MS (QUEST) or NUTS2 (RHOMOLO) level. As noted by Pellegrini et. Al (2013), they also typically find positive results of the policies. The focus for these models is not only the effect on GDP, employment or several other macro-economic variables, but also on the channels through which the policy produces its effects, including the temporal component (short vs long term effects) with interaction between demand and supply factors and labour and capital over time, spillovers, trade flows etc. Importantly, these models do not check experimentally how much a given index has changed because of the intervention but rather mimic within a controlled environment what happens to the equilibrium values if a range of parameters shifted as a result of the policy intervention. This means that the way in which the policy shock is translated into the model has a pivotal role in determining reliability of the estimates. In particular, recent estimates of ESF/YEI impacts (see for instance DG REGIO and Joint Research Centre Seville (2016) and Sakkas, Conte and Salotti, 2018)) also included in the EC's Impact Assessment of Human Capital Investments, consider what would happen if that money was spent in two main ways, notably General education and public expenditure. Demand and supply effects have been simulated through multiplying, e.g., investment in education by some coefficients on higher productivity per euro spent borrowed from the literature. This significantly limits both the reliability of the estimates as well as the adherence to the policy whose effects are tested.

- **Categorisation criteria:** macro level, focus either on whole cohesion policy or on specific funds, effects on a range of macro-economic variables, disaggregated results by region.

One can see that the methods applied to examine the effects of cohesion policy in general, and of the ESF in particular, vary considerably. It is also chiefly a matter of which research question is being investigated each time.

If the attention goes to the overall effect of a rather large policy such as the ESF/YEI, then it is reasonable to argue that macro-economic models should be at a minimum involved in the estimation, as it becomes important to emphasise the interplay between the policy and the environment in which it unfolds. To name just one, in a region with a shortage of skilled labour, rigid wages, and elastic labour demand, a retraining program may have a large net employment

effect,⁴² whereas a similar program in a different region may have a very small net impact, according to standard economic theory.

To sum up the issues that might affect micro-econometric (counterfactual) impact evaluation in the domain on which this research focuses (activation measures for youth employment) one could say that:

- it is a “partial equilibrium” estimation, that is, it produces estimates of what would have happened to a given individual if they had not received support, but it assumes away the interplay between the intervention and the environment in which it unfolds, thus (see for example Escudero, 2014):
 - it does not look at changes to the wage-setting structure following provision of support (no impacts on the reservation wages, downward effect on wage setting due to better matching, etc);
 - it does not look at changes to the demand of labour following the provision of support (no substitution effects due to higher productivity of workers, no displacement effects of those trained vs those who have not received training, no impact on productivity, and thus, as a result, on the economy’s clearing levels);
 - it cannot address spillover effects; it actually assumes that spillover effects do not exist, while, e.g., input-output tables focus on the opposite (i.e. indirect and induced effects which are sometimes even larger than the direct effects);
- it tells little about alternative ways to employ the money which funded the intervention. In other words, it tends to focus on benefits rather than cost-opportunity
- it has strong internal validity (especially with RCT⁴³ or RDD) but somewhat lower external validity as it can be affected by contextual conditions to a large extent

Yet, the different macro-economic/econometric approaches are affected by different sources of bias and even in the case of the possibly more reliable counterfactual RDD macro-econometric designs, some important limitations – such as effect heterogeneity, lack of a clear logic link between the (different forms of) policies and the effects and no regional estimates – mean that different approaches should be explored.

In essence, one could argue that, somewhat trivially, micro-economic evaluations are more reliable but not suitable for a reasoning at a broader scale and macro-econometric estimates are too broad-cut, so they fail to consider the micro-level nature (and especially variation) in the composition of support and the target group to which this is directed.

As per Pellegrini et al. (2013) there seems to be a strong need to better integrate the micro and macro level with model specifications which can, at once, reconcile the need for disaggregation and adherence to the variety of investment forms even within the same umbrella funding initiative and the need to look at the big picture through macroeconomic lenses which can better

⁴² This is due to the fact that if the wages have sluggish adjustments to the new increased productivity of the labour force and there is an elastic labour demand, much more upskilled unemployment will become attractive for employers and thus employable, all the more so if skilled labour is scarce in the area.

⁴³ Randomized control trials

inform on whether, all things considered, the effects on the economic context are positive or otherwise.

In general, better evidence on the impact of the ESF was highlighted also by the Regulatory Scrutiny Board⁴⁴ (hereafter RSB) in its opinion provided to DG EMPL about the ex-post 2007-2013 ESF evaluation, where further efforts to assess the effectiveness of the fund were requested, including by making broader use of Counterfactual Impact Evaluations. The RSB draw attention to the fact that “[the report] should indicate to what extent the results represent a realisation of the objectives and explain them with regard to the deviation from a clear and explicit baseline that has to be developed”. It continued by inquiring into whether the ESF results could be attributable to the ESF or to any other development of the socio-economic context outside of the ESF.⁴⁵

In this line, DG EMPL published in 2017 a call for proposals which aimed at fostering the use of CIE in the EU. According to the tasks specifications⁴⁶ “*the current granularity of evaluations at national and regional levels generates the difficulty in synthesising ESF impacts at EU level. Therefore there is a need to identify and solve the problems related to data that prevent from fully exploiting the increased potential for conducting more CIE evaluations at Member State level and also analyse the feasibility of using CIE to reach valid conclusions about effectiveness and sustainability of ESF at country and EU level.*” This call has proved necessary as, according to Evaluation helpdesk data as of late September 2017⁴⁷ only 46 evaluations had been carried out that could be classified as CIE, of which just 1 on the ESF in 2014-2020 in LV “*Evaluation of unemployed profiling systems for job placement*” and 1 on the YEI in IT. The situation has henceforth improved especially within the framework of the YEI, especially with the regulation’s obligation of carrying out an impact evaluation of the YEI by the end of 2018. This has meant some 10 CIEs studies being produced by a range of MS, as better described in section 3.3.6 below.

However, and even more importantly for the context of this research project, even where the granularity of CIE methods should increase substantially in the next years – and this is to be expected - one should be wary of the other limitations that are linked to the aggregation of results at the national and European level.

In this train of thoughts, attention is paid in this research to the possibility of improving on some of the weaknesses of the macro-economic general equilibrium approach, which currently employs some unrealistic assumptions in terms of input data, but has some strong points when

⁴⁴ Ref. Ares(2016)5694203 - 30/09/2016, DG EMPL – Ex-post evaluation of the 2007-2013 European Social Fund programmes, Draft version of 2 September 2016 available at https://ec.europa.eu/info/sites/info/files/rsb-opinion_en.pdf

⁴⁵ “to what extent improved success indicators (reference value to be provided) can be attributed to the action of the ESF, how important was the contribution from other actions/funds/programmes and to what extent were they influenced by other developments in the EU outside of the ESF (e.g. in how far is the reported employment of 9.2 million participants attributable to the ESF intervention only?)”

⁴⁶ Request for services under the Multiple Framework Contracts VT/2016/027 for the provision of services related to the implementation of Better Regulation Guidelines - Pilot and feasibility study on the sustainability and effectiveness of results for ESF participants using CIEs. Identification n.7 - VT/2017/045

⁴⁷ Based on an extraction provided to FGB in the frame of the ESF Performance and Thematic Reports, updated as of late September 2017 (FGB, 2017)

it comes to producing results that are regionalised, deal with effect heterogeneity and help consider the link between the intervention and its effects.

Thus, the key issue remains how to link the micro and the macro level; in other words, how to make sure that (i) the lesson drawn from the counterfactual is incorporated in the input data (efficiency/effectiveness) for the macroeconomic models whilst (ii) a high level of detail in terms reflecting the actual outcomes and microlevel/short term impacts of the ESF/YEI is kept as the data is inputted in General Equilibrium models, avoiding too far-stretched assumptions on how money is spent.

Particularly relevant to the work carried out in liaison with the JRC were a few issues linked to the features of the ESF/ YEI monitoring system, and particularly:

- the lack of precise data on output and results at the NUTS2 level;
- the lack of conditional distributions of output and results (i.e. not just how many low skilled have found a job, but also how many of these were also unemployed, or inactive);
- the lack of a time profile for the materialisation of output/ results;
- the lack of knowledge on input data (what exactly have the ESF / YEI funded, and where?); and
- the uncertainty about the fact that results measured by the monitoring system can actually be ascribed to the intervention, or might just be deadweight.

These issues will be addressed after having briefly presented the stylised structure of the RHOMOLO model.

3.1.2 Key features of RHOMOLO, JRC’s Regional HOListic MOdel

3.1.2.1 The JRC’s RHOMOLO model⁴⁸ - main features

The RHOMOLO model is a spatial dynamic computable general equilibrium model which was devised by DG REGIO and the JRC and employed over recent years in the context of both ex post and ex ante evaluation of cohesion policy. A more extended description of general equilibrium models and of the RHOMOLO structure is developed in Annex III. Here, the key elements are recalled, to keep the chapter readable.

Although being a CGE, RHOMOLO may rely for its calibration on an extended timeframe upon another EC model, namely QUEST (Ratto et. Al, 2008) a DSGE model employed by DG REGIO, with only one sector and 28 MS. It can thus derive part of its parameters from a DSGE model to then further the analysis at the regional level, which answers to the fundamental critique of Barca (2009) on the importance of the local dimension of the cohesion policy. In addition, it is dynamic, in that it includes expectations (myopic) and both financial markets and labour markets are described through functions that allow for “sluggish adjustments” to equilibrium levels. Yet it cannot, at this stage, incorporate fully-fledged dynamic behaviour with forward-looking agents, due to the recursive dynamics of the model which link changes to “*the continuous accumulation of human, knowledge and physical capital, which extend past accumulation rules to the future.*” (Di Comite et Al., 2015:25). It thus ties - in some simulation

⁴⁸ Based on the various documents published by the JRC (see references), and particularly Mercenier et. Al (2016)

exercises - its regional results so that they are consistent with national results computed through QUEST, with fully rational economic agents. A new version of the model (V3) is currently under finalization and is expected to include also forward-looking agents.

In its currently used V3 version (Lecca et al, 2018), RHOMOLO covers 267 NUTS2⁴⁹ regions in the EU-28, whose economies are broken down into ten NACE Rev. 2 sectors.⁵⁰ Given it is essentially a CGE model, it is based on the EU interregional Social Accounting Matrix⁵¹ derived from Thissen et al., (2019), with data on firms derived from the European Structural business statistics (Eurostat, 2017) and structural and behavioural parameters borrowed from the literature.

As factors are mobile, can be traded and move across the EU (at a cost), the regional dimension as a determinant of economic performance is key as transport cost “*triggers the aggregation or dispersion forces which, through the availability of cheaper intermediates, access to consumers and degree of competition*”. In other words, this means that the spatial dimension affects the performance and competitive environment in which firms operate (cf. “spatial effects” below).

In annex III, a range of features of the model are summarised (but a full discussion can be found in Lecca et al, 2018) but here below the main equation describing the functioning of labour market in RHOMOLO is described as this is key to understand the way in which HC investment was modelled in the past.

- Labour market: in RHOMOLO v3, there are three different options for the labour market, namely a static wage curve,⁵² a dynamic wage curve and a Phillips curve. These can be used alternatively shifting the parameters of the baseline equation for the wage curve below

$$(1) \quad rw_{r,e,t} = \alpha_e + \alpha rw_{r,e,t-1} - \beta u_{r,e,t} + \zeta \Delta p_{r,t} - \theta \Delta u_{r,e,t} + \omega \Gamma_t$$

Where

- “ rw ” is the real wage, negatively correlated with unemployment levels and positively correlated with the price of output and the productivity trend
- “ $u_{e,t}$ ” is the unemployment rate
- “ $\Delta p_{r,t}$ ” is the price of output
- “ Γ_t ” is the productivity trend
- “ e ” is the differentiation by skill level,
- “ r ” is the region

⁴⁹ Nomenclature of territorial units for statistics - the version implemented in RHOMOLO is the NUTS 2006. Tables are normally used to update the allocations to the more recent NUTS2016 classification.

⁵⁰ NACE rev. 2: A, B_E, C, F, G-I, J, K-L, M-N, O-Q, R-U. See specificities in Lecca et. Al (2018)

⁵¹ A social accounting matrix (SAM) can be defined as an organized matrix representation of all transactions and transfers between different production activities, factors of production, and institutions (households, corporate sector, and government) within the economy and with respect to the rest of the world

⁵² A wage curve is a reduced-form representation of structural models of imperfect labour markets, such as union wage bargaining models, efficiency wage models, or matching models labour markets

- “a” is an exogenously fixed productivity parameter

This means that the level of unemployment is endogenized and made dynamic through a wage curve. In other words, unemployment levels are linked to real wages but also to changes in output prices and past real wages. If the parameters α , ζ and θ are set to zero, then the curve becomes static, and real wage only depends upon unemployment rate. If then α is set to 1 the equation becomes a Phillip curve, with changes between rw in $t=0$ and rw in $t= -1$ are dependent from levels of unemployment in $t=0$ and unemployment and employment in absence of changes to exogenous parameters (i.e. the volume and composition of the labour force) come back to the steady state. Importantly, the β coefficient, which is basically the slope of the wage curve, is borrowed from the literature⁵³ and is not differentiated across regions. The same applies to the dynamic parameters. Labour demand is determined by the firms, by region, sector and skill level, and, at clearing conditions, the situation is as per the equation below:

$$(2) \quad \sum_j N_{r,j} ld_{r,j,e} = (1 - u_{r,e}) L_{r,e}$$

Where the labour demand (ld , by region economic sector and skill level) for all firms (N , by region and economic sector) equals the labour force (by region and skill level, considering the unemployment rate). Importantly, labour supply is determined exogenously and assumed not to vary (no natural population change). Nevertheless, there is a module for labour mobility originally developed in Persyn et. al (2014), “*where workers' migration is governed by expected differences in the real incomes, and is also dependent on the probability to be employed in a given region.*” (Lecca et. Al 2018: 16). This means that the labour supply in $t+1$ equals the original labour supply plus the net migration rate ($m_{r,e}$), which is, in turn, contingent on a vector of features of the regional socio-economic context such as unemployment, real wage, distance etc.

Thus, the labour market supply evolves as in

$$(3) \quad L_{r,e,t} = L_{r,e,t-1}(1 + m_{r,e})$$

So, the model allows controlling for the effect of policies also in terms of how these affect skills mobility.

In addition to the structure of the labour market, it is also worth recalling here that in RHOMOLO there are spatial effects which are due to iceberg type transport costs and, especially, myopic expectations of the agents. In fact, “*RHOMOLO assumes that the behaviour of inter-temporally optimising agents depends only on the current and past states of the economy, not on the inter-temporal optimisation.*”. It is in this sense myopic, with backward

⁵³ The value of β is set at 0.1, based on Nikjamp and Poot (2005) and typical of a long-run wage curve. The dynamic parameters α , ζ and θ are set to 0.1, 0.25 and 0.03 respectively, based on Montuenga-Gómez, and Ramos-Parreño, 2005 and Nymoene and Rødseth, 2003.

looking agents. Future versions of the code will allow for rational (forward looking) expectations (Lecca et al., 2018).

3.1.2.2 How has investment in HC been modelled in RHOMOLO in the past?

The structure of RHOMOLO and particularly its flexibility when it comes to mimicking changes in the labour market means that it can be a rather helpful tool to understand both the extent of impacts of HC investments as well as how these materialise, i.e. through which channels and building upon which economic theory. Spatial effects (both in terms of firms and workers decisions) are also a key strength if one thinks of Barca (2009) and the need for place-based policies. Nevertheless, some features of the model mean that it is quite complex to translate investment in HC into policy shocks that can be “digested” by RHOMOLO. Whilst the actual approach to the simulation is discussed in section 3.2 below, here one should recall that:

- Decisions of participation in the labour market are not yet modelled within RHOMOLO v3. What this means is that LM participation (and thus shares of inactive, unemployed and employed) is exogenous. The reason for this is mainly that participation to education and training as well as the labour market are inherently forward-looking decisions, whereas agents in RHOMOLO are backward looking due to the computational issues with such a huge model (hundreds of thousand equations within the model given 267 regions, 10 NACE sectors etc.).
- There is no R&D sector as yet. In RHOMOLO v2, a national R&D sector was foreseen to which a share of high skill workers was dedicated. Thus, long-term growth was also the result of, e.g., people becoming high skilled and contributing to the R&D sector within a semi-endogenous growth model. In the current v3 model, there are constant returns to scale for labour and different sources of capital investment (public and private), so the underlying macro-economic theory remains similar, but there is no clear link between changes in TFP and better supply of high skill workers within a region or sector. The productivity parameter is also a shift and share parameter, determined exogenously.

As a result, the impact of the ESF was so far analysed by:

- Shifting exogenously the labour productivity parameter for regions who benefitted from the ESF support, assuming that:
 - ESF/YEI allocation in TO8 and TO10 was equivalent to a general education policy and estimating how many additional school years could be purchased with ESF/YEI money.
 - Each additional year of schooling increases productivity of the whole economy by a given factor (borrowing estimates from the literature), adjusted at the MS level according to indexes such as quality of administration, PIAAC values etc. to better match the different effectiveness of Education and Training systems.
 - The entire population was affected by the policy and increased their productivity, or that only average levels of productivity matter to firms (whereas increased in productivity, if any, benefit participants only)

Let's briefly recall the wage curve described in equation (1) above. The previous mechanism basically operated through exogenously shifting the a_e factor of productivity, for each "e" skill level and "r" region based on values borrowed from the literature of how much **one additional year of schooling** (i) costs and (ii) increases productivity.

- Translating TO9 expenditure in general government spent, thus assuming:
 - No impact on participant's occupational status (i.e. neither from inactivity to unemployment, nor from inactivity or unemployment into employment)
 - No effect on participants' productivity
 - Mainly just demand-side short term effects due to higher government expenditure

In terms of the model's equations this basically means a shift in aggregated demand, which in turns triggers a reaction in the endogenous parameters of the model describing the different markets and sectors.

Available studies on the ESF/YEI (see for example FGB et. Al, 2016; Metis et. Al, ongoing; FGB et. Al, ongoing), however, clarify that the nature of the related investment cannot be subsumed under the above typologies of interventions.

In particular:

- for TO8, measures chiefly concern measure to improve matching (guidance, counselling) access to employment (job or start-up subsidies, work-based learning like internships, traineeships and apprenticeships) vocational education and training (for occupational skills) and other job-mobility or institutional capacity (reforming PES) measures;
- for TO9, measures aim at active inclusion, so especially bringing people closer to the labour market, i.e. having them starting to actively seek employment or find one, including by removing obstacles for their social inclusion; and
- within TO10, in addition to measures preventing early-school leaving (which can actually be seen as providing people with additional years of schooling) much work is carried out that deals with occupational skills within life-long learning approaches and with improving the labour market relevance of workers' skills. So again, general education seems to be too far stretched to mimic this.

In view of the above, it's easy to argue that, for TO8 and TO10, shifting a parameter of labour productivity based on the literature which is based on the econometric assessment of a rather different typology of intervention (general education vs access to employment measures) might be misleading, and so is assuming away the very purpose of TO9 investment, which is active inclusion and cannot be treated in the same way of non-productive government expenditure.

Based on this understanding of the RHOMOLO v3 model and of the current assumptions used in the simulations carried out with it so far, in the next section a framework is developed that: (i) shows which alternative approaches might be used to reproduce a policy shock within RHOMOLO that is more suitable to the ESF/YEI reality and (ii) presents the specific methodological choices made within the ensuing research.

3.2 The analytical framework for this simulation exercise

In the light of the RHOMOLO structure as quickly presented above and of joint discussions with the JRC, a range of possible modelling strategies have been examined.

This review, together with an assessment of what information (and when) can be drawn from the ESF/YEI monitoring system has led to the definition of an operational strategy which aims to test a novel way of translating investments in youth employment measures within the ESF/YEI into a shock that can be handled by the RHOMOLO model. This investment strictly corresponds to money allocated to Investment Priority 8.ii – “Sustainable integration into the labour market of young people[.]”.

This experiment, with just some adjustments, could be replicated in the future for investments that go beyond Investment Priority 8.ii, notably those concerning access to employment at large (TO8) and social inclusion (TO9) to start with.

But the starting point is to understand which alternative ways can be used to shock the RHOMOLO model, which are described here below.

3.2.1 How to model HC investments in RHOMOLO

In essence, there are three broad avenues through which ESF/YEI investment can be mirrored into RHOMOLO. These are listed below in order of increasing use of the endogenous mechanisms of the model.

As it can be expected, adjusting some exogenous “shift and share” parameters is the quickest way to compare the baseline scenario with a policy shock, i.e. a scenario in which parameters are amended so as to reflect permanent changes to the environment under observation (e.g. increased productivity, changes to the labour force) as a result of the intervention.

On the other hand, fully endogenous shocks would allow a more faithful and nuanced representation of how investment (the policy shock) affects relationships between agents and how, in turn, their behaviour means changes to the market clearing levels, growth etc.

Along these lines, the three main research avenues have been identified in liaison with the JRC for RHOMOLO (see also Sakkas, 2018), notably:

- avenue 1: Change some calibration parameters to force the model to replicate some measured programme outcomes for an endogenous variable (e.g. employment) by, e.g. changing values of the workforce as a result of the intervention, other things being equal;
- avenue 2: To estimate econometrically how the intervention might alter some specific exogenous parameters which can be freely set in the model, such as the labour productivity; and
- avenue 3: To specifically model targeted areas as an economic activity, e.g. to set up a production structure for education and training.

The three research avenues, and the details of how to best exploit the data we have, were discussed with the JRC and greatly benefitted from their inputs, with the following results:

- avenue 1:

- alternative a: to estimate econometrically the employment effects at the regional level. This would be along the lines of the RDD approach seen in Pellegrini et. Al (2013), but also in Bondonio (2016) etc. The problems with this approach are the following:
 - low progress of the implementation, which means that data on the employment effects may be still unreliable and employment impacts still in the process of materialising;
 - overlap with previous programming period and other policies in the same field: difficult to understand what is due to the ESF/YEI and what is due to a mix of other policies, including national policies and previous funding. So, it might be hard to identify a relationship between allocations and their effects;
 - no effect heterogeneity, only average values across the EU; and
 - no consideration of labour migration/crowding out effects.
- alternative b: to estimate the effects on the treated, through either remodelling (i.e. creating a fourth category of individuals with separate features) or adding new people to the labour force (or to the employed) as a result of the intervention. Main problems with this approach are:
 - remodelling can be very complex as it requires knowledge on a range of features such as the different productivity levels of the groups of treated and non-treated, on their elasticity of substitution and how these are affected by the intervention. This is not available data; and
 - adding people to the labour force is feasible (assuming it is known how many participated in the interventions and who, among those, have changed in a sustainable manner their labour market status as a result of the intervention), but the estimation of displacement and substitution effects is left with RHOMOLO. The displacement size will depend on the shape of labour demand, the wage curve, and the shape of labour supply etc. The wage (or Phillip) curve is arguably the main weakness in the sense that it is not regionalised in RHOMOLO at the moment.
- avenue 2:
 - alternative a: to increase productivity, either of the entire workforce (as it was done so far) or of participants only. Main problems:
 - need of estimates of how much the interventions affect productivity, which are missing for ALMP (no info on wages of the treated, which also means no counterfactual studies on it); and
 - in the case of increasing productivity of the total workforce, it is assumed that the productivity of the total workforce is increased (including the non-treated), or that perhaps only the average matters to firms.
 - alternative b: to consider how ALMP affect the wage curve, through reducing frictional unemployment and increasing the efficiency of matching. Main problem:
 - Need of estimates of how different forms of intervention affect the wage curve and reduce its rigidities. Ambitious empirical project which goes beyond the scope of this research.

- alternative c: to consider how (some) ALMP affect the cost of participation, and how this affect the behaviour of the population. Main problems:
 - not all ALMP affect this, or not all of them in the same measure;
 - participation to the LM is not currently modelled within RHOMOLO v3, so this would be for future reference anyways; and
 - when the module is finished, estimates are needed of how measures have affected the cost of participating. This can be straightforward for some typologies of ESF measures (e.g. childcare) but harder elsewhere.
- avenue 3: to set up a production structure for education and training. Policies would be then modelled as a subsidy to this activity. Main problems:
 - this module does not exist at present, nor it is foreseen in the near future; and
 - the reason for the lack of such a model is that education and training choices are inherently forward looking decisions. In RHOMOLO, agents are backward looking due to the computational structure of the model which needs to solve simultaneously an enormous range of equations, which is a trade-off for the regionalised approach.

In consideration of the above, it was decided to explore, within the context of this research, the avenues 1(a) and 1(b), namely changing employment levels and the number of people in the labour force based on elaborated ESF/YEI data. This is in line with what envisaged by Sakkas (2018) in the latest simulations carried out through RHOMOLO in the context of the Impact Assessment for the ESF. In the related paper, he suggested the need for precise data on the effects of ESF/YEI operations funded within the scope of thematic objective 8 – which is exactly the case of youth employment support under Investment Priority 8.ii, object of this research.

It will be for RHOMOLO then, in future research, to calculate how changing these parameters affects other endogenous variables of the model so that second-order effects (i.e. impact on aggregate output, labour demand, wages, migration flows etc.) are estimated. The next section will describe more in detail the approach followed to obtain this data from the current ESF/YEI monitoring and evaluation system, together with the underlying assumptions.

3.2.2 ESF monitoring data and main issues with a view to the simulation

In order to pursue Avenues 1(a) and 1(b), there is a need of very detailed information coming from the ESF/YEI monitoring system. This section clarifies whether and how such information can be retrieved.

Whilst a detailed description of the ESF/YEI monitoring and evaluation data is not the purpose of this section (better information can be found in chapter 2 of this research project), it is worth briefly recalling the main issues related to its suitability to be used as input to estimate macro-economic impacts that this specific part of the work aims at addressing, notably:

- **data granularity**, that is, whether ESF/YEI data is sufficiently detailed and disaggregated to provide input into the RHOMOLO model. As explained in section 3.1.2.2, labour market data should be disaggregated, as a minimum, by labour market status, level of skill and NUTS2 region. With respect to this, the main issues are as follows:

- **disaggregation at NUTS2 level:** in the frame of the ESF/YEI data is collected at the level of ESF/YEI Operational Programmes. These can be defined at the NUTS0, 1, 2 or 3 level. This implies that:
 - overall allocations (programming) can be defined at a higher level than NUTS2; and
 - financial progress of expenditure (implementation) can defined at a higher than NUTS 2.

In order to address this, some proxies need to be identified that can help ascribe not only expenditure, but also the related participations, results and effects at the NUTS2 level;

- **disaggregation by participants' background characteristics:** although detailed micro-data on participants is collected by Managing Authorities, the ESF monitoring system makes available only rather aggregated info on participants, especially with a view to limiting administrative burden for the collation of micro-data and avoiding privacy issues. This means, however, that:
 - for output data (number of participants) only marginal distributions are available and not conditional (or cross) distributions exist by age, labour market status and educational attainment. In other words, one could know how many participants to a given programme are low skilled OR inactive within a specific strand of a programme (Investment priority, by category of region) but not low skilled AND inactive;
 - for result data (number of participants improving their educational or labour market status) only aggregated data is available (in other words, one could know how many participants have found employment but not how many of them were previously inactive or unemployed, or low skilled etc.); and
 - data from existing Counterfactual Impact evaluations is not standardised, it typically focuses on average effects and only in some cases it gives some information which can help understand how to distinguish effects by typology of participant (heterogeneity of effects not always tackled). .
- **distribution of participations and results over time:** the ESF/YEI monitoring system tells us (with some lag and approximation)⁵⁴ how many participations the funds have supported as well as the occupational and educational results thereof (immediately after the end of the operation and at six months). However, in order to run a proper simulation for, i.e., the current programming period, one should also get to know how much more is to be expected once the entire financial envelope is put to use. Thus, a measure is needed that tells us how far we have got into the implementation. However, existing financial as well as physical indicators can be affected by bias in this respect.
- **disaggregation by type of activity and results pursued:** one key question concerning the effects of a given intervention is “how to measure success”, which

⁵⁴ According to the ESF regulation, data on output and results can be reported in relation to either partly or fully implemented operations. In practice, this means that some ongoing operations which have been funded have no participations or results reported. In addition, participations are measured at the entry to the operation, whilst results only at the exit, which can be a few months later. So yearly information on the ESF/YEI output/results can be affected by time lags.

implies having some clear ideas on what outcome variables should be affected by a given form of support. Inevitably, different interventions will produce not only different results with respect to a given set of outcome variables, but may also affect completely different variables not previously considered as relevant. However, given the degree to which ESF interventions are diversified, a reasoning at the macro level can only be done by looking at some stylised classes of results, and notably those where data is available. In the specific case of TO8 operations (which include Youth Employment ones), the overarching goal of TO8 being employment, one could argue that changes of Labour Market status are the most notable result, but, in fact, changes in individual productivity, in the cost of participating to the labour market or more systemic changes (e.g. investments in PES to reduce search frictions) can not only produce different results but also potentially call for different modelling approaches.

- **data suitability to measure effects**, that is, whether we can safely assume that the data is correct and capture only what happens because of the ESF support, which includes:
 - **causality and attribution issues:** it is widely acknowledged that the results tracked by the ESF/YEI monitoring system (e.g. participants finding employment upon completing the intervention) may not be generated by the intervention itself (issue of causality/attribution) but by spontaneous dynamics;
 - **sustainability:** information within the ESF/YEI is available that reports how many have, i.e., found employment upon the end of the intervention and also (estimated) at 6 months of exiting from support. But the chosen way of shifting parameters in RHOMOLO means trying to understand how many of the participants have stably changed employment status, so that they can be added to the workforce or to the equilibrium levels of employment;
 - **lock-in effects:** when results are measured upon the exit from the operation effects might, especially for some typologies of interventions such as work-based learning or VET, materialise at a later stage. On the other hand, spontaneous dynamics might increasingly affect results when these are measured at 6 months of exiting the intervention, making it harder to disentangle their gross and net effects; and
 - **general equilibrium effects:** no data is available within the ESF/YEI monitoring system that can provide standardised information on, e.g., displacement, substitution, migration or other spillover effects.
- **data comprehensiveness**, that is, the data available faithfully represents the entire range of benefits that are generated by the ESF/YEI support. Unfortunately, at this stage data on intermediate outcomes (increase in awareness, self-esteem) is not available nor can it be reconstructed. Thus, in principle only occupational results are tracked, which is however quite consistent with the purpose of youth employment spend. Nevertheless, as occupational results are only measured for the inactive and unemployed, the fraction of expenditure which is dedicated to the participants already employed needs to be translated into something else, added into the RHOMOLO model through a different channel. Also CIE tends to be quite limited in this respect, as they only measure the probability of being in employment at different points in time without, e.g., looking into other potential benefits related to individual productivity, active citizenship, health, crime rates etc.

Against this backdrop, a framework is described in the next section that illustrates how each issue has been addressed in the proposed modelling approach.

3.2.3 Analytical framework and methods

The table below, summarises the solutions envisaged in order to address the issues recalled above. The table focuses on the object of this exercise, that is, support provided for Youth Employment through money invested under the thematic objective 8, investment priority 8.ii “sustainable integration of young people in the labour market” leveraging upon ESF and YEI financial support. The choice of a more circumscribed scope has to do with higher homogeneity of the target group and typology of support provided, which ultimately aims at bringing young people closer to or into employment. Another reason for this focus is that, due to programming requirements, there is a broader availability of counterfactual impact evaluations within the frame of youth employment initiatives.

Though the rationale of this approach can be replicated elsewhere under the ESF (especially under TO8) the framework might need fine tuning especially where the aim of the interventions is not explicitly (or directly) bringing people closer to or into employment.

Table 10 – Overview of the proposed methodological approach

Broad Issue	Specific Issue	Proposed solution in the modelling approach
Data Granularity	Disaggregation at the NUST2 level	Information either on financial progress or financial allocation might be available (or reconstructed) at the NUTS2 level, as it is contained in Annual Implementation Report (table 7, progress of costs and expenditure) and Operational Programmes. This is considered as a proxy for the regional distribution of output and results too, assuming that within the same MS and category of region progress of output and results is distributed evenly across regions covered by the same OP.
	Disaggregation by participants' background characteristics	<p>Conditional distributions of participants are reconstructed econometrically, using the model that is believed to best adhere to the hierarchical structure of the data as well as variability of approaches between MS. This is notably a mixed-effects with both fixed and random effects, along the lines of</p> $(4) Y_{ij} = \beta_0 + \beta_1 X_{ij} + \mu_0_j + \mu_1 x_{ij} + \theta_1 D2_i + \theta_2 D3_i + \theta_3 D4_i + \epsilon_{ij}$ <p>Where, Y = depending on the regression is the share of low skilled, medium skilled, high skilled in each observation. X = depending on the regression, is the share of inactive, unemployed, employed in each observation β_0 = constant μ_0_j = country fixed effects $\mu_1 x_{ij}$ random effects (country level) D1..3_i cat. of region</p> <p>Through such a model it is assumed that (and tested if) a linear correlation exist between a given labour market condition and the educational status, which can be leveraged upon to estimate how many are say, low skilled AND unemployed, using data on the marginal distribution of LM status and Educational attainment across all programmes and categories of regions. A similar reasoning applies to success rates, whose correlation with the LM status and educational level at the entry of the operation is tested to predict the conditional distributions of results by LM status and Educational attainment.</p>
	Distribution of results over time	A mix of qualitative reasoning and information on the intensity of correlation between a measure of progress and some measured indexes is exploited to select the one that is believed to better represent the advancement of the programmes (and how much more one can expect that it will be delivered from now onward).
	Disaggregation by typology of intervention	At this stage of the research, there is no clear information available that indicates how much was spent where and on what. This information is being analysed in the context of the mid-term evaluation of the YE and might be used in subsequent revisions of this research when the YE evaluation will be published (expectedly over the course of 2020). So at the moment there is no ad-hoc treatment of the different typologies of intervention, but this is somewhat less relevant as the focus is already on their measured or estimated partial equilibrium outcomes.

Broad Issue	Specific Issue	Proposed solution in the modelling approach
Data Suitability to measure effects	Causality and attribution issues	This is a pivotal aspect of the exercise. The key point is that partial equilibrium effects are estimated in this research through monitoring data and coefficients from CIEs on youth employment programmes (as per the three scenarios described below) and then the general equilibrium dynamics will be estimated through RHOMOLO in future research (see below).
	Sustainability of results	Three alternative approaches are devised to calculate (and compare) short/medium term partial-equilibrium effects of youth employment policies supported through ESF/YEI, namely:
	Lock-in effects	<p>a. By using immediate ESF/YEI result indicators, net short-term partial equilibrium effects are proxied as:</p> <ul style="list-style-type: none"> o From inactivity to unemployment: the share of inactive participants switching to unemployment (CR01) minus the average LM transition rate in each MS from inactivity to unemployment (proxy of spontaneous dynamic)⁵⁵ o From inactivity or unemployment to employment: the share of inactive or unemployed participants shifting to employment (CR04), minus the respective average LM rates <p>b. By using longer term ESF/YEI result indicators, net effects (employment only) are proxied as the share of originally inactive or unemployed participants in employment at 6 months of the operation minus their respective LM transition rate</p> <p>c. By using available Counterfactual Impact Evaluations in combination with ESF/YEI output indicators, the net effects (on employment only) are calculated as the increase in the probability of being employed in the medium term (6 to 24 months) times the number of participants, by area and skill level.</p> <p>In (a) the main issues are that (i) sustainability of results and lock in effects cannot be controlled for, (ii) the proxy for causality is rather rough, also as there might be endogeneity⁵⁶ and LM transition data is not disaggregated by regions. At the same time, this is considered to be less sensitive to spontaneous dynamic than (b), which, in turn, can better capture the effects once lock-in effects are arguably vanished. The superior model is clearly (c) where the net effects on participants are estimated empirically, thus accounting for the deadweight effect, lock-in effects as well as sustainability of results. However, the quality, availability and level</p>

⁵⁵ Eurostat, Labour market transitions - quarterly data [lfsi_long_q]. According to the Eurostat webpage “Labour market transitions show the movements of individuals between the labour market statuses of employment, unemployment and economic inactivity. They help to understand and interpret changes in the levels of labour market indicators based on the EU-Labour Force Survey (LFS).” Main issues of this data for the research: (i) Statistics are gathered in a context where public interventions exist, both at the EU and the national level, that may bias the estimates. (ii) Interventions carried out within ESF/YEI, tend to address those lingering at a greater distance to the labour market. This should be taken into account in interpreting the results (iii) Data is available at the MS level while regional LM can vary to a great extent.

⁵⁶ LM transition data might be biased due to the existence of support in a given territory, especially if the relative size of the programme is large compared to the LM in a country

Broad Issue	Specific Issue	Proposed solution in the modelling approach
		of disaggregation of CIE results means that they can cover only partly the data as explained in 3.3.6 (especially beyond YE measures). ⁵⁷
	General equilibrium effects ⁵⁸	<p>At any rate, the task of determining general equilibrium effects lies with RHOMOLO, where the wage curve and the endogenous mechanisms (described in 3.1.2 and Annex III) jointly determining the ripple effect generated by the increased labour participation and level of employment in each region on labour demand, aggregate output, regional spillovers, skills mobility and so forth. This includes the feedback on workers not receiving support (displacement) the consequences of increasing productivity of the workers (substitution effects) and other spillovers.</p> <p>In terms of the model's equation this approach aims to mimic an exogenous shift of the labour force. Using the RHOMOLO model notation we had the wage curve described in (1) and, in equilibrium the situation described in (2) where the labour demand (per region, economic sector and skill level) for all firms (per region and economic sector) equals the labour force (by region and skill level, considering the unemployment rate). Importantly, labour supply is determined exogenously and assumed not to vary (no natural population change). What is suggested here is to shift the exogenously the labour force parameters based on net partial equilibrium effects on participants, by year, skill level, LM status and region. This means shifting the unemployment rate (u) in each region (r) and for each skill level (e), adding yearly those youth supported into employment through the ESF/YEI.</p>
Data Comprehensiveness	Range measured benefits connected to support	In this research design, the focus is on the effects of the Youth Employment measures in terms of changes of occupational status, which is the ultimate goal of the interventions. What it is assumed in the model by exogenously shifting the values of the workforce is that participants, originally unemployable, become employable thanks to the intervention just like the rest of the population. So benefits in terms of upskilling or better motivation are already incorporated in this modelling approach. However, additional benefits from the support (i.e., health, crime rates etc.) cannot be estimated, due to both the lack of information on them as well as the absence of model specifications which can cover also these aspects.

Source: author's elaboration also through contacts and exchanges with the JR

⁵⁷ Indeed, it is clear that monitoring data, just like it was in the case of counterfactual evaluations, is focused on those who have received assistance, it cannot paint a comprehensive picture of what happens to the system as a whole already in the short to medium term (displacement effects) nor can it look at substitution effects (which happen due to the increased productivity of workers). In addition, it cannot fully account for spontaneous dynamics (deadweight), differently from the counterfactual. On the other hand, it helps paint a much more comprehensive picture of the achievements that may have been generated by a programme and, differently from standard approaches employed in macroeconomic modelling, focuses on results rather than inputs (expenditure) and is thus expected to better adhere to the specificity of the actions carried out within, in this case, the ESF

⁵⁸ (displacement, substitution, spillovers)

As shown in the table, this research represents in essence an attempt at making micro-level information from the ESF/YEI monitoring evaluation system a suitable input for the RHOMOLO model to jointly take into account micro and macro level dynamics. Such an attempt, as it is briefly clarified above and better explained in the following sections, is inevitably punctuated by a range of assumptions and need for proxies which aim to fill several information gaps, but are liable to carry bias with them.

Finally, it should always be borne in mind that other positive effects stemming from social inclusion activities can hardly be evaluated quantitatively (e.g. improvements in self-confidence, better health, better participation in the society, reduced crime rates etc.). Thus, they will need to be taken into account qualitatively when assessing the simulation results.

3.3 Methods and methodological steps

3.3.1 Overview

To follow the approach described in 3.2.3, a number of steps have been carried out, which have followed the below stylised process:

1. **Getting input data from the ESF monitoring system:** data from the ESF 2018 Annual Implementation Reports (submitted in 2019) has been retrieved, with information up to 31 December 2018
2. **Estimating conditional distributions of participants characteristics based on ESF/YEI monitoring data:** correlations between participants background characteristics have been tested to identify a parametric form relationship among their LM status and educational attainment to econometrically predict their conditional distributions. This involved:
 1. Identifying a suitable model which could fit ESF data and structure
 2. Determining whether the residual ‘NO ISCED’ category of the ESF monitoring system could be treated as ‘low skilled’
 3. Testing parameters and predicting conditional probabilities of being, Low, medium or high skilled if inactive/unemployed/employed on that basis (by category of region and fund)
 4. Creating samples of participants with predicted conditional probabilities
 5. Adjusting aggregated values of the predicted sample to match the aggregated values for participants by LM status
 6. Checking the extent to which the resulting aggregated values of participants by level of skills in the predicted (adjusted) sample differ from the original values
 7. Calculating cumulated (2018) average success rates by category of region and Fund for
 - 1 Inactive becoming unemployed (success rate job seeking or SRJS)
 - 2 Inactive or unemployed becoming employed (success rate employment or SRE)
 8. Testing parameters and predicting conditional probabilities of becoming
 - 1 Unemployed if inactive, by skill level
 - 2 Employed if inactive or unemployed, by skill level and previous LM status

9. Creating samples of predicted conditional probabilities for participants' results and their background characteristics
 10. Adjusting the predicted figures for the breakdown of successful participants so that these match the actual values
- 3. Identifying a proxy for the level of implementation of the OPs, which involved,**
 1. Understanding which indexes that are available can be used and what level of progress do they measure
 2. Testing econometrically the intensity of their correlation with a synthetic measure of progress
 3. Identifying the most promising proxy based on a quali-quantitative reasoning.
 - 4. Bringing data at the NUTS2 level, which involved several operational steps which are however reported in Annex III for conciseness.**
 - 5. Discounting values by coefficients of net effects based on counterfactual impact evaluations, which involved:**
 1. Screening available evaluations on the Youth Employment Initiative or on Investment priority 8.ii
 2. Identifying coefficients which could fit the monitoring data (i.e. net increases in probability of being employed for participants)
 - 6. Discounting values by LM transition data, which involved**
 1. Identifying and drawing data from Eurostat (MS level)
 2. Discounting non zero values for the LM transition data

It is worth mentioning that steps 5 and 6 are alternative, in the sense that they produce different scenarios (respectively (1) and (2) based on results from monitoring data discounted by LM transition data and (3) based on counterfactual impact evaluation).

3.3.2 Data used and how to get them

The first step has been to get input data from the ESF/YEI monitoring system.⁵⁹ As anticipated, this is used for exemplificative purposes only, as the data is not yet validated by the Managing Authorities. Data on output and results was aggregated at the OP level by category of region/fund. What this means is that there are 103 observations in the dataset, related to 73 Operational Programmes in 23 Member States. Of these, 37 include YEI investment and 53 ESF investment across three category of region (less developed, transition, more developed). 11 of such observations show however 0 participants, either due to monitoring issues, late start of the programmes or to changes in the structure of the OP (which means that they had previously selected IP 8.ii for support but no longer do so, and monitoring data was either never reported or shifted to other investment priorities).

⁵⁹ This is based on an extraction of Annual Implementation Reports 2019 as of 19 July 2019 from SFC, the System for Fund Management in the European Union, which is used within the frame of the study VT/2018/014 'ESF synthesis of Annual Implementation Reports and Progress Reports submitted in 2018 and 2019 and ad-hoc reports' and the study VC/2018/0715 'Study supporting the 2019 Evaluation of Youth Employment' carried out by FGB, with the author's support, and our partners Metis and Ockham. Data on output and results of the ESF by OP is however also available in the opencohesion portal and in Annual Implementation Reports prepared yearly by all managing authorities, which need to be published according to Art. 50(9) of the Common Provision Regulation, Reg (EU) No 1303/2013

Data is drawn from common output indicators (COI) and common result indicators (CRI), as per the below list:

1. Common Output Indicators (COI)
 - Breakdown of participants by LM status:
 - COI01 – unemployed, including long term unemployed
 - COI03 – inactive
 - COI05 – employed, including self-employed
 - Breakdown of participants by Educational attainment
 - COI09 – with primary (ISCED 1) or lower secondary education (ISCED 2)
 - COI10 – with upper secondary (ISCED 3) or postsecondary education (ISCED 4)
 - COI11 – with tertiary education (ISCED 5 to 8)
 - COI11a (calculated as COI01 + COI03 + COI05 – the sum of participants with educational attainment) – with no ISCED⁶⁰
2. Common Result Indicators (CRI)
 - CRI01 – Inactive participants engaged in job searching upon leaving
 - CRI04 – participants in employment, including self-employment, upon leaving
 - CRI06 - Participants in employment, including self-employment, six months after leaving

Though not quite the focus of this exercise, participation is reported for approx. 3.8 million people, of which approx. 0.9 million are inactive, 2.8 unemployed and approx. 0.05 employed. Low skilled and medium skilled are quite evenly distributed (1.5 million each) whilst the number of high skilled is approx. 0.65 million.⁶¹

3.3.3 Econometric estimation of conditional distributions of participants characteristics

This step is concerned with the estimation of cross-distributions or conditional distributions of participant characteristics, both with respect with the output and result data of the ESF/YEI, based on monitoring data.

Below a synoptic table is presented showing the choices eventually made to estimate them. The following sections describe the steps undertaken to get there and the underlying methodological choices.

⁶⁰ The ESF regulation includes a “completeness requirement” that means that data on participants can only be recorded when the underlying micro-data are complete for all non-sensitive variables (gender, employment status, age, level of education and household situation). The only variables that can be incomplete are those dealing with special categories of personal data (i.e. sensitive data) that are required for indicators marked with “***” in Annex I of the ESF Regulation (Regulation (EU) No 1304/2013) (i.e. those related to migrants, participants with a foreign background, minorities, participants with disabilities, and other disadvantaged). This means that the core data based on non-sensitive variables should always be complete and internally consistent (e.g. sum of employed+unemployed+inactive should equal total participants). However, participants for which the ISCED is not known may potentially be encountered so the total of COI09-10 and 11 does not equal that of COI01-3-5. For this reason, an indicator was derived, called COI11a which is equal to the difference between the two abovementioned sums. This should represent the number of those with no ISCED registered.

⁶¹ Values should be intended as exemplificative only as these are not yet validated by the EC and Managing Authorities. The aim of this research is only to showcase how ESF/YEI could be used in practice to feed into the RHOMOLO model, and not to discuss actual values, which might be subject to revision.

Table 11 - Overall approach to the reconstruction of ESF/YEI participants' background features

Data/ status	LM	Output data		Outcome data	
		Aggregated	Disaggregated data by LM	Disaggregated data by LM and ED. Attainment ⁶²	Aggregated
Inactive	Sum of COI 01-03-05	COI 03	<ul style="list-style-type: none"> • Estimated Inactive low skilled (LS) = COI03*probability of being low skilled if inactive by Member states (MSs) and Cat of region • Estimated Inactive medium skilled (MS) = COI03* probability of being medium skilled if inactive by MSs and Fund/Cat of region • Estimated Inactive high skilled (HS) = COI03* probability of being high skilled if inactive by MSs and Fund/Cat of region 	CRI01 for the inactive starting to look for a job and CRI04 and CRI06 for the inactive or unemployed becoming employed respectively upon leaving the intervention or at 6 months of it.	<ul style="list-style-type: none"> • Becoming unemployed if inactive: pro-quota estimated distribution⁶³ • Becoming Employed if inactive LS = Estimated Inactive LS*probability of being employed by MSs if originally inactive and LS⁶⁴ • Becoming Employed if inactive MS = Estimated Inactive MS*probability of being employed by MSs if originally inactive and MS • Becoming Employed if inactive HS = Estimated Inactive HS*probability of being employed by MSs if originally inactive and HS
Unemployed			COI 01		<ul style="list-style-type: none"> • Estimated Unemployed LS = COI1*probability of being low skilled if unemployed by MSs and Fund/Cat of region • Estimated Unemployed MS = COI1*probability of being medium skilled if unemployed by MSs and Fund/Cat of region • Estimated Unemployed HS = COI1*probability of being medium

⁶² In all cases, Predicted conditional probabilities are then adjusted linearly to match actual data of marginal distribution, with one iteration (actual data are matched at the level of the LM breakdown). The level of deviation from the actual data on educational attainment is reported in the analysis.

⁶³ No statistically significant difference found among skill levels when it comes to inactive becoming unemployed, as described below in section 3.3.3.8.

⁶⁴ Here and below, two probabilities are calculated, one for CRI04 (in empl upon exit) and the other one for CRI06 (in empl after 6 months) as shown in section 3.3.3.8

⁶⁵ Here and below, two probabilities are calculated, one for CRI04 (in empl upon exit) and the other one for CRI06 (in empl after 6 months) as shown in section 3.3.3.8

Data/ status	LM	Output data		Outcome data	
		Aggregated	Disaggregated data by LM	Disaggregated data by LM and ED. Attainment ⁶²	Aggregated
				skilled if unemployed by MSs and Fund/Cat of region	
Employed		COI 05	<ul style="list-style-type: none"> • Estimated Employed Low skilled=COI5* probability of being low skilled if employed by MSs and Fund/Cat of region • Estimated Employed Medium Skilled=COI5* probability of being medium skilled if employed by MSs and fund/Cat of region • Estimated Employed High Skilled=COI5*probability of being high skilled if employed by MSs and fund/category of region 	No measured outcomes for the employed as this is not relevant for participants in receiving youth employment support. Activities are not directed towards the employed them and, consistently, this category only accounts for 1.8% of the total participants in investment priority 8.ii (YE)	

Source: author's elaboration

3.3.3.1 Identifying a suitable model which could fit ESF data and structure

As it is logical for a territorial policy, the ESF/YEI is structured in the form of areas and sub areas of interventions, both geographically and in the sense of the structure of activities it carries out. As a result, the data stemming from its monitoring system is obviously nested, that is, data on participants is nested within an action that belongs to an Investment Priority that falls within a Thematic Objective and is carried out in a region located within a Member State. One could identify even more levels but the main idea is that the data is nested. Another key feature of the data at hand is that there isn't a specific interest in observing its dynamic over time. Rather we would like to know on average how many people have been supported and with what result. Because ESF data may be quite volatile and subject to time lag (e.g. there's a time form the entry to an operation to the reporting of the potentially related result, reporting may be subject to revisions etc.) then it was deemed best to use cumulative 2018 data. This means that the database at hand is cross-sectional, that is, a number of observations across different entities but no longitudinal (yearly) info on them.

Thus, hierarchical models seem the obvious choice. The main rationale underpinning such models is that part of the measured variation across, say, the outcome (dependent) variables of the observed units is not explained by differences of the independent variable but by idiosyncratic factors, that is, by factors that are germane to the environment in which the units are embedded. To put it more simply, it may be that, e.g., the variation of success rates for participants is not (only) explained by a given form of intervention, or labour market status or thematic objective to which it belongs but, rather, by the labour market or administrative practices of a given region or MS. If we assume these as sufficiently constant with respect to the outcome variable, then the best way to address for them is simply to exclude them from the computation, by using different "intercepts" for the regression line (the trend we want to observe). Mixed models in particular allow combining random intercepts (which is the case of a fixed effects model) and random slopes as necessary. In the first case, there is one common trend that however moves from different starting points depending on the MS in which it is observed. In the second case, not only the starting point but also the actual trend (i.e. the slope of the regression line) is allowed to vary. This latter hypothesis is also quite consistent with the structure of the ESF, as different programming choices and differences in the LM structure of each MS can affect also the slope of the regressor.

In the light of the above mentioned reflection, the model used throughout is a Mixed effects model⁶⁶ as per equation (4), introduced in section 3.2.3.

Because two funds (ESF/YEI) may support YE interventions one of which (the ESF) provides different intensities of financing according to, a.o., 3 different category of regions, dummies are introduced in the model to check whether differences may be explained not only by the MS but also by the category of region or fund.

⁶⁶ In stata, the code was typically along the lines of: `xtmixed dependentvariable independentvariable i.category of region, || ms: independentvariable`

Such model specification was tested with respect to other functional forms and have consistently returned smaller confidence intervals and, thus, higher significance of the estimates.

3.3.3.2 *Determining whether the residual ‘no isced’ category of the ESF/YEI monitoring system could be treated as ‘low skilled’*

The first step was to apply the Hierarchical linear model to check whether a correlation exists under the ESF in EU MS between the share of inactive within a programme and the shares of participants by skill level.

A multi-level mixed-effects linear regression was applied that took into account the share of inactive (computed as COI03/(COI01+COI03+COI05)) and the share of low skilled. The assumption here is that inactivity is correlated with lower levels educational attainment. The results are shown below.

Table 12 – STATA output for the mixed effects model run to estimate the correlation between skill level and employment status upon entry to the operations

VARIABLES	(1) no isced	(2) isced 1-2	(3) low skilled combined	(4) medium skilled	(5) high skilled
Share of inactive	0.109** (0.0531)	0.374*** (0.0906)	0.469*** (0.0796)	-0.207*** (0.0691)	-0.239*** (0.0624)
More developed (dummy)	-0.0221 (0.0388)	0.130* (0.0662)	0.105* (0.0576)	-0.0814 (0.0500)	-0.0239 (0.0444)
Transition (dummy)	0.0158 (0.0504)	0.0425 (0.0861)	0.0568 (0.0746)	-0.0470 (0.0646)	-0.00995 (0.0572)
YEI (dummy)	0.0108 (0.0358)	0.0303 (0.0612)	0.0406 (0.0531)	0.00174 (0.0460)	-0.0437 (0.0408)
Constant	0.0263 (0.0353)	0.190*** (0.0602)	0.214*** (0.0547)	0.516*** (0.0482)	0.271*** (0.0470)
RANDOM EFFECTS					
lns1_1_1 Constant	-19.20*** (4.661)	-28.04*** (6.184)	-26.96*** (4.742)	-26.79*** (4.943)	-19.87*** (6.111)
lns1_1_2 Constant	-2.741*** (0.270)	-2.208*** (0.309)	-2.142*** (0.285)	-2.223*** (0.300)	-2.035*** (0.234)
lnsig_e Constant	-2.107*** (0.0793)	-1.572*** (0.0810)	-1.722*** (0.0833)	-1.868*** (0.0856)	-2.000*** (0.0856)
Observations	92	92	92	92	92
Number of groups	23	23	23	23	23
LR test vs. linear regression	9.403	7.186	9.795	5.692	14.04
Prob <	0.00908	0.0275	1.42e-08	0.00514	0.000893
Wald chi	5.504	24.17	42.34	14.80	15.74
Prob <	0.239	7.37e-05	0.00747	0.0581	0.00339

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Source: Author’s elaboration based on tentative AIR 2019 data

The model identifies a statistically significant correlation between the share of inactive participants and different average skill levels in the programmes. The LR test vs. linear regression suggests that this model improves upon the linear (non-hierarchical regression). In addition, the coefficients for the random effects are also statistically significant on average. If one assumed linearity and that the entire population were inactive (100%) then the estimated share of low skilled (combining no isced and isced 1-2) would equal on average 68% (Share of inactive* β + constant). The fund and typology of region don't seem to affect the estimates. The same share would be around 30% and close to 5% for the high skilled.

It is interesting to notice that no isced and isced 1-2 are similarly correlated to inactivity levels across programmes. This provides further support to the already intuitive assumption that the no ISCED should be considered as low skilled much like those with ISCED 1-2. Thus, in the subsequent analyses, the share of low skilled has been considered as the sum of no ISCED and ISCED 1-2 in any given observation.

3.3.3.3 Testing parameters and predicting conditional probabilities

Based on the above and assuming linearity, one could thus estimate the probability of being low skilled if inactive, unemployed or employed. The same applies to the medium and high skilled. The key idea is that when estimating values based on the above model, at the right end of the distribution (i.e., where the share of inactive, or unemployed or employed equals to 100%) one can find the probability of being low, medium or high skilled (depending on the regression) AND inactive, unemployed or employed. This is always controlled for by the fund and category of region. Clearly, the underlying assumptions are quite strong. But in absence of more detailed data on background features of participants, econometric estimation seems the only way forward.

The results of the estimations are reported in the STATA output below.

Table 13 - STATA output of the mixed effects models run to estimate the conditional probabilities of participants background features (by LM status and skill level)

VARIABLES	(1) inactive LS	(2) inactive MS	(3) inactive HS	(4) unemployed LS	(5) unemployed MS	(6) unemployed HS	(7) employed LS	(8) employed MS	(9) employed HS
% of inactive (1-3), % of unemployed (4-6), % of employed (7-9)	0.469*** (0.0796)	-0.207*** (0.0691)	-0.239*** (0.0624)	-0.457*** (0.0784)	0.202*** (0.0712)	0.249*** (0.0669)	0.615 (0.380)	-0.410 (0.300)	-0.166 (0.273)
More developed	0.105* (0.0576)	-0.0814 (0.0500)	-0.0239 (0.0444)	0.0831 (0.0579)	-0.0740 (0.0497)	-0.0105 (0.0418)	0.0979 (0.0685)	-0.0737 (0.0544)	-0.0252 (0.0488)
Transition	0.0568 (0.0746)	-0.0470 (0.0646)	-0.00995 (0.0572)	0.0489 (0.0749)	-0.0465 (0.0643)	-0.00887 (0.0537)	0.120 (0.0841)	-0.0743 (0.0668)	-0.0413 (0.0600)
YEI	0.0406 (0.0531)	0.00174 (0.0460)	-0.0437 (0.0408)	0.0392 (0.0534)	0.00229 (0.0464)	-0.0404 (0.0398)	0.0251 (0.0605)	0.00746 (0.0480)	-0.0355 (0.0431)
Constant	0.214*** (0.0547)	0.516*** (0.0482)	0.271*** (0.0470)	0.676*** (0.0758)	0.310*** (0.0632)	0.0165 (0.0492)	0.313*** (0.0613)	0.472*** (0.0467)	0.218*** (0.0496)
RANDOM EFFECTS									
Ins1_1_1 Constant	-26.96*** (4.742)	-26.79*** (4.943)	-19.87*** (6.111)	-2.698* (1.404)	-2.281*** (0.582)	-1.815*** (0.216)	-24.79*** (6.421)	-26.09*** (6.667)	-24.84*** (5.957)
Ins1_1_2 Constant	-2.142*** (0.285)	-2.223*** (0.300)	-2.035*** (0.234)	-2.365*** (0.582)	-2.842** (1.347)	-28.88 (0)	-1.876*** (0.254)	-2.254*** (0.316)	-1.867*** (0.215)
Insig_e Constant	-1.722*** (0.0833)	-1.868*** (0.0856)	-2.000*** (0.0856)	-1.718*** (0.0841)	-1.866*** (0.0899)	-2.041*** (0.0828)	-1.597*** (0.0834)	-1.822*** (0.0852)	-1.950*** (0.0853)
Observations	92	92	92	92	92	92	92	92	92
Number of groups	23	23	23	23	23	23	23	23	23
LR test vs. linear regression	9.795	5.692	14.04	8.322	5.804	21.90	14.52	5.280	20.73
Prob <	0.00747	0.00514	0.000893	1.50e-08	0.00459	0.00484	0.000702	0.114	3.15e-05
Wald chi	42.34	14.80	15.74	42.22	15.05	14.94	7.883	7.458	1.335
Prob <	1.42e-08	0.0581	0.00339	0.0156	0.0549	1.75e-05	0.0960	0.0713	0.855

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Source: Author's elaboration based on tentative AIR 2019 data

The above regressions show that the models are statistically significant for the inactive and unemployed, whilst values are not statistically significant for the employed, most likely due to their very small presence (less than 1% of participants) across YE actions. Once more, random effects have statistically significant coefficients on average and that the distribution across skill levels of the inactive and unemployed is rather complementary, with a much stronger presence of low skilled among the inactive and of medium to high skilled among the unemployed.

3.3.3.4 Creating samples of participants with predicted conditional probabilities

Once estimated, the statistically significant values from the regression were used to reconstruct the conditional distributions of skill levels by LM status.

In particular, the actual beta coefficients and fixed effects coefficients are estimated by STATA while running the regressions, stored and exported to the database with monitoring data for all observations. In addition to these, random effects are predicted through a Best Linear Unbiased Predictor (BLUP) and added to the estimates. Fitted values from the mixed effects model are then complemented, whenever a coefficient for dummies of the fund/category of region is found to be statistically significant, with its value. What this means an estimate is produced (partly predicted through BLUP) for each observation that takes into account: (i) differences in constants at the MS level (ii) differences in the slope of the parametric relation at the MS level (iii) differences for the (statistically significant) constants of typology of regions/fund. This is similar to assuming that within the same MS, fund and category of region, background features of participants are distributed evenly. In the table below, an overview (MS average, for conciseness) is provided.

As it can be seen, the sum of the estimated probabilities of the skill levels by labour market status is quite close to 100% in most cases.

Table 14 - estimated conditional probabilities for participants' labour market status and skill level, by MS

MS	Inactive				Unemployed				Employed			
	LS	MS	HS	Total	LS	MS	HS	Total	LS	MS	HS	Total
BE	78.7%	27.9%	0.0%	106.6%	27.8%	50.5%	18.0%	96.3%	50.3%	38.7%	10.5%	99.5%
BG	65.8%	31.7%	5.4%	102.8%	19.3%	52.3%	30.2%	101.7%	29.8%	47.4%	23.3%	100.5%
CY	60.7%	21.2%	34.0%	115.9%	9.6%	40.7%	62.1%	112.5%	11.8%	40.0%	58.4%	110.2%
CZ	71.4%	31.3%	0.0%	102.7%	24.8%	51.7%	19.5%	96.0%	33.9%	48.0%	16.4%	98.3%
DE	99.7%	15.7%	0.0%	115.4%	40.3%	37.7%	12.9%	90.9%	56.4%	34.1%	8.5%	99.0%
ES	79.9%	13.6%	12.4%	105.9%	31.4%	30.7%	39.1%	101.2%	37.3%	32.3%	34.0%	103.6%
FR	82.0%	26.7%	0.0%	108.7%	37.1%	46.3%	11.9%	95.3%	49.4%	41.5%	9.0%	99.8%
GR	61.8%	36.3%	5.8%	103.9%	15.4%	57.3%	30.0%	102.7%	20.3%	53.2%	27.2%	100.7%
HR	66.0%	20.4%	20.4%	106.7%	19.4%	40.0%	46.8%	106.2%	23.7%	39.4%	43.7%	106.8%
HU	71.2%	42.9%	0.0%	114.1%	21.5%	64.1%	10.4%	96.0%	25.7%	60.3%	10.2%	96.2%
IE	75.0%	24.8%	0.7%	100.5%	27.6%	46.6%	24.5%	98.7%	47.3%	38.9%	13.1%	99.3%
IT	70.1%	32.2%	4.5%	106.8%	16.9%	54.8%	29.3%	101.0%	28.1%	48.9%	22.8%	99.8%
LT	66.3%	33.0%	3.3%	102.6%	19.9%	53.6%	26.8%	100.3%	27.7%	49.5%	22.8%	100.1%
LU	70.2%	40.2%	2.5%	112.9%	14.8%	57.0%	26.7%	98.5%	27.8%	53.3%	15.5%	96.6%
LV	65.2%	35.5%	1.3%	102.0%	18.9%	55.9%	24.4%	99.2%	29.0%	50.8%	18.6%	98.4%
MT	60.1%	41.4%	0.2%	101.6%	15.3%	57.4%	25.5%	98.2%	27.6%	54.5%	13.4%	95.5%
PL	60.5%	40.3%	6.2%	107.0%	10.9%	61.5%	30.3%	102.6%	16.1%	56.9%	26.8%	99.8%
PT	64.4%	25.5%	17.4%	107.3%	18.0%	45.3%	44.6%	107.9%	23.5%	43.5%	40.1%	107.1%
RO	71.1%	39.7%	0.0%	110.8%	24.4%	60.9%	8.9%	94.2%	29.6%	57.2%	8.6%	95.3%
SE	77.4%	33.5%	0.0%	111.0%	25.9%	54.3%	14.4%	94.6%	38.1%	48.7%	10.5%	97.3%
SI	64.3%	33.1%	10.3%	107.7%	14.3%	53.8%	34.9%	103.0%	17.0%	51.6%	33.2%	101.7%
SK	66.0%	31.8%	5.7%	103.5%	19.5%	52.3%	29.9%	101.6%	25.4%	49.1%	27.0%	101.5%
UK	79.5%	32.1%	0.0%	111.6%	30.6%	52.3%	9.6%	92.4%	43.3%	47.6%	7.7%	98.6%
Overall	75.1%	29.4%	3.8%	108.3%	25.5%	50.2%	22.4%	98.2%	36.5%	45.4%	18.0%	100.0%

Source: Author's elaboration based on tentative AIR 2019 data

The subsequent step, is that of adjusting these probabilities in order that they add up to exactly 1 so as to be used as coefficients which multiply the number of participants by LM status for each observation. Once all the values are adjusted to 100%, actual values of inactive, unemployed and employed are multiplied by the estimated probability of being Low, medium or high skilled.

3.3.3.5 Adjusting aggregated values of the predicted sample to match the aggregated values for participants by LM status

As the probabilities of being low, medium or high skilled given a labour market status are just estimated and not observed values, there will always be a difference between the total estimated values by educational attainment and the actual ones.

Accordingly, the subsequent step was to adjust values so that they were proportionally increased or decreased in order to match their total by skill level. This allowed, on the one hand, to keep the relative probability of being low/medium/high skilled within a programme as close as possible to the estimates,⁶⁷ while, on the other hand, matching the actual figures by skill level. Nevertheless, this implied having values that would no longer match the actual aggregated figures by LM status.⁶⁸

Thus, one additional iteration of the process was applied this time constraining on the total number of participants by LM status. Clearly, this made the total values by Skill level depart again from the actual one, but to a lower extent. Multiple iterations that progressively bring value closer to the actual ones may be an option, but this should be made somehow automatic in the future if necessary.

3.3.3.6 Checking the extent to which the resulting aggregated values of participants by level of skills in the predicted (adjusted) sample differ from the original values

To evaluate the extent to which the predicted adjusted samples departed from the aggregated values of the original observation by skill level across the entire distribution two steps are carried out:

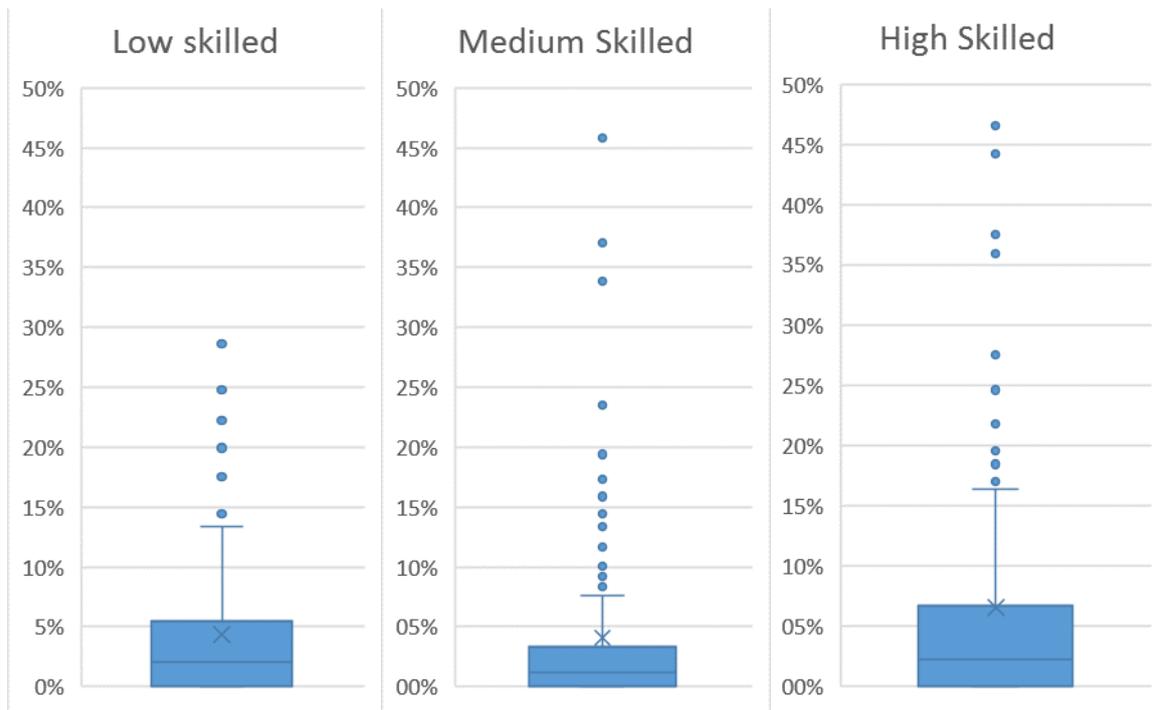
- measuring the relative distance from the predicted value; and
- plotting the distribution of relative distances from the predicted value using box plots.

The goal here is appraise not only the average deviation, but also its distribution over the different OPs and observations.

⁶⁷ The fact that it is proportional is aimed to ensure that the ratio between the probability of being low skilled if inactive and the probability of being low skilled if unemployed/employed remains to the extent possible constant and the same of what found econometrically.

⁶⁸ As this is a linear transformation, it may be that increasing proportionally the number of, e.g., inactive being low skilled means that the number of inactive low skilled needed to keep the ratio constant is higher than the number of inactive themselves, or that the latter is exceeded when adding up the remainder.

Figure 8 - box plots describing the distribution of the deviation of the aggregated estimated samples by skill level from the actual aggregated monitoring data by skill level.



Source: author's elaboration based on tentative AIR 2019 data

For the Low Skilled, we have an average deviation of approx. 4%, with a maximum value of the distribution (excluding outliers) of approx. 13%. $\frac{3}{4}$ of the distribution, however, sit just above the 5% threshold, with a median value of approx. 2%.

For the medium skilled, the average deviation is of 4% and the maximum value of the distribution excluding outliers of around 8%. $\frac{3}{4}$ of the distribution sit below the average, with a median value of approx. 1 percent.

For the high skilled, values are slightly higher and the average deviation is of 7% and the maximum value of the distribution excluding outliers of 8.6%. $\frac{3}{4}$ of the distribution sit below the average, with a median value of 1.4%.

If we take into account absolute figures:

- for the low skilled, the total difference is approx. 50 thousand over 1.5 million participants (3%);
- for the medium skilled, the total difference is approx. 10 thousand over 1.5 million participants (0.6%); and
- for the high skilled, the total difference is approx. 5 thousand over 0.65 million participants (0.8%).

These values represent a substantial improvement if compared with the imbalances in predicted data but could potentially be improved with further iterations of the adjustment procedure. Nevertheless, the higher the number of iterations the further the distance from the predicted ratio of LS to MS to HS in a given observation. Thus, this is deemed as good balance for the time being.

It is important to notice that at this stage the estimated figures of participants by background characteristic and programme are known. This is the basis for applying coefficients drawn from the Counterfactual Impact Evaluation. Hence, the steps described in sections 3.3.3.7 – 10 do not apply to scenario 3, but only to scenario 1 and 2.

3.3.3.7 Calculating cumulated (2018) average success rates

This step simply consisted in linking results to participants, that is, calculating the share of:

- inactive people starting to look for a job upon completion of the intervention (Success rate for job seeking);
- inactive or unemployed people becoming employed (Success rate for employment); and
- inactive or unemployed people becoming employed.

These success rates are included in the dataset and were used in the subsequent steps. However, the share of inactive people starting to look for a job should not be calculated based on the whole pool of inactive participants, as some of them have expectedly found a job, so these need to be subtracted from the denominator. In the next section this point is further clarified.

3.3.3.8 Testing parameters and predicting conditional probabilities for result indicators

This step involved understanding:

1. the extent to which results - and notably an upward shift in the labour market status of the participant upon exit to the operation – can be ascribed to participants with a given labour market status at the entry to the operation; and
2. the distribution of such results also by the skill level of the participants, proxied, as above, by the highest educational attainment of the participants at the entry of the operation.

Similarly to the approach used in previous estimates, a multi-level mixed effects model is used, but this time with two predictors, namely the shares of low, medium and high skill participants as well as their labour market status (which is binary, either inactive or unemployed, so the share of inactive has been used).

The analysis was replicated for both the success rate at 6 months of the operations and immediately upon exit.

The results are shown below, with model specifications 1-3 focusing on short-term success rates (CR04) and 4-6 testing the correlation among longer term success rates (CR06) and participants' background features.

Table 15 - STATA output of mixed effects models used to estimate the success rates of inactive and unemployed, by skill level

VARIABLES	(1) % LS in empl	(2) % MS in empl	(3) % HS in empl	(4) % LS in empl LT	(5) % MS in empl LT	(6) % HS in empl LT
Share of low skilled (1) Share of medium skilled (2) Share of high skilled (3)	-0.185*** (0.0676)	0.146 (0.0945)	0.197** (0.0863)	-0.173* (0.0977)	0.0660 (0.114)	0.166 (0.120)
share of inactive on inact + unempl	0.0244 (0.0559)	-0.0545 (0.0493)	-0.00957 (0.0520)	-0.138 (0.0852)	-0.198*** (0.0752)	-0.172** (0.0783)
More developed	0.0392 (0.0366)	0.0234 (0.0365)	0.0263 (0.0364)	0.0376 (0.0563)	0.0274 (0.0563)	0.0237 (0.0555)
Transition	-0.0317 (0.0465)	-0.0452 (0.0475)	-0.0390 (0.0468)	0.0303 (0.0725)	0.0283 (0.0723)	0.0236 (0.0722)
YEI	0.0257 (0.0332)	0.0211 (0.0337)	0.0269 (0.0336)	-0.0107 (0.0513)	-0.0169 (0.0511)	-0.0106 (0.0514)
Constant	0.285*** (0.0415)	0.172*** (0.0503)	0.193*** (0.0459)	0.369*** (0.0527)	0.299*** (0.0770)	0.288*** (0.0586)
RANDOM EFFECTS						
lns1_1_1 Constant	-18.56 (4.813)	-1.554*** (0.267)	-22.75*** (4.135)	-17.92*** (5.885)	-26.89*** (4.796)	-18.77*** (4.815)
lns1_1_2 Constant	-2.180*** (0.212)	-2.963*** (0.645)	-2.149*** (0.205)	-2.644*** (0.650)	-2.377*** (0.409)	-2.508*** (0.493)
lnsig_e Constant	-2.215*** (0.0842)	-2.205*** (0.0861)	-2.207*** (0.0839)	-1.740*** (0.0906)	-1.752*** (0.0867)	-1.746*** (0.0876)
Observations	92	92	92	92	92	92
Number of groups	23	23	23	23	23	23
LR test vs. linear regression	26.67	29.64	27.47	0.845	2.688	1.575
Prob <	1.62e-06	0.127	1.08e-06	0.655	0.261	0.455
Wald chi	11.93	8.583	9.466	13.16	9.191	11.34
Prob <	0.0358	3.67e-07	0.0919	0.0219	0.102	0.0450

Source: author's elaboration based on tentative AIR 2019 data

First of all it is interesting to notice that the hierarchical structure of the data seems to be a lot more nuanced when dealing with longer-term result indicators, as shown by the LR test vs. linear regression. The reasons for this might vary largely. One is that spontaneous dynamics increasingly affect the data so that background features of the target group become the key driver of their labour market condition, irrespective of the YE support to which they have been exposed. Differences might also be due to the way data is collected for the longer term indicators, which is survey based instead of for the entire population, but assuming compliance to the criteria set in regulations and related guidance documents this should not be the case.

In any event, the other notable difference is that the LM status at the entry of the operation seems irrelevant for immediate result indicators, whilst inactivity is negatively correlated (on average in a statistically significant manner) with success rates for longer-term results indicators. In most cases, skill level matters when it comes to success rates.

Once identified the coefficients, fixed effects that were stored are estimated, as in previous cases, also with predicted BLUP values for random effects (these are described in section 3.3.3.9).

Having estimated how many inactive have found employment, their immediate term success rate is subtracted from the pool of inactive to calculate the success rate for job seeking. This is because a positive outcome on the inactive can be traced as a) “becoming unemployed” and b) “becoming employed” and in principle these are mutually exclusive (if you become employed you are not registered as “jobseeker upon exit of the operation”). Thus, success is defined as:

$$\frac{\text{initially inactive participants becoming unemployed} / (\text{initially inactive participants} - \text{initially inactive participants becoming employed})}{\text{initially inactive participants becoming unemployed} / (\text{initially inactive participants} - \text{initially inactive participants becoming employed})}$$

A model is then applied that considers whether skill level affects the probability of switching from inactivity to unemployment, with the results shown below.

Table 16 – STATA output of mixed effect models to estimate the probability of starting to look for employment, by skill level

VARIABLES	(1) % of inactive LS jobseeking	(2) % of inactive MS jobseeking	(3) % of inactive HS jobseeking
Share of low skilled (1)/ Share of medium skilled (2) / Share of high skilled (3)	0.0528 (0.0913)	-0.0570 (0.115)	-0.00634 (0.169)
More developed	0.0466 (0.0547)	0.0499 (0.0549)	0.0553 (0.0540)
Transition	0.00242 (0.0730)	0.00499 (0.0727)	0.0111 (0.0726)
YEI	0.0625 (0.0515)	0.0653 (0.0517)	0.0650 (0.0520)
Constant	0.0632 (0.0540)	0.110 (0.0701)	0.0837* (0.0499)
RANDOM EFFECTS			
lns1_1_1 Constant	-2.592** (1.224)	-16.78*** (5.178)	-22.22*** (6.103)
lns1_1_2 Constant	-3.246** (1.268)	-2.982*** (0.658)	-2.979*** (0.653)
lnsig_e Constant	-1.827*** (0.0916)	-1.821*** (0.0895)	-1.820*** (0.0894)
Observations	76	76	76
Number of groups	18	18	18
LR test vs. linear regression	1.343	1.269	1.212
Prob <	0.511	0.530	0.545
Wald chi	2.321	2.318	2.065
Prob <	0.677	0.678	0.724

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Source: author's elaboration based on tentative AIR 2019 data

The results suggest that the skill composition does not affect the probability of switching from inactivity to unemployment for the inactive. It is worth mentioning that this test is only based on the level of skills for the entire sample of participants, thus not by all means conclusive, but the best one could do at this stage.

3.3.3.9 Creating samples of predicted conditional probabilities for participants' results and their background characteristics.

Thus, results are compiled following the approach just presented.

In particular:

- a. for what concerns the inactive starting to look for a job results are ascribed pro-quota by skill level. One should recall that the “quotas” are those estimated in the first part of this simulation, that is, as conditional probabilities of being low, medium or high skilled if inactive;
- B. for what concerns the inactive and unemployed, the estimation involved 2 steps
 1. estimating probabilities of becoming employed if low, medium or high skilled for the inactive and unemployed separately; and
 2. multiplying such probabilities for the estimated number of participants determined in 3.3.3.5.

Particularly for step B.1, the table below shows the results of the econometric estimations at the MS level.

Table 17 - estimated success rates of inactive and unemployed, by skill level and MS

MS	Success rate upon exit			Success rate at 6 month of exiting from support					
	Inactive and unemployed			Inactive			Unemployed		
	LS	MS	HS	LS	MS	HS	LS	MS	HS
BE	2.7%	6.1%	30.8%	2.6%	11.6%	24.0%	16.4%	31.4%	41.1%
BG	14.3%	42.5%	43.4%	12.0%	26.1%	35.9%	25.8%	45.9%	53.0%
CY	0.1%	16.5%	26.0%	4.1%	16.3%	25.2%	17.9%	36.1%	42.3%
CZ	20.1%	48.3%	49.4%	9.1%	21.4%	32.4%	22.9%	41.2%	49.5%
DE	10.1%	26.6%	35.9%	6.6%	15.1%	27.2%	20.4%	34.9%	44.3%
ES	15.0%	46.4%	39.8%	3.6%	13.1%	23.4%	17.4%	32.9%	40.6%
FR	3.1%	12.3%	31.2%	0.1%	8.1%	21.3%	13.9%	27.9%	38.5%
GR	0.0%	14.5%	30.0%	4.0%	14.4%	26.4%	17.8%	34.2%	43.6%
HR	18.4%	47.9%	44.5%	12.4%	27.2%	35.2%	26.2%	47.0%	52.4%
HU	24.9%	52.3%	56.9%	8.7%	20.1%	32.9%	22.4%	39.9%	50.1%
IE	7.4%	28.0%	34.4%	4.4%	13.8%	25.8%	18.2%	33.6%	42.9%
IT	0.0%	13.8%	24.4%	1.5%	12.1%	23.7%	15.3%	31.9%	40.9%
LT	5.8%	24.2%	35.3%	3.7%	13.5%	25.7%	17.5%	33.3%	42.9%
LU	4.8%	26.0%	36.4%	5.7%	17.1%	28.9%	19.5%	36.9%	46.0%
LV	5.9%	24.3%	36.1%	6.9%	18.6%	30.0%	20.7%	38.4%	47.1%
MT	7.8%	30.8%	39.7%	4.9%	15.7%	28.0%	18.7%	35.5%	45.1%
PL	34.9%	75.3%	66.3%	15.8%	31.7%	41.2%	29.6%	51.5%	58.3%
PT	17.3%	48.4%	44.5%	8.9%	22.3%	31.6%	22.7%	42.1%	48.8%
RO	0.0%	5.1%	26.3%	0.0%	5.4%	20.3%	12.5%	25.2%	37.5%
SE	16.3%	44.6%	46.0%	6.4%	16.8%	29.1%	20.2%	36.6%	46.3%
SI	11.2%	39.4%	40.5%	4.6%	15.8%	26.9%	18.4%	35.6%	44.1%
SK	14.1%	39.9%	43.3%	5.7%	16.8%	28.2%	19.5%	36.6%	45.3%
UK	5.0%	18.7%	34.2%	3.1%	11.6%	25.1%	16.9%	31.4%	42.3%
EU avg	6.8%	23.3%	34.2%	3.6%	13.5%	25.5%	17.4%	33.3%	42.7%

Source: author's elaboration based on tentative AIR2019 data

As there is no need this time around that the sum of probabilities matches 100%, these are the coefficients that can be directly used to multiply the number of participants by skill level.

As mentioned in A, the distribution by skill level of success rates in terms of participants starting to look for a job is just pro quota, so the values seen in 3.3.3.4 apply also here.

3.3.3.10 Adjusting the predicted figures for the breakdown of successful participants

After having estimated the number of low skilled, medium skilled and high skilled (by labour market status, pro quota) the estimated conditional distributions are increased or decreased proportionally so as to match actual aggregated value of inactive or unemployed becoming employed. This allows to keep the ratio of probabilities of becoming employed steady across different skill levels, but reduces linearly success rates for observations where the success rate was lower than the expected value (estimated econometrically).

3.3.4 Quali-quantitative identification of a proxy for the level of implementation of the OPs

3.3.4.1 Why it matters

Identifying an index that mimics the progress of implementation of ESF/YEI OPs is a pivotal exercise within this data preparation part as it affects:

- **the geographical distribution of participations and results at the NUTS2 level:** as participations and results are not registered at the NUTS2 level, but only at the higher level of OP/FUND/Category of region, there's a need to use some reference value to split proportionally aggregated participations/results into disaggregated NUTS2 regions; and
- **the distribution over time:** as we are only 5 years into the implementation of the ESF/YEI with the data available but this will be continue to be measured until 2023, there's a need to understand what more could be expected and to what level of implementation do the current shares of participations and results correspond.

In the ESF monitoring practice, four main measures exist that have a 2023 target and could be considered as proxies for the level of implementation,⁶⁹ notably:

- **the expenditure rate** (total eligible expenditure declared by beneficiaries to Managing Authorities/ total allocation);
- **the project selection rate** (total eligible cost of operations selected for support/ total allocation);
- **the achievement rate for output indicators** (total cumulative value of an output indicator with a target/2023 target value); and
- **the achievement rate for result indicators** (total cumulative value of a result indicator with a target/2023 target value).

3.3.4.2 What the data tells us

One way of understanding whether any of those is a good proxy for measuring how advanced is a given programme is to test their correlation with some measure of advancement in participations or results. Intuitively, the more money spent/ the higher the achievement rate, the higher the number of participations and results. Obviously, the absolute number of participations and results cannot be a good backdrop, as it is heavily affected by the scale of the

⁶⁹ One could also consider the indicators selected for the Performance Framework but: (i) their level of disaggregation is not suitable for this study, as they are set at the PAX and not IP level; (ii) their coverage is somewhat narrower and they may carry stronger bias, i.e. if MAs give priority to them as they are subject of a specific monitoring procedure

programme. So, to get rid of this first “scale” bias, participations and results are divided by the overall allocation, in order to have participations per euro allocated and results per euro allocated. One can assume that other things being equal, the more a programme is advanced the higher the number of participations/results per euro allocated.

Next, because the two proxies of progress would have themselves different scales (participations are a lot higher than results) values are first standardised (thus their variation is tied to the standard deviation) and then normalised between 0 and 1 (this way their variation is comparable to the 0 to 1 range of achievement and expenditure rates). As the values are now comparable one another (and so should be their variation), it remains to be checked whether there might be different sources of bias. Among others, at least the following can (significantly) affect unit costs (thus the number of participants per euro allocated):

- typologies of intervention;
- target group;
- MS and purchasing power; and
- category of region.

The best way of controlling for this is use a hierarchical three-level model with mixed effects (observations are nested within categories of regions which are nested within MS – random effects are allowed thus the slope of each MS may differ), similar to what described by equation (4). In addition, covariates for the share of inactive and share of low skilled can be included. With such a model, one could argue that under IP 8.ii the target population is comparatively homogeneous and that similar types of interventions are offered to people in the same country, category of region, labour market status and with the same level of education, thus with similar costs. This also accounts for the parity of purchasing powers, as within the same country and category of region the costs should be comparable.

The result of this is a set of regressions which use as dependent variable alternatively the number of participations per euro (standardised and then normalised between 0 and 1) and the number of participants getting into employment (standardised and then normalised between 0 and 1) to see where the correlation is stronger.

Another thing to consider is that participations and results tend to occur (and be recorded) at different points in time. By definition, the former is measured at the beginning of an operation, whilst the latter upon its completion. Furthermore, results may take more time to be registered simply as that’s a more difficult data to obtain. So, the regressions are run separately for the two measures of progress.

Firstly, let’s see how the first set of regressions work with participations. The table below shows the beta coefficient and the p-value for all the variables included in the model. The differences among the columns are simply due to the different proxy variables put to test (identified in the first row of each table).

Table 18 - output of the regressions (multi-level mixed effects) on proxies for progress of output indicators

	Project selection rate	Expendit. rate	synthetic index 1 ⁷⁰	synthetic index 2	AR (output)	AR (output adjusted) ⁷¹	AR (results)	AR (results adjusted)
Proxy variable (see first row)	$\beta=.060$; $P=0.062^{**}$	$\beta=.1443$; $P=0.009^{***}$	$\beta=.0808$; $P=0.029^{**}$	$\beta=.1349$; $P=0.006^{***}$	$\beta=.0788$; $P=0.000^{***}$ $Z=6.18$	$\beta=.1444$; $P=0.000^{***}$ $Z=4.05$	$\beta=.0542$; $P=0.001^{***}$	$\beta=.0727$; $P=0.005^{***}$
SI ⁷²	-.0405; P=0.452	.0031; P=0.943	-.0365; P=0.483	-.0170; P=0.714	-.0518; P=0.310	-.0627; P=0.187	-.0630; P=0.223	-.0136; P=0.791
SLS ⁷³	0.0008; P=0.089*	.0434; P=0.401	.0515; P= 0.373	.0396; P=0.459	.0236; P= 0.655	.0365; P= 0.479	.04177; P= 0.436	.0432; P=0.439
constant	.0260; P=0.392	.009; P=0.685	.0173; P=0.559	.0045; P=0.860	.0250; P=0.277	-.003; P= 0.892	.0332; P=0.148	.0253; P=0.324

Source: author's elaboration based on tentative AIR2019 data

It would seem that proxy which is best correlated with the progress in implementation (measured as standardised number of participants per euro of allocation) is the achievement rate of output indicators, but other variables have confidence intervals of above 99% too, including the expenditure rate.

When it comes to results, the situation looks slightly differently, as per the table below.

⁷⁰ Because one could assume that participations occur after the operation is selected but before its expenditure is declared, a midway point could be calculated as the mid-way point between project selection and expenditure declared. The synthetic index 1 is based 25% on project selection and 75% on expenditure rate, whereas synthetic index 2 is based on 75% expenditure rate 25% project selection rate.

⁷¹ Because some OPs report target achievement of >>100%, it was decided to fix them at 100%; in other words their implementation is considered as completed.

⁷² Share of inactive

⁷³ Share of low skilled

Table 19 - output of the regressions (multi-level mixed effects) on proxies for progress of result indicators

	Project selection rate	Expenditure rate	synthetic index 1 ⁷⁴	synthetic index 2	Achievement rate (output)	Achievement rate (output adjusted) ⁷⁵	AR (results)	AR (results adjusted)
Proxy variable (see first row)	$\beta=.0217$; P=0.314	$\beta=.3239$; P=0.000*** Z=3.98 WaldX= 30.5	$\beta=.2175$; p=0.001*** Z=3.21	$\beta=.3153$; p=0.000*** Z=3.92 WaldX=29.2	$\beta=.1604$; p=0.000*** Z=3.67 WaldX=16.83	$\beta=.2867$; p=0.000*** Z=4.12 WaldX=27.3	$\beta=.0942$; p=0.002*** Z=3.07 WaldX=12.9	$\beta=.2079$; P=0.000*** Z=3.98 WaldX=28.4
SI	.1854; P=0.046***	.2692; P=0.002***	.2214; p=0.014**	.2610; p=0.002***	.1134; p=0.159	.1938; p=0.026**	.1180; p=0.144	.2285; P=0.006***
SLS	.0629; P=0.550	.0293; P=0.760	.0530; p=0.601	.0299; p=0.758	.0114; p=0.900	.0373; p=0.697	.0196; p=0.829	.0403; P=0.660
constant	.1568; P= 0.004	.0069; p=0.879	.0044; p=0.934	-.0138; p=0.777	.0958; p=0.056**	.0047; p=0.918	.1044; p=0.030	.0265; P=0.527

Source: author's elaboration based on tentative AIR2019 data

One can see that most proxy variables except the achievement rate for results are statistically correlated with a measure of progress of results. Nevertheless, the strongest correlation is found for the eligible expenditure alone, which is significant at over the 99% level. If one looks at the best proxy for participations, namely the adjusted achievement rate of targets for output indicators, that remains sufficiently correlated with the progress of results (further confirming it is generally a good proxy) but to a lesser extent if compared with eligible expenditure alone.

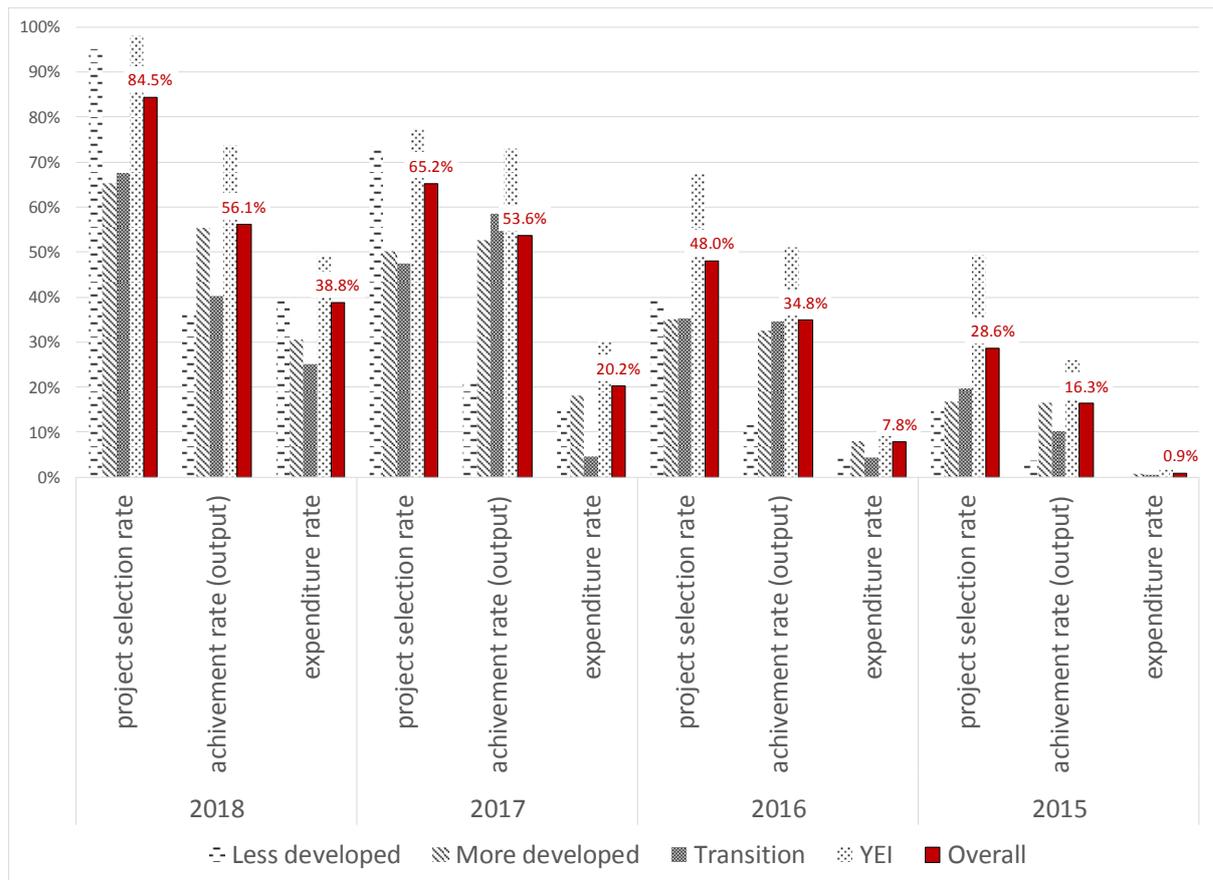
Other than correlations, one may want to consider how these variables have progressed over the first years of the implementation.

The graph below shows this advancement by category of region.

⁷⁴ Because one could assume that participations occur after the operation is selected but before its expenditure is declared, a midway point could be calculated as the mid-way point between project selection and expenditure declared. The synthetic index 1 is based 25% on project selection and 75% on expenditure rate, whereas synthetic index 2 is based on 75% expenditure rate 25% project selection rate.

⁷⁵ Because some OPs report target achievement of >>100%, it was decided to fix them at 100%; in other words their implementation is considered as completed.

Figure 9 - progress of selected financial and physical indicators, by year and fund/category of region



Source: author's elaboration based on tentative AIR2019 data

As it can be seen, the average achievement rate (AR) for output indicators shows values that are typically midway between those of eligible cost and eligible expenditure, which is consistent with the assumptions on how progress of implementation is reported. The difference with the eligible expenditure rates is however decreasing over time, as it is to be expected. It is interesting to notice that when looking at the distribution by category of region, the AR is however lower than the level of eligible expenditure in less developed regions in 2016 and 2017. This means that when using it as a proxy of progress, it would produce slightly more positive projections than eligible expenditure for less developed regions in 2016 and 2017. Another thing to be taken into account when comparing the achievement and the eligible expenditure rates is what values do they report for programmes that are having troubles setting up their administrative procedures for expenditure declaration.

- If one looks at instances in which cumulative expenditure rates at 2018 are 0 (14 obs), 60% of them (8 obs) have data on participations and approx 35% of them (5 obs) have data on participants in employment
- If one looks at instances in which achievement rates of output indicators at 2018 are 0 (8 obs), less than 40% (3 obs) have values on participations and approx. 25% (2 obs) have values on results.

Interestingly, there are instances in which 0 achievement rates are coupled with some eligible expenditure declared, and vice versa, which may be useful to increase the coverage of the proxies.

3.3.4.3 Summing up and proxy selection

As anticipated, there are two different areas for which proxies are used to distribute participation and results across regions:

- Geographical distribution (i.e. to ascribe Operational Programmes data to NUTS2 regions)
- Time profile (i.e. to understand how many participants and results should be measured each year for each operational programme and especially and how much more is to be expected for the future years of implementation).

3.3.4.3.1 Scenarios for the geographical distribution

As to the former, only financial data are available or can be reconstructed at the NUTS2 level, thus info on participation (achievement rates of output indicators) cannot be used to the end. According to section 3.3.4.2 analysis eligible costs are not strongly correlated with implementation at the moment. Thus, a better option would be to use eligible expenditure. Nevertheless, we know that eligible expenditure has gaps.⁷⁶ Thus, a slightly different scenario is used that typically relies on eligible expenditure but uses the eligible cost shares by NUTS2 region when the certified expenditure is missing.⁷⁷

3.3.4.3.2 Scenarios for the distribution over time

When it comes to the distribution over time, again there are two indicators that have showed to perform better than the average, notably eligible expenditure and achievement rates. However, it was noticed that both have (mutually different) data gaps. So two scenarios are suggested which can be used.

- a. Eligible expenditure with values from achievement rates where the first is missing (**preferred as time profile for RESULTS**).
- b. Achievement rates for output indicators with values from eligible expenditure where the first is missing (**preferred as time profile for OUTPUTS**).

It is important to notice that these different scenarios are mainly used in the actual RHOMOLO simulation (which needs yearly data) whereas in this research, only cumulated 2018 values are discussed.

⁷⁶ Some gaps, which cause values for some OP to be omitted, exist also for eligible costs due to the fact that there are some instances in which eligible costs are not reported but we have participations.

⁷⁷ It is important to consider that the sheer size of eligible costs (which is probably overestimated with respect to output and results as it relates to operations that are still ongoing) does not matter for the geographical distribution the relative NUTS2 proportion of expenditure incurred that matters to assign one programme's output/results to a given region.

3.3.5 Transformation of ESF data in NUTS2 level info

As abovementioned, the unit of analysis of this exercise so far has been that of the Fund/category of region by OP. But OPs can be defined at the NUTS0, 1, 2 or 3 level. So one key issue is how to convert such multilevel data into NUTS2 level data, especially where only the NUTS0 and NUTS1 breakdown is provided.

Several steps have been followed, including through VBA macros, to do so. In essence, values have been distributed pro-quota allocation at the NUTS2 level. A more thorough discussion is provided in Annex III.

3.3.6 Identifying coefficients to estimate net effects from Counterfactual Impact Assessment studies

Member States are required at least twice during the programming period, according to article 19(6) of the Reg (EU) 1304/2013 to produce and evaluation that “*shall assess the effectiveness, efficiency and impact of joint support from the ESF and the specific allocation for YEI including for the implementation of the Youth Guarantee. The first evaluation shall be completed by 31 December 2015 and the second evaluation by 31 December 2018*”.

This section is concerned with a brief survey on them which is aimed at identifying coefficients that might be suitable to estimate net effects for participants. In particular, the goal is to identify how many participants have found employment because of the intervention, ideally with coefficients by NUTS2 region, skill level and market status at the entry of the operation, to match the structure of the AIR2019 data elaborated for RHOMOLO according to the previous steps. However, as it will be shown, results are often not presented in such a way, or not estimated with a suitable breakdown. In addition, in some cases the evaluation focuses on specific measures, whose intensity within the funds cannot be easily reconstructed at this stage. As a result, a number of simplifying assumptions are used, which should be further improved upon in the future.

To the best of the author’s knowledge, 10 MS have followed counterfactual approaches to estimate the effects of youth employment measures, as per the list below.

Table 20 - overview of the methods used in the 2nd mandatory evaluations of the YEI, by MS

MS	Main technique used
Hungary	Regression Discontinuity Design
Bulgaria ⁷⁸	Propensity Score Matching and Difference in Differences
Sweden	Coarsened Exact Matching for the regional measures and RCT for the national one
France	Propensity score and difference in difference, depending on the strand evaluated (both regional and national measures under examination, through different approaches)
Spain	Propensity score matching

⁷⁸ A macroeconomic model is also used, but the goal of this survey is to identify coefficients suitable for an estimation of the net effects on participants.

MS	Main technique used
Italy	Covariate matching on a balanced counterfactual group
Latvia	Propensity score matching
Croatia	Propensity score matching
Poland	Coarsened Exact Matching
Slovenia	Difference in differences

Although linguistic obstacles hamper the possibility of an in-depth review of their methodological approaches within this research, the authors themselves in some cases (e.g. Slovenia) point out that they could only roughly balance a small share of participants with a control group. In Hungary, several outcomes labelled as “net effects” are de facto drawn from self-reported counterfactual approaches, and there is only one result (with no clear statistics presented) which compares the treated with a balanced control group. There, it is also unclear how many months after the intervention the positive employment result is measured. So caution is warranted in relying upon them.

Another issue lies with the heterogeneity of the outcome-variable tested. For instance in Latvia, only the number of days in unemployment is taken as a reference to compare the performance of the treated with the control group. In general, there is little investigation on inactive having started to look for a job, perhaps because it is inherently harder to track such a status in a comparative manner. This however means that one of the two changes of occupational status measured in scenario 1 and 2 will not be assessed under scenario 3.

Another concern relates to the breakdown of findings on the net effects, which is a key element within this research. The various reports point in some cases to effects which vary according to participants’ background features, as it is to be expected. However, in just a few cases these are explicitly reported in the evaluation published. In addition, differences across regions are also often not spelled out (with some notable exceptions, as it is the case for Italy). Regardless, one should consider that when looking for the net effects on the participants, considerations on the environment are slightly less compelling, as this is chiefly considered for the possible feedback loop that is however to be estimated by the RHOMOLO model in future research.

Against this backdrop, the following information could be ultimately drawn from the CIE studies screened, as shown by the Table below.

Table 21 - Synoptic table on findings of the second mandatory evaluations carried out by Member States on the YEI and related YE evaluations, by MS

MS	Name of the measure	Timeline of the impact (expressed as the change in probability of being employed with respect to the counterfactual situation)			by skill level or LM status			Other measured elements and notes
		Short term (1-3 month)	Medium term 6-9 month	Long term 12-month or more	LS	MS	HS	
HU	Entire programme	+ 15%	+ 6%					NO The results vary considerably by measure, from -5% to +10%. This is based on a regression discontinuity design. There are differences measured in the effects between cities and the countryside, but no differentiation by NUTS2 and nothing that could be used for the current research in that sense.
BG	Entire programme			+ 7.7%				NO Higher “net effects” on the participants are mentioned from other estimation methods (including self-reported counterfactual), which are in the range of 30 to 50%. The deadweight effect is estimated at 2%, the substitution effect at 13% and the displacement effect at 12%. Estimations based on the propensity score actually show a 0.077 times higher probability of the treated to be employed by the end of 2018. Not clear how the previous estimates are computed, so only the longer-term result based on propensity score matching is retained. Interesting to notice that according to the authors significant net effects are observed in employment expectations among young people with higher education
SE	Regional measures			+ 6 p.p.				NO Bigger impact for men and migrants (up to 13 p.p.). Coarsened Exact Matching used here
	Ung Framtid “Young future”			+3 p.p.				NO 10 less days of unemployment but only for women, but especially 3 p.p. less of unemployment for participants (as the probability of leaving unemployment for the following 12 months). Randomised experiment used for the national “Young Future” measure
FR	OVERALL (lot2)		neutral					NO There are 5 different evaluation strands which are then comprised under the umbrella report from KPMG et. Al (2019).

MS	Name of the measure	Timeline of the impact (expressed as the change in probability of being employed with respect to the counterfactual situation)			by skill level or LM status			Other measured elements and notes
		Short term (1-3 month)	Medium term 6-9 month	Long term 12-month or more	LS	MS	HS	
	AIJ		+10 p.p., 28% more					The last strand (lot 6) is a review of the overall approach . It is worth noticing that the absence of impact in regions that benefit from the YEI is attributed by the author of lot 6 (Farvaque, 2018) to the multiple forms of interventions which support YE also in non-YEI regions. Following this reasoning, the below counterfactual impact evaluation on specific measures are taken as a reference, instead of the overall CIE carried out in Lot2.
	APEC (Accompagnement 26 p.p. (63% vs renforcé des jeunes 37%) diplômés)		0 p.p. (79% vs 79%)					In terms of the national measures, it is mentioned that the Garantie jeunes and the AIJ represent over 80% of the allocation, whereas the APEC just 1%. This means the coefficients of the other two should be the focus. According to the evaluation, the vast majority of regional funds (85%) is dedicated to the regional VET programmes, which have not been tested. In addition, there is another third of the expenditure invested in the “volet denconcentré”, that is a mixture of several forms of intervention at the local level, for which no effects are estimated.
	Garantie jeunes			+ 7.3 p.p. (36,3 %, vs 29,2) at 24 months				
	Parcours autonomie		Neutral (+ 5 in some regions and -5 in others)					Positive in Nord-Pas-de Calais (32.5% vs 28.3%) but negative in Champagne ardenne (22.8% vs 27.8%)
	Cap avenir – Cap métiers (Languedoc-Roussillon)		+ 8.5 p.p. 22.6% vs 14.1%					Positive but just short term. This measure is meant to be quite relevant within the “regional” strand.

MS	Name of the measure	Timeline of the impact (expressed as the change in probability of being employed with respect to the counterfactual situation)			by skill level or LM status			Other measured elements and notes
		Short term (1-3 month)	Medium term 6-9 month	Long term 12-month or more	LS	MS	HS	
	Ecole de deuxième chance (E2C)	-6.1 p.p. (3 - 6 moths : -7,2; 9 months : -6,8 pt)		12 months : -6,4 p.p.				Negative effect, but it is acknowledged that the TG differs (likely harder to reach individuals than those in the control group) and the goal of this measure is not directly empl
	SAS Apprentissage							What was measured here was the probability of leaving an apprenticeship rather than job placement. Results are negative overall but do not look at the employment probabilities, so not very relevant.
ES	Apprenticeships			24 months +27p.p. (72% vs 47%)	NO			Applies only to the money invested into the specific objective 8.2.2 (there are other three funded), which supports employability through VET and apprenticeships actions. It is reported that there are significant differences in the effects across regions (without further information disclosed that can be used here)
	Training			24 months +20 p.p. (58% vs 38%)				
SI	First Challenge 2015		+ 21 p.p. (79% vs 48%)		NO			There are only 8 groups for which a balanced control group could be found, according to the authors. Various doubts about the CIE design (DID). Unclear how many months after the end of the intervention this was checked
IT	YEI overall		+ 8.9 p.p. (46% vs 37%)	+ 12.3 p.p. (54.7% vs 42.4%)	NO			There's a slightly different timeframe in the evaluation as the reference point is the moment when youths are taken in charge, not the moment they leave the support. Assuming an intervention typically lasts 6 months, everything was adjusted accordingly. The values retained are those for the YEI overall, broken down by macro-region.
	YEI (North) overall		+ 4 p.p. (56% vs 52%)	+4.25 p.p. (63.4% vs 59.1%)				
	YEI (Centre) overall		+ 13.8 p.p. (47.9% vs 34.1%)	+17.1 p.p. (56.2% vs 39.1%)				

MS	Name of the measure	Timeline of the impact (expressed as the change in probability of being employed with respect to the counterfactual situation)			by skill level or LM status			Other measured elements and notes
		Short term (1-3 month)	Medium term 6-9 month	Long term 12-month or more	LS	MS	HS	
	YEI overall (South)		south: 34.6 vs 23.6 (11 p.p.)	+12.5 p.p. (42.4% vs 29.9%)				This covers approx. 60% of the young people having participated in the YEI in Italy
	YEI: traineeship		40.7 vs 37.7% (3 p.p.)	53.4% vs 43.9% (9.5 p.p.)				
	YEI: VET			31.2 vs 40.4 (-9.2 p.p.)				
	YEI: Civil service			26.9% vs 29.3% (-2.4 p.p.)				
	Friuli Venezia Giulia PIPOL (only results for those less than 30 y.o.)			+6p.p. for the entire sample, but neutral or negative for VET and +13p.p. for traineeships	7	5.5	1.5	63% of PIPOL participants are younger than 30 y.o., and the results retained are those focusing on them. The coefficient of the net effect is not statistically significant for the high skilled
LV	first job experience		20% of more working days	10% more working days				
	subsidised jobs for unemployed young people (pNPG)		10% more working days	5% more working days			NO	
	Support activity aid for self-employment or business start-ups		No statistically significant impact	No statistically significant impact				

MS	Name of the measure	Timeline of the impact (expressed as the change in probability of being employed with respect to the counterfactual situation)			by skill level or LM status			Other measured elements and notes
		Short term (1-3 month)	Medium term 6-9 month	Long term 12-month or more	LS	MS	HS	
HR	YEI		5 p.p.					They calculate the net effect also giving information on the number of participants per typology of measure. The calculated net effects vary broadly by typology of measure, (from -13 p.p. to + 27 p.p.). What is reported here is the average, weighted per number of participants. A caliper of 0.01 was used by the authors to match participants through a nearest neighbourhood matching approach.
	ESF			-24.6 p.p.				Same as above, but the authors say that in a similar study on the same measure they had found neutral effects about the measure the determines the negative effect here. Accordingly, this result should not be retained as they say it's mainly explained by the low numerosity of the sample and should be considered as neutral effect instead.
PL	ESF + YEI all three strands (mostly integrated pathways)		7.6 p.p. to 13 p.p. (lower) 17.5 pp to 20 p.p. (higher)		10.7 / 20.1	7.7 19.9	6.2 17.8	There are two results tracked: the higher of the two is about those who leave unemployment and declare to do so having found employment. The lower one results if one takes into account also “no reason mentioned”. The range is determined by the fact that they have broken down the results by length of the unemployment condition of the participants. The effect is larger for 25-29 than for 18-24 y.o. and for the LTU. Here, like in other countries stronger effects for the low skilled, despite lower gross results. No breakdown by NUTS2 although in rural areas the effects are deemed to be higher. No clear distinction between inactive and unemployed, but there is one between unemployed for the first time and having already been unemployed before.

Source: author's elaboration based on the second mandatory (counterfactual impact) evaluations carried out by Member States on the YEI and other evaluations involving support in Investment priority 8.ii

As shown in the table above, there is a broad range of estimates and related approaches. It is not straightforward to identify sufficiently standardised coefficients that can be used, especially if one considers instances such as France (where the National “macro” evaluation identifies no effects on YEI regions, and most counterfactual impact evaluations carried out on specific measures showing positive results) but also Croatia, with some strong negative results that are already flagged as unreliable by the authors but on which little background information can be gleaned.

On the other hand, some studies seem to rely on rather sound approaches, benefit from the very high numerosity of the samples they compare, and go quite in depth also on the disaggregation of result, which is a clear benefit for this analysis.

In identifying the coefficients for the analysis, some criteria have been followed. In general, longer-term net effects have been favoured over short-term ones, disaggregation by region and skill level is always retained, whereas if effects of different forms of intervention are estimated, the average for one country/programme is either calculated based on available information or based on an informed assumption. In addition, in countries, such as Italy, where the national evaluation for the YEI does not in principle cover all programmes, the results are stretched over other programmes if there is a disaggregation by typology of region. Otherwise, the data of programmes not covered is dropped. Whenever a specific evaluation insist on a single region, the coefficients from that evaluation are used.

All in all, with the above simplifying assumptions, the following coefficients are retained for the analysis.

Table 22 - Coefficients on net employment effects retained from available CIEs on YE

MS	Main coefficient	Explanation
HU	+6%	The actual net effect in p.p. is calculated based on the success rate at 6 months, multiplying the gross success rate for a coefficient that equals $1-(1/1.06)$
BG	+7.7%	Same, but the coefficient is $1-(1/1.077)$
SE	+4.5 p.p.	This is the average of the national and regional measures, considering that the share of participants is about the same in both measures based on tentative AIR 2019 data
FR	+ 5 p.p.	This is based on the fact that the two main national measures belonging to the YEI in France have effectiveness coefficients ranging between 7.3 and 10p.p. According to the Lot2 evaluation on the YEI, at least 60% of the participants up to 2017 are involved into those. As a result, the estimated effect for 60% of the population is around 8.5. There are some regional measures analysed that have worked well but others with neutral or negative effects. So one reasonable estimate is to decrease partially the success measured for the national measures, also in consideration of the overall evaluation on the support in YEI regions, which shows less favourable effects.
ES	+20 p.p.	The lower end of this range of estimates is taken as a reference as the evaluation only covers one of the specific objectives pursued by YEI activities. So a more prudent approach is chosen.
SI	+21 p.p.	This is however a rather “weak” estimate, according to the authors of the evaluation.

MS	Main coefficient	Explanation
IT	+4.25 p.p.	This applies to the northern regions, irrespective of the programme
	+17.1 p.p.	Centre
	+12.5 p.p.	South
	+7 p.p. (LS), +5.5 p.p. (MS), +1.5 p.p. (HS)	This is for Friuli Venezia Giulia, broken down by skill level, and only on those younger than 30 y.o.
HR	+4 p.p.	This is an average of the effect for the YEI and ESF, based on the information contained in the evaluation on the number of participants on which it was estimated. The effect on the ESF measures has been considered as neutral, as suggested by the authors.
PL	+15.4 p.p. (LS), +13.8 p.p. (MS), +12 p.p. (HS)	Estimates are a mid-point, by skill level, between the two results measured in the evaluation, notably the difference between treated and non-treated leaving unemployment (For any reason) and the difference between the two groups while declaring they did so as they went into employment. The authors of the evaluation declare that it is not possible to tell which of the two is most reliable.

Source: author's elaboration based on information from CIEs on the YEI (2nd mandatory evaluation) and other relevant CIEs identified on ESF IP 8.ii

In general, one can see that estimated values vary from nearly neutral effect to over 20p.p. of net effects. Highest net effects are estimated for Spain, Poland, especially for the low skilled, and Italy, especially in the central and southern/less developed regions, and Slovenia.

Yet, it also appears quite clearly that the current level of information from available evaluation does not suffice to estimate reliably coefficients that can adhere to the actual typology and volume of operations on the ground, especially due to the mix of measures that are supported (which can have very different effects) as well as techniques used to produce the results. A more “meta-analytical” approach, which however goes beyond the purpose of this research, would be necessary to improve upon the current coefficients, together with standardised instructions for MS on how to carry out evaluations and a systematic collection of information on typologies of interventions and impacts by target groups.

Nevertheless, some notable trends already appear, including:

- net effects typically larger for the low skilled or those at a broader distance from the labour market, which is contrary to what seen in terms of gross effects (success rates are negatively correlated with inactivity and lower levels of skills among participants). In two cases results are available that are already disaggregated in the evaluations published; and
- some differences between typologies of regions (see, e.g., Italy and Spain) and between cities and rural areas, which could be better exploited if disaggregated data becomes increasingly available.

It should also be recalled that some of the assumptions currently used (e.g. TO8 treated the same way as general education, with no region specific information) are clearly much more broad-cut than those used here. So, in such simulations there is an inherent need for simplifying the assumptions given the huge complexity of the policy initiative under examination.

Finally, despite CIEs not being available for many MS and provided that strong assumptions are needed in some cases, the coverage of their estimates can be considered as quite broad if

measured as the share of participants over the overall EU figure. This is notably approx. 2.9 million participants over 3.8 million participants overall (approx. 75%).

Based on these coefficients, multiplied by regionalised data on participation (by skill level whenever possible), the net number of participants who have found employment due to the ESF/YEI support for youth employment measures was produced. Results of this strand of the simulation (scenario 3) are discussed in section 3.7 below, also in comparative terms with the other simulation approaches.

3.3.7 Discounting values by LM transition data

The last step for scenario 1 and 2 was to discount monitoring data on results (both immediate and longer-term) for LM transition data, as a proxy of spontaneous dynamics. It should however be anticipated that some of the findings of the counterfactual impact evaluations (i.e. net effects larger where gross effects are smaller, although this is not always the case) mean that these two scenarios should be treated with caution. At the same time, this is probably the best one could do when CIE data is not available.

3.3.7.1 Identifying and drawing data from Eurostat (MS level)

The first step of this working package concerned identifying the right values with which to discount the changes in the Labour Market status as traced by the ESF/YEI monitoring system. As anticipated, the rationale behind this choice is that the monitoring system can only accurately report what has happened **while** the intervention was taking place, but it is (potentially a lot) less reliable in measuring what has happened **because of** the intervention. Given the insufficient coverage and availability of counterfactual impact evaluations on the ESF support, one way that is available to roughly control for the deadweight effect is to have a look at what happens in general, irrespective of the ESF support. Thus Eurostat LM transition data⁷⁹ is used as a proxy of the counterfactual (of LM spontaneous dynamics) for the three changes of LM status mapped by the ESF monitoring system, notably:

- from inactivity to unemployment;
- from inactivity to employment; and
- from unemployment to employment.

In general, the data used is MS level data (the only available) and measures quarterly transitions as a percentage of the relevant population, by gender. Values taken are averages 2015-2018, to match the time when the bulk of current ESF activity was carried out. These values are reported in the sheet 'LM transitions elab'. As for the Germany no data is available, the EU average is taken instead.

It should be considered however that:

- **there is an endogeneity problem**, as the changes of LM status measured by Eurostat may be due to the ESF or to similar policies, but there is at present no way of distinguishing between 'spontaneous' changes and 'supported' changes. It is thus assumed that ALMPs

⁷⁹ Labour market transitions - quarterly data [lfsi_long_q]

have a sufficiently small influence on changes of LM status, but this this may be a wrong assumption. The values of LM transitions are experimental;

- **there is no breakdown by skill level**, thus the same rates are applied to the entire population (but there's likely a downward bias as the ESF tends to focus on harder to reach individuals, which are likely to linger in their LM status a lot more than the remainder of the population);
- **the current simulations treat the unemployed as a homogeneous ensemble with no breakdown for the long term unemployed**. This, again, may exert downward pressure on the discounted figures, as significant shares of the ESF participants are long term unemployed;
- some beneficiaries may interpret the requirements of the ESF monitoring system strictly and already draw off some deadweight; and
- there is no breakdown by NUTS2 region, but only MS data.

One way of addressing such issues is to consider discounted values as a pavement, while the non-discounted values as a ceiling, so to define a range of possible impacts of the ESF. The issue may be particularly relevant in some MS, where the dynamic of LM transitions is quite strong. In general, this is especially meant for comparison with Scenario 3. It can also help us appraise the “scale” of the potential impact, if not a reliable value.

3.3.7.2 Discounting non zero values for the LM transition data

Once the right data from Eurostat was identified, the subsequent step was to reduce pro-quota values for the result indicators identified. To note that:

- first, it was calculated the final success rate for the five changes of labour market status under examination (inactive to unemployed, inactive to employed (both immediate and longer term), unemployed to employed (both immediate and longer term); and
- second, this success rate was reduced by the ‘deadweight’, that is the share of ‘spontaneous’ LM transitions as per Eurostat data.

A coefficient was calculated that equals the ‘counterfactual impact’ and corresponds to the (success rate – deadweight)/success rate.

In case of negative values (where the deadweight is higher than the ‘gross success rate’, the coefficient was capped at 0, assuming that the support received should not have a negative impact on single individuals). This coefficient then multiplied all the relevant observations.

The final results are described in section 3.4 below, together with those from scenario 3 (based on the coefficients form CIE).

3.4 Discussing what the estimated data (input for RHOMOLO) looks like

This section aims to compare the results estimated in the different scenarios described in table 10 (section 3.2.3) with a view to understanding their differences, the implications in terms of their reliability and the extent to which the assumptions used can bias the estimates. It should be reminded here just like elsewhere that the interest is not in the values per se (given these are tentatively based on AIR2019 but not yet validated and only used for exemplificative purposes)

but about understanding whether any of the approaches presented within this research could be built upon in future research carried out with RHOMOLO. Once more, it should be reminded that the benchmark approach is based on the assumption used in the ex-post evaluation of the European social fund 2007-2013 and in the recent impact assessment on the Human Capital Investments in which RHOMOLO dealt with investment in youth employment. To briefly recall the main limitations of that approach: (i) the effects were estimated based on an intensity of support offered across regions which was equal to the original allocations – thus assuming no difference between the original allocation and the actual implementation; (ii) investment is treated as a standardised general education policy, so its effectiveness is borrowed from the literature on education rather than on actual micro-data on the effectiveness of active labour market policies within the ESF/YEI 2014-2020 (iii) the time profile of the investment (distribution across 2014-2023) is drawn from info of the previous programming period and not based on the current one.

Now, to discuss the limitation of the approach presented in this research, first of all a distinction should be drawn between data on participations and data on labour market changes generated by the YE support, which are discussed separately in the sections below. This is a key difference in that, on the one hand, information was missing as to the volume of the support offered (how many people, in which regions, with what background features) and, on the other hand, information is needed on the effects of the support on them. As described in section 3.2.3 the underlying difficulties are rather distinct between the two.

3.4.1 Data on participation

As to data on participations, three main issues had to be tackled, namely:

- the distribution over time (yearly);
- the geographical distribution (NUTS2 level); and
- the distribution of participants' background features by labour market status and educational level.

With respect to (i) **identification of a proxy of progress over time**, this clearly remains complex as there is simply no bullet-proof way of understanding how much more will be produced in the future based on current data. However, the quali-quantitative analysis shows that the expenditure rate⁸⁰ in 2018 seems to be the available data which more strongly correlates to progress, together with a measure of progress towards the targets of output indicators (achievement rate of output indicators). Their respective advancement is not so different if one looks at the EU average (at approx 38% and 56%) and in general, one could consider that the expenditure rate slightly underestimates the volume of activities actually implemented (especially when it comes to the number of participants in partly implemented operations, as these can be registered before the expenditure is reported by beneficiaries) whilst the achievement rate slightly over estimates actual implementation, due to the fact that MAs might be placing stronger emphasis on those actions for which a target is set, leaving somewhat behind the rest. The two measures of progress could be also used in combination to develop two

⁸⁰ Share of eligible expenditure declared by beneficiaries to the Managing Authorities over the total allocation

scenarios, to act as the pavement and ceiling for the distribution of output and results over time. At any rate, the difference will decrease year after year as all available allocation is put to use towards the end of the programming period.

As to (ii), the **geographical distribution by NUST2 level**, preliminary data shows that the share of expenditure declared in operations that are measured at the NUTS3 or 2 level is around 36%, with an additional 9% at the NUTS1 level. For the latter and the remaining 54%, that is only recorded at the national level, it should be recalled that available information on expenditure is nevertheless broken down by typology of region, which allows for a further disaggregation of the expenditure (so it is not just NUTS1, but NUTS1 by typology of region, and the same applies to NUTS0). Finally, the proxy based on allocation data which was used to further assign output and results within the same NUTS1/0 and typology of region, should work well and is already an improvement over previous approaches to the simulation (as it uses a combination of programming and implementation data). Overall, the regionalisation approach shows potential to adhere a lot better than in previous attempts to the reality of the ESF/YEI implementation, as it shifts the focus from programming to actual implementation. The shares of expenditure by MS and NUTS level at which they are reported are displayed below.

Table 23 – share of eligible expenditure declared by beneficiaries, by NUTS level and MS

MS	NUTS0	NUTS1	NUTS2	NUTS3
BE	0.0%	100.0%	0.0%	0.0%
BG	100.0%	0.0%	0.0%	0.0%
CY	100.0%	0.0%	0.0%	0.0%
CZ	0.0%	0.0%	0.0%	100.0%
DE	0.0%	66.0%	0.1%	33.9%
ES	100.0%	0.0%	0.0%	0.0%
FR	29.8%	0.0%	69.9%	0.3%
GR	0.0%	0.0%	100.0%	0.0%
HR	100.0%	0.0%	0.0%	0.0%
HU	0.0%	0.0%	0.0%	100.0%
IE	100.0%	0.0%	0.0%	0.0%
IT	75.4%	0.0%	24.6%	0.0%
LT	100.0%	0.0%	0.0%	0.0%
LU	100.0%	0.0%	0.0%	0.0%
LV	0.0%	0.0%	100.0%	0.0%
MT	0.0%	0.0%	0.0%	100.0%
PL	41.1%	0.0%	58.9%	0.0%
PT	0.0%	0.0%	94.0%	6.0%
RO	100.0%	0.0%	0.0%	0.0%
SE	100.0%	0.0%	0.0%	0.0%
SI	0.0%	9.2%	90.8%	0.0%
SK	0.0%	100.0%	0.0%	0.0%
UK	0.0%	69.7%	27.3%	3.0%
Overall	54.5%	9.1%	30.7%	5.8%

Source: author's elaboration based on tentative AIR2019 Table 7

In terms of geographical distribution, the approach seems generally suitable for small countries (with just a few NUTS2 regions) and should work well also for some large countries with rather detailed data (e.g. Greece, France, with over 70% of the expenditure at the NUTS2, but also Portugal, Poland and some countries with good NUTS1 level data). Most likely, better data should be sought for in countries such as Spain and Italy, which have several NUTS2 region each and little information on expenditure at that level. This could be done in future revisions of the research for instance in the case of the national YEI OP in Italy, that has a dedicated monitoring system which allows understanding, with some approximation, how many people have participated to the interventions for each region.

Finally, with respect to (iii), the conditional distribution of participations per labour market status and skill level, also in this case it is difficult to find a benchmark that is based on other modelling approaches. The multi-level fixed effects model with fixed and random country level effects used has typically proven to be a good fit in terms of significance of the regressions (Wald Chi) and regressors used, despite the low numerosity of the sample of observations in the database. It is also a good model in theory, as it can adhere to the nested structure of the data in the ESF/YEI and to the different approaches that are followed by MS in identifying their target group, whilst examining some common background features of the participants of the ESF (i.e. typically skills are related to the labour market status). Whilst the distribution of participants by MS, LM and skill level is already reported in section 3.3.3.4 (and not replicated in additional tables here), it remains to be said that the more micro-data becomes available, the easier it will become to test the reliability of the estimates.

3.4.2 Data on changes of labour market status

Differently from the case of participations, which are obviously just generated by the ESF/YEI support, changes of occupational status can be either a result of the support received or just spontaneous dynamics, i.e. changes that would have occurred to participants even if they had not been supported. Because other general equilibrium effects such as alternative ways to employ the money, displacement and substitution effects or other spillovers are left to RHOMOLO to be determined in future research, here the key question is that of the net effects on participants, in line with the framework approach described in 3.2.3. It is thus a question of identifying to the extent possible the partial equilibrium effect in terms of changes of occupational status of ESF/YEI participants to YE support. As explained earlier, three scenarios are considered which can be summarised as:

- (1) Net effects on the participants = share of participants reporting to be in employment upon completion of the intervention – share of inactive(unemployed) who change labour market status according to LM transition data
- (2) Net effects on the participants = share of participants reporting to be in employment at 6 months of exiting from support - share of inactive(unemployed) who change labour market status according to LM transition data

(3) Net effects on participants = number of participants * net effect⁸¹ of the intervention as estimated in CIEs

In addition to these, one additional measure of success is estimated, namely inactive participants starting to look for a job, but, given there is no longer term result nor CIE study on this, only the first scenario (immediate results discounted by LM transition data) can be used to that end.

The starting point is to have a look at the distribution of results by MS according to the different scenarios, as shown in the Table below.

Table 24 – overall employment results measured as per the different monitoring indicators and scenarios, by MS

MS	Gross results/effects		Net results/effects		
	In empl (immediate)	In empl (6 months)	Scenario 1 (in empl immediate)	Scenario 2 (in empl 6 months)	Scenario 3 (in empl CIE)
BE	13,083	14,154	2,001	1,932	
BG	13,657	24,797	19,445	15,977	976
CY	967	1,534	118	-	
CZ	1,806	2,105	863	778	
DE	7,814	7,322	1,570	2,045	
EL	6,509	13,645	13,251	8,525	
ES	149,794	124,772	393	14,390	151,434
FR	131,234	160,575	58,903	40,658	24,686
HR	14,179	21,011	11,221	7,947	1,355
HU	42,469	36,941	2,059	9,797	2,404
IE	1,604	439	-	1,303	
IT	157,699	206,368	133,780	121,117	106,047
LT	11,577	9,293	35	313	
LU	674	871	4,316	551	
LV	5,521	10,349	231	149	
MT	411	411	16,641	9,681	
PL	289,171	295,020	207,810	200,972	61,718
PT	25,381	32,982	11,288	8,101	
RO	32	32	-	-	
SE	15,740	13,449	537	1,054	1,956
SI	3,373	3,178	78	53	2,727
SK	31,434	28,380	13,649	14,692	
UK	28,590	34,051	2,636	1,289	
Tot for CIE MS	817,316	886,111	434,227	411,965	353,303
Tot	952,719	1,041,679	500,825	461,326	353,303

Source: author's elaboration based on tentative AIR2019 data CIE studies and Eurostat LM transition data

As it can be seen the rough amount of people who found employment as a result of their participation to the ESF/YEI support in YE according to the three scenarios developed isn't excessively different. The three scenarios share a similar distance with the pure monitoring values of result indicators (not discounted by deadweight, i.e. employment results which would have occurred in absence of the support), just about halving their values. However, if one looks

⁸¹ The coefficient use is the net increase (p.p.) of probability of being employed in the medium term (6 to 24 months, depending on the availability of estimates) for the treated.

into their geographical composition, this varies broadly. The charts below compare the three “net effects” scenarios one another by displaying their values as a share of the number of people reporting to be in employment at 6 months of the intervention (gross result), by MS. So 100% is in case no deadweight is measured, whereas lower values mean that the deadweight is higher.

Figure 10 – net employment effects as a share of participants declaring to have found employment at 6 months of exiting the support, by scenario (1 left, 2 centre, 3 right)

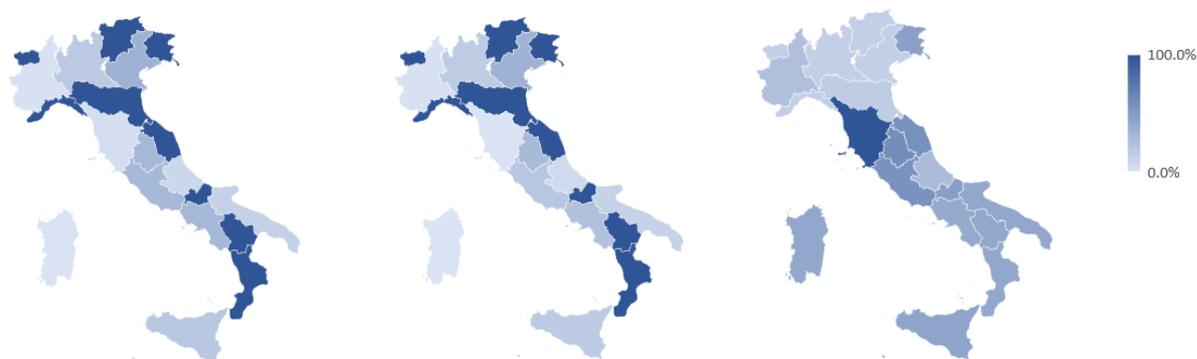


Source: author’s elaboration based on tentative AIR2019 data CIE studies and Eurostat LM transition data

One can see that whilst Spain displays much stronger measured effects on the employment levels of participants in scenario 3 (CIE) than in 1 and 2 (discounted monitoring data), the opposite holds true for countries such as France, Poland and Bulgaria. Hence the cross country variation appears to be rather strong, and this is due to the fact that the finding of counterfactual impact evaluation vary significantly across countries. The main problem there is whether this could be considered as a measure of actual differences or just due to bias in the research designs.

In addition, even if one took as a reference a single MS showing relatively comparable cross-scenario values such as Italy, inter-regional variance would remain marked, as shown in the Figure below.

Figure 11 - net employment effects as a share of participants declaring to have found employment at 6 months of exiting the support, by scenario (1 left, 2 centre, 3 right). Focus on NUTS2 regions in Italy



Source: author’s elaboration based on tentative AIR2019 data CIE studies and Eurostat LM transition data

This is the likely result of the fact that the Italian CIE showed much stronger net effects for regions in the central and southern/insular part of the country, whereas monitoring data

discounted for LM transition values depicts a very different scenario, with gross effects stronger in the north.

In addition to the differences in the total net employment effects on the treated, one should give attention also to their distribution per participants' background features, as shown in the charts below. These display the share of employment effects for the inactive and the share of effects for low skilled, respectively on the overall effects measured, comparing these across scenarios. This time, the gross results too are included in the figure, to appraise how they compare with those in which deadweight effects are (in principle) removed.

Figure 12 – share of inactive on participants in employment, in order gross results upon completion and at 6 months, net results upon completion and at 6 months (scenario 1 and 2) and net results based on CIE (scenario 3), by MS



Source: author's elaboration based on tentative AIR2019 data CIE studies and Eurostat LM transition data

As it can be seen, also the variation in terms of composition is rather strong, especially for countries such as Spain, but also Sweden, France and Italy. Generally, an increasing incidence of inactive for whom the intervention had a positive employment effect over unemployed can be found in scenarios 1 to 3 (where the estimated deadweight effect is removed from the estimates), as compared to the gross results. This is also because according to LM transition data inactive switch to employment more seldom than unemployed. In addition, when compared with the estimations done with the CIE, it should be recalled that the latter does not use predicted shares of inactive participants having found a job, but the CIE coefficients are just multiplied by data on participation. And, because inactivity is negatively correlated with

success rates in scenario 1 and 2, the gross success rate for the inactive is estimated as lower on average than that of the unemployed, whilst this is not the case in scenario 3. Regional variation is strong too, though not presented for conciseness.

Figure 13 - share of low skilled on participants in employment, in order (from left to right and from the top downward): gross results upon completion and at 6 months, net results upon completion and at 6 months (scenario 1 and 2) and net results based on CIE (scenario 3), by MS



Source: author's elaboration based on tentative AIR2019 data CIE studies and Eurostat LM transition data

Similarly to the example above on the share of inactive participants over aggregated effects for the inactive and unemployed, also the skill composition of those finding employment tends to vary geographically depending on the modelling approach followed, but the share of low skilled increases significantly (across all regions) only when counterfactual impact evaluations are taken into account (scenario 3). This is most likely due the fact that in comparison with scenario 1 and 2, labour market transitions tended to overestimate the spontaneous LM transitions of the low skilled, thus overestimating the deadweight effect for the low skilled and underestimating, in turn, the net effects. In addition, the net effects estimated in the Italian CIE are higher in regions of Italy where the number of low skilled is high and both in Friuli and Poland the low skilled had distinctively better net effects than the average population. So, most likely, differences would be even more pronounced if more disaggregated results from CIE were available, in line with the finding from the literature on ALMP (see for instance Card and Kluge, 2018) suggesting that net effects are higher for those at a certain distance from the labour market.

All in all, one could conclude that when it comes to the rough size of the partial equilibrium net effects on participants, also scenario 1 and 2 could, with some approximation, come closer to

the “actual” net effects than a simple look at gross results from monitoring indicators. However, when it comes to the distribution, especially the regional distribution of effects, scenario 1 and 2 risk to carry far too much bias. Unfortunately, estimates from scenario 3 based on CIE too should be considered with great caution lacking both an in-depth quality assessment of the CIEs used, sufficient granularity of the results presented (especially for a regionalised analysis, if not for Italy) and heterogeneity of the approaches. So strong dependency from the assumptions is still present. But further research efforts and a broader use of counterfactual impact evaluation focusing on the heterogeneity of effects could easily improve upon this. More importantly, despite the limitations, this approach seems to already produce a much more nuanced picture of the activity supported by ESF/YEI than what would be generated using the old RHOMOLO assumptions on ESF/YEI being a general education policy and not information on the actual implementation.

3.5 Concluding remarks

In section 3.1.1, it was clarified that evidence on the overall long-term effects of cohesion policy remains at best mixed. Furthermore, it is often untimely (due to the (s)low availability of detailed financial data on implementation) and frequently not place-based (just averages at the EU-level calculated). The macroeconomic theory behind some of the macro-econometric simulations is also quite debatable (decreasing returns of scale, model à la Solow) especially if investments in HC are considered. Concerns have been raised about the fact that results are highly model or assumptions dependent and that it would be necessary to better bridge micro and macro evaluation studies, especially for broad policy initiatives such as the cohesion policy. Quantitative evidence on the overall effects of the ESF and YEI in particular is still very limited, if not absent at all. Some recent notable examples include the results stemming from the JRC’s computable general equilibrium RHOMOLO model which were used in the recent Impact Assessment on the Human Capital Investments, carried out by DG Employment in the context of the new proposals for the funds for the Multiannual Financial Framework 2021-2027. The RHOMOLO model, consistently used in the most notable evaluations and visibility material of the funding initiatives related to the cohesion policy and promoted by DG Employment and DG Regio, arguably comes with several practical advantages with respect to the modelling of policy impacts within the frame of cohesion policy. These include the fact that (i) it produces a range of results for many macroeconomic variables, (ii) it presents disaggregated effects by NUTS2 region (place-based) and over time, (iii) it sheds light on the logical link between the policy under examination and its effects and, (iv) particularly for human capital policies, it includes an articulated labour market module which accounts for labour market rigidities, changes in the composition of the labour market, skill migration and so forth. Nevertheless, as clarified in section 3.1.2, some of the assumptions used both in the context of the ex-post evaluation of the 2007-2013 programming period and in the impact assessment of the Human Capital Investments cannot match the logic and typology of initiatives funded under the ESF/YEI, nor the assumable national and regional heterogeneity in the effectiveness of the support offered.

Against this background, in this research a process was defined in section 3.2 and showcased in section 3.3 that aims to improve the provision of data on which RHOMOLO bases its

macro-simulation approach, in general for the ESF/YEI but in particular within the frame of Youth Employment policies.

The rationale is that of trying to bridge the gap between micro-level and macro level evidence in line with Pellegrini et. Al (2013), relying on ESF/YEI monitoring data cleaned by the deadweight effect to produce input data (net – partial equilibrium - effects on participants, by NUTS2 region, LM status and skill level to the extent possible) that is then elaborated, accounting for general equilibrium dynamics, through RHOMOLO, in line with what suggested in Sakkas (2018). This partial equilibrium data (changes to the labour force) is suitable to be used as an exogenous shock to RHOMOLO (the labour force is a shift and share parameter of the model) which then affects its endogenous mechanisms and variables.

As results from RHOMOLO are listed as a mandatory source for the Mid-Term Evaluations of the different thematic objectives of the ESF/YEI, and most notably in the context of the Youth Employment and TO8 evaluations, the data elaborated within this research (section 3.3) are being shared with the JRC and are feeding into those. In fact, the methodological approach described in 3.2.2 and 3.2.3 was devised in collaboration with the JRC, who are already using regionalised monitoring data on participants (produced within this research with their estimated conditional distribution by LM status and Skill level) and their results to improve the simulations, together with data on the typologies of interventions collected within the frame of the studies commissioned by DG EMPL (which is not included in this research as it is funded).

In the previous sections, several obstacles were described along with some envisaged solutions and the implications these carry along, and more will be schematically described in the next “strength and weaknesses” section.

But in essence, the current level of knowledge on the composition of the youth employment support impedes to reliably infer what are the net employment effects for participants to youth employment support in all countries and regions, by Labour Market status and skill level. This is chiefly due to the vast variety of forms of support offered, differences in participants’ features as well as diversity in the methodological approaches to their evaluation. However, with some assumptions and approximation, it is possible to get a sense of how things are proceeding and of the volume of effects generated for the participants, especially in some countries and regions. Most importantly, there seems to be significant room for improving the estimates, as more in-depth research is carried out on available counterfactual impact evaluations, information on typology of interventions is collected and micro-data begins to become available.

3.5.1 Strengths of the approach and how this improves upon the literature

As clarified in section 3.1.1, the three main approaches beside general equilibrium models which have tested the effects of active labour market or cohesion policy have significant shortcomings when it comes to discussing the overall effects of the ESF/YEI. In particular:

- (i) micro-level counterfactual impact evaluation (CIE) typically assume the absence of spillover effects and interactions among agents, which violates the intervention logic of EU funds. As a result, their results are “partial equilibrium” in the sense that only focus on those receiving support whilst ignoring both what happens to the context (non-recipients of the policies) as well as alternative ways of using the intervention’s

money. This is for instance clarified in the meta-analysis of over 200 recent studies of Active Labour Market Policies from Card and Kluve (2018), which states that “these studies are at best partial equilibrium”, as also stressed by Crepon et Al. (2013). In addition, external validity of such approaches is typically limited, which reduces aggregability at the programme or Member State level;

- (ii) macro-econometric models (see for instance the survey from Dall’Erba & Fang, 2017) tend to: (a) focus on cohesion policy as a whole, (b) be based on a growth model à la Solow (not very suitable for human capital development investments), (c) ignore the differential contribution of specific typologies of interventions (or even specific funds) and (d) focus on the impact on GDP only; in addition, (e) there is no explanation of the logical link between the intervention and its effects, no regionalised estimates and limited or absent effects heterogeneity; and
- (iii) recent macro-level counterfactual impact evaluations (based on RDD) face some external validity issues (it is a local average treatment effect for less developed regions close to the cut-off, see for instance Pellegrini et al, 2013) and, again, only focus on average GDP impacts and assume away effects’ heterogeneity.

It was also highlighted that a fourth approach exists, which is that of general equilibrium models such as JRC’s RHOMOLO. However, as per Sakkas (2018), the most recent attempts at capturing the effects of the ESF need to rely on strongly simplifying assumptions.

Against this backdrop, the main strength of this approach is that of suggesting a framework that seeks to bring together knowledge on the effects of ALMP policies, on the current structure of the RHOMOLO model and on the monitoring and evaluation system of the ESF/YEI in an integrated manner. As anticipated, this should aid a much better informed estimation of the macroeconomic effects of the ESF/YEI through RHOMOLO, with significant improvements upon the current assumptions. Input data produced according to this approach can in particular shift the focus from just programming data (as it was in the past) to actual implementation data when it comes to both the “where” (NUTS2) and “whom” (participants background features) dimensions of evaluation. In addition, once used in combination with the RHOMOLO model, this approach is suitable to generate timely information on the ESF/YEI impacts that is: (i) disaggregated by region; (ii) focuses on the channels through which the policy intervention pursues its goals; (iii) takes into account participants heterogeneity and (iv) can account for general equilibrium effects such as spillovers, displacement, substitution effects etc. Finally, in section 3.2.1, other modelling strategies through RHOMOLO are outlined that can be used in the future to test from different angles a given HC policy, thus decreasing model-dependency (sensitivity analysis).

Importantly, this approach addresses directly the need of detailed data on the micro-level effects on the labour force generated by the ESF/YEI as identified by Sakkas (2018) in the latest simulations carried out through RHOMOLO, whilst making a first step into addressing Pellegrini et al. (2013) concern about the need to better integrate the micro and macro level of the evaluation of cohesion policy.

3.5.2 Weaknesses and room for development

However, also in the light of the complexity of the issues at stake, several weaknesses appeared too, and the main ones are recalled below:

- according to the methodological framework proposed in this research, a limited use of the endogenous mechanisms of the RHOMOLO model is currently envisaged, as explained in section 3.2.1. Data on the net effects on the employability of participants are in fact used to change exogenously “shift and share” parameters of the model, which will then determine how such a shock enters multiplicatively in the model to affect equilibrium levels (including determining displacement or substitution effects). However, the labour market in RHOMOLO is necessarily a reduced-form representation of the reality and some coefficients of the wage curve are just assumed equal for all regions. This is however a limit of RHOMOLO which applies also to other modelling strategies and could potentially be addressed in the future;
- assumedly, the econometric reconstruction of the background features of participants can only roughly replicate the actual structure of the micro-data, although there is no straightforward measure available of how close (far) the estimates presented fall from it. The box plots on estimated data (section 3.3.3.6) show a reasonable degree of deviation of estimated aggregated data from the actual aggregated data, which is an indirect measure of plausibility of the estimates. In addition, nearly all estimated figures are based on statistically significant coefficients. Of course, tests should be carried out leveraging on micro-data as this becomes available even in just a few programmes, to understand whether the current approach identifies good proxies to fill some gaps. This is particularly relevant as it might be that in some MS micro data will be easy to collate, while in others the form of their monitoring systems means difficulties in aggregating information from such detailed sources. So, a mixed approach (estimated data for some countries and actual data for others) might be an actual option for future research;
- the coefficients for the deadweight effect used for scenario 1 and 2 and based on labour market transition data can only roughly indicate the scale of the effects at the EU level, but due to its limited level of disaggregation and given endogeneity issues, can hardly be considered as a sound proxy of the deadweight effect. In terms of the distributional impact (by region, by target group), its reliability is possibly even lower. This might however be mitigated using proxies of deadweight which take better into account the distance from the labour market of the different target groups addressed;
- the coefficients borrowed from the 2nd mandatory evaluation of the YEI screened (and related evaluations) might carry bias due to model-dependency and because their scope (timeframe, typology of operations and target group covered) in some cases may not reflect the full extent of the support provided. This has created the need for assumptions being made about which coefficients to actually use which can themselves affect the estimates. In addition, the lack of granularity (effects by NUTS2 region, LM status and skill level) significantly hampers the suitability of the estimates for the RHOMOLO model. This can however be improved in the future as more CIEs become available and if action is taken

from the European Commission with respect to the standardisation of the studies and the request for disaggregated results.

In connection with the weaknesses described above, the following developments are suggested that would build upon the current methodological framework and increase its reliability:

- to increase accuracy in the identification of coefficients for net-effects of YE measures, by:
 - fully fledged meta-analysis on the effects of the YEI, including if possible getting in touch with the authors of all studies to retrieve estimates of their studies by NUTS2 regions, LM status and skill level;
 - to extend the literature review on the effects of ESF funded YE measures, to draw from additional CIE sources and possibly cover more countries;
 - a dedicated counterfactual (ideally promoted by DG EMPL) that employs a consistent estimation approach to multiple countries and aims to retrieve information broken down as highlighted;
 - as information on the composition of support (typologies of intervention, also from the Mid-Term Evaluation currently being carried out) becomes available, to fine tune some of the assumptions on the coefficients borrowed from the CIEs, better taking the diversity of effects into account;
- to improve the estimation of participants' background features, by:
 - requesting and employing to the extent possible micro data from the ESF/YEI monitoring system; and
 - exploiting existing data from dedicated reporting activities which are based on micro-data (see for instance yearly monitoring on the YEI in Italy, or some evaluations).

3.5.3 Policy recommendations

This research activity has also highlighted some gaps in the ESF/YEI monitoring and evaluation systems. Some of them have already been described in Chapter 2 but are here briefly recalled to the attention of those in charge of the monitoring and evaluation system of the ESF/YEI:

- there is a need to streamline CIEs, i.e. to ask Managing Authorities commissioning such studies to introduce in their requests for services that results of counterfactual impact evaluations should be broken down, whenever feasible, by target group, especially by level of skill. Effect heterogeneity plays a key role in understanding what works and for whom, so this must be an explicit aim of each evaluation. In addition, evaluation studies should always come with a reference to the typology of interventions they are examining. Finally, the outcome variables should also be streamlined (i.e. probability of being in employment at a given point in time). This would greatly facilitate the creation of a map of partial equilibrium effects which can then feed into macroeconomic modelling to generate evidence on spillovers.
- there is a need to invest in collating and streamlining micro-data available, which should at least distinguish by skill level and labour market status, but also by form of intervention to which participants have been exposed. Without this background information it is very hard

to understand whether the right people are being addressed and take into consideration how this impact the economy as a whole. Even if micro-data proved too hard to be collated, it would be important to standardise at least information on typologies of intervention, even in case micro-data were not available. This would allow using multiple modelling approaches to the analysis of the impacts of ESF investments;

- there is a need to clarify what it means to “gain a qualification” in the ESF/YEI monitoring system. At present, this measure is not meaningful for policy analysis as it doesn’t inform on the extent of the upskilling. It would be important to link this especially to ISCED levels, as at the moment this indication cannot be used to effectively measure success of the supported interventions; and, finally
- to think of some strategies to understand the effect on the participants’ productivity levels, or encourage/promote dedicated counterfactual impact evaluations to this end which could leverage on administrative sources or survey approaches. This would allow different modelling approaches to the evaluation of human capital development, in line with the literature on the measurement of human capital through individual productivity.
- to consider the need of tools and studies which take more openly into account the interplay between the support provided and its implementing environment, to consider the materialisation of general equilibrium effects, especially displacement effects as it might often be the case of active labour market policies. In lack of dedicated guidance and information on the issue this is an issue which seems to remain overlooked by policy analysts working on assessing the effects of cohesion policy.

3.6 Bibliography

Barca F. (2009), An agenda for a reformed Cohesion Policy: Independent report prepared at the request of Danuta Hübner, the Commissioner for Regional Policy. Available at: http://ec.europa.eu/regional_policy/policy/future/barca_en.htm

Becker S.O., Egger P.H. and Von Ehrlich M. (2010) Going NUTS: The Effect of EU Structural Funds on Regional Performance, *Journal of Public Economics*, 94, 578-590. Available at: http://www.parisschoolofeconomics.eu/docs/koenig-pamina/going-nuts_beckereggervonehrlich.pdf

Becker S.O., Egger P.H. and Von Ehrlich M. (2013) Absorptive Capacity and the Growth and Investment Effects of Regional Transfers: A Regression Discontinuity Design with Heterogeneous Treatment Effects, *American Economic Journal: Economic Policy*, 5, 4, 29-77. Available at: <https://www.aeaweb.org/articles?id=10.1257/pol.5.4.29>

Bondonio D. (2016), The impact of varying per capita intensities of EU Funds on regional growth: Estimating dose response treatment effects through statistical matching - Work package 14d: Propensity score matching. Ex post evaluation of Cohesion Policy programmes 2007-2013, focusing on the European Regional Development Fund (ERDF) and the Cohesion Fund (CF), European Commission, April 2016. Available at: <https://op.europa.eu/en/publication-detail/-/publication/afb88bdf-2303-11e6-86d0-01aa75ed71a1/language-en>

Card, D., J. Kluve and A. Weber (2018), What Works? A Meta-Analysis of Recent Active Labor Market Program Evaluations, *Journal of the European Economic Association* 16(3). Available at <https://academic.oup.com/jeea/article-abstract/16/3/894/4430618?redirectedFrom=fulltext>

COMMISSION IMPLEMENTING DECISION of 18 February 2014 setting out the list of regions eligible for funding from the European Regional Development Fund and the European Social Fund and of Member States eligible for funding from the Cohesion Fund for the period 2014-2020 (notified under document C(2014) 974)

Crépon, B., Duflo, E., Gurgand, M., Rathelot, R., and Zamora, P. (2013) “Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment *.” *The Quarterly Journal of Economics* 128, no. 2 (May 2013): 531–80. <https://doi.org/10.1093/qje/qjt001>.

Crescenzi, R. and M. Giua (2017): “Different approaches to the analysis of the EU Cohesion Policy. Leveraging complementarities for evidence-based policy learning,” in Bachtler J., Berkowitz P., Hardy S. and Muravska T. (eds) *EU Cohesion Policy. Reassessing performance and direction* (NY: Routledge), pp. 21-32. Available at: <https://www.taylorfrancis.com/books/e/9781315401867>

Dall’Erba, S. and F. Fang (2017), “Meta-analysis of the Impact of European Union Structural Funds on Regional Growth”, *Regional Studies* 51, 822–832. Available at: <http://hdl.handle.net/10.1080/00343404.2015.1100285>

Darvas Z., Mazza J. and Midoes C. (2019), How to improve European Union cohesion policy for the next decade, Bruegel Policy Contribution No. 8, available at: https://bruegel.org/wp-content/uploads/2019/05/PC-08_2019.pdf

Di Comite, F., Kanacs, d'A., Torfs, W. (2015) Macroeconomic Modelling of R&D and Innovation Policies - An application of RHOMOLO and QUEST, JRC Technical Reports, p25 available at: <http://publications.jrc.ec.europa.eu/repository/bitstream/JRC89558/Ifna27084enn.pdf>

DG EMPL, country factsheets on the Updated ESF Ex-post evaluation 2007–2013 Synthesis Report, available at: <https://ec.europa.eu/social/keyDocuments.jsp?advSearchKey=ESFker0713&mode=advancedSubmit&langId=en&policyArea=&type=0&country=0&year=0>

DG REGIO (2016a), The impact of cohesion policy 2007-2013: model simulations with Quest III – Work Package 14a Ex post evaluation of Cohesion Policy programmes 2007-2013, focusing on the European Regional Development Fund (ERDF) and the Cohesion Fund (CF), European Commission, available at: https://ec.europa.eu/regional_policy/sources/docgener/evaluation/pdf/expost2013/wp14a_final_report_en.pdf

DG REGIO (2016b), The impact of cohesion policy 2007-2013: model simulations with RHOMOLO – Work Package 14b Ex post evaluation of Cohesion Policy programmes 2007-2013, focusing on the European Regional Development Fund (ERDF) and the Cohesion Fund (CF), European Commission, available at: https://ec.europa.eu/regional_policy/sources/docgener/evaluation/pdf/expost2013/wp14b_final_report_en.pdf

Escudero, V. (2014) Are active labour market policies effective in activating and integrating low-skilled individuals? An international comparison, MPRA Paper No. 55319, posted 16. April 2014 03:56 UTC available at: <https://mpra.ub.uni-muenchen.de/55319/>

European Commission (2016), Ex-post evaluation of the 2007-2013 ESF Programmes, SWD (2016) 452 Final, p. 35. Available at: https://ec.europa.eu/jrc/sites/jrcsh/files/swd2016_452_final_-_ex_post_evaluation_of_the_european_social_fund_2007-2013.pdf

FGB, Ceps and COWI (2016), The analysis of the outcome of the negotiations concerning the Partnership Agreements and ESF Operational Programmes, for the programming period 2014-2020, European Commission, DG Empl, Luxembourg: Publications Office of the European Union, 2016 ISBN: 978-92-79-62769-9 doi: 10.2767/90132, available at: <https://op.europa.eu/en/publication-detail/-/publication/b2c01d15-ffef-11e6-8a35-01aa75ed71a1/language-en>

FGB, Metis, Applica and COWI (2018b), Study supporting the update of the data reported in the 2007-2013 ESF ex-post evaluation, European Commission, DG EMPL, Publications Office of the European Union, 2018 ISBN: 978-92-79-93945-7 doi: 10.2767/64328

FGB, Metis, Applica and Ockham (ongoing), Study supporting the evaluation of employment and mobility, European Commission, DG EMPL

Hagen T., Mohl P. (2009) Econometric evaluation of EU Cohesion Policy: A survey. Discussion Paper 09-052. ZEW, Centre for European Economic Research, Mannheim, available at: <http://ftp.zew.de/pub/zew-docs/dp/dp09052.pdf>

Lecca, P., Barbero, J., Christensen, M., Conte, A., Di Comite, F., Diaz-Lanchas, J., Diukanova, O., Mandras, G., Persyn, D. and Sakkas, S. (2018). RHOMOLO V3: A spatial modelling framework. JRC Technical Reports JRC111861, EUR 29229 EN. Publications Office of the European Union, Luxembourg. Available at: https://publications.jrc.ec.europa.eu/repository/bitstream/JRC111861/jrc111861_rhomolo_may_2018_final_2.pdf

Lee D. S. (2008). "Randomized Experiments from Non-random Selection in U.S. House Elections", *Journal of Econometrics*, Vol. 142, N.2, pp. 675-697. Available at: <https://www.princeton.edu/~davidlee/wp/RDrand.pdf>

Marzinotto, B. (2012), 'The growth effects of EU Cohesion Policy: a meta-analysis', Bruegel Working Paper 2012/14, October 2012. Available at: <https://bruegel.org/2012/10/the-growth-effects-of-eu-cohesion-policy-a-meta-analysis/>

Mercenier, J., Álvarez-Martínez, M., Brandsma, A., Di Comite, F., Diukanova, O., Kancs, d'A., Lecca, P., López-Cobo, M., Monfort, Ph., Persyn, D., Rillaers, A., Thissen M., and Torfs, W. (2016). "RHOMOLO-v2 Model Description: A spatial computable general equilibrium model for EU regions and sectors," JRC Technical reports JRC100011, European Commission, DG Joint Research Centre, EUR 27728 EN, doi:10.2791/184460, p. 14 - available at: https://ec.europa.eu/jrc/sites/jrcsh/files/rhomolo_technical_description.pdf

Metis, FGB, Applica and Ockham (ongoing), study supporting the evaluation of youth employment, European Commission, DG EMPL

Mohl, P. and Hagen, T. (2011), 'Do structural funds promote regional employment? Evidence from dynamic panel data models', *Working Paper Series* No. 1403, European Central Bank, December 2011. Available at: <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1403.pdf>

Montuenga-Gómez, V.M. and Ramos-Parreño, J.M., (2005). Reconciling the wage curve and the Phillips curve. *Journal of Economic Surveys*, Vol. 19(5). Available at: <https://onlinelibrary.wiley.com/doi/10.1111/j.0950-0804.2005.00266.x>

Nijkamp P., and Poot, J., (2005). The Last Word on the Wage Curve? *Journal of Economic Surveys*, vol. 19(3). Available at: <https://onlinelibrary.wiley.com/doi/10.1111/j.0950-0804.2005.00254.x>

Nymoer, R., and Rødseth, A., (2003). Explaining unemployment: some lessons from Nordic wage formation. *Labour Economics*, Vol. 10(1), p. 1-29. Available at: [https://www.sciencedirect.com/science/article/pii/S0927-5371\(02\)00133-1](https://www.sciencedirect.com/science/article/pii/S0927-5371(02)00133-1)

Pellegrini, G., Terribile, F., Tarola, O., Muccigrosso, T., & Busillo, F. (2013). Measuring the effects of European regional policy on economic growth: A regression discontinuity approach. *Papers in Regional Science* 92, 217–233. Available at: <https://rsaiconnect.onlinelibrary.wiley.com/doi/full/10.1111/j.1435-5957.2012.00459.x>

- Pellegrini, G. and Cerqua, A. (2016), Measuring the impact of Structural and Cohesion Funds using regression discontinuity design in EU27 in the period 1994-2011 – Work Package 14c – Tasks 2 and 3. Ex post evaluation of Cohesion Policy programmes 2007-2013, focusing on the European Regional Development Fund (ERDF) and the Cohesion Fund (CF), European Commission, April 2016. Available at: https://ec.europa.eu/regional_policy/en/information/publications/evaluations/2016/measuring-the-impact-of-structural-and-cohesion-funds-using-regression-discontinuity-design-in-eu27-in-the-period-1994-2011-work-package-14c-ex-post-evaluation-of-cohesion-policy-programmes-2007-2013-focusing-on-erdf-and-the-cohesion-fund
- Persyn, D., Torfs, W. and Kancs, d. (2014). Modelling regional labour market dynamics: Participation, employment and migration decisions in a spatial CGE model for the EU. *Investigaciones Regionales*, Vol. 29(2014). Available at: <https://ideas.repec.org/a/ris/invreg/0262.html>
- Pieńkowski, J. and Berkowitz, P. (2015) Econometric assessments of Cohesion Policy growth effects: How to make them more relevant for policy makers?, European Commission, Regional Working Paper 2015, WP 02/2015. available at: https://ec.europa.eu/regional_policy/sources/docgener/work/2015_02_econ_assess.pdf
- Ratto, M., Roeger, W., Veld, J. (2008) QUEST III: An Estimated Open-Economy DSGE Model of the Euro Area with Fiscal and Monetary Policy, European Commission, DG ECFIN. Available at: https://ec.europa.eu/economy_finance/publications/pages/publication12918_en.pdf
- Rodrik D. (2012) Why We Learn Nothing from Regressing Economic Growth on Policies, *Seoul Journal of Economics* 2012, Vol. 25, No. 2. Available at: <https://scholar.harvard.edu/files/dani-rodrik/files/why-we-learn-regressing-nothing-by-regressinggrowthonpolicies.pdf>
- Rutherford, T. F., (2000) Getting Started with CGE Modeling. Available at: <http://web.mit.edu/paltsev/www/docs/intro.pdf>
- Sakkas, S (2018) : The macroeconomic implications of the European Social Fund: An impact assessment exercise using the RHOMOLO model, JRC Working Papers on Territorial Modelling and Analysis, No. 01/2018, European Commission, Joint Research Centre (JRC), Seville
- Sakkas, S., Conte, A., and Salotti, S. (2018). The impact of the European Social Fund: The RHOMOLO assessment. Territorial Development Insights Series, JRC113328, European Commission. Available at: <https://ec.europa.eu/jrc/sites/jrcsh/files/jrc113328.pdf>
- Thissen M., Ivanova O., Mandras G., and Husby T. (2019). European NUTS 2 regions: construction of interregional trade-linked Supply and Use tables with consistent transport flows. JRC Working Papers on Territorial Modelling and Analysis No. 01/2019, European Commission, Seville, 2019, JRC115439. Available at: <https://ec.europa.eu/jrc/sites/jrcsh/files/jrc115439.pdf>

Varga, J, and in t'Veld; J (2009) A Model-based Analysis of the Impact of Cohesion Policy Expenditure 2000-06: Simulations with the QUEST III endogenous R&D model, European Economic Policy papers n.387. Available at: https://ec.europa.eu/economy_finance/publications/pages/publication16016_en.pdf

CIEs on the evaluation of Youth Employment programmes

Angelova R., Simeonova-Ganeva R., Ganev K., Dimitrova M., Uzunova E., Taskova I., Ganeva I (2019), The Evaluation of the Measures of Youth Employment Initiative under Priority Axis 1 of Operational Program "Human Resources development" 2014-2020 - Summary of Final Report, Ministry of Labour And Social Policy Operational Program "Human Resources Development", February 2019

Baran J., Hardy W. and Kalinowsky H. (2017), Badanie efektów wsparcia zrealizowanego na rzecz osób młodych w ramach Programu Operacyjnego Wiedza Edukacja Rozwój - II Raport Wskaźnikowy, Ministerstwo Rozwoju, 2017, ISBN: 978-83-7610-621-2

Deloitte Slovenija (2018), Vrednotenje Pobude za zaposlovanje mladih (program »Prvi izziv 2015«) Znižanje brezposelnosti mladih, ki niso zaposleni in se ne izobražujejo ali usposablajo v starosti od 15 do 29 let v kohezijski regiji vzhodna Slovenija, December 2018

Desović P., Horvat M., Hodak B. Ivišić, J. (2019), Evaluacija Inicijative za zapošljavanje mladih u sklopu Operativnog programa Učinkoviti ljudski potencijali 2014. – 2020. za 2018. Godinu - Završno izvješće, Mirovinskoga Sustava, Zagreb, March 2019

Di Battista G., Landi R., Lion C., Lupo V., Pizzuti S., Santomieri K., Sciatta V., Senatore A. M., Stocco P., Toti E. (2019), Secondo Rapporto di Valutazione della Garanzia Giovani e del Programma Operativo Nazionale Iniziativa Occupazione Giovani al 31 dicembre 2018, ANPAL, January 2019

Equinox Consulting Kft. (2018), A Z Ifjúsági Foglalkoztatási Kezdeményezés Eredményeinek Értéke Lése, Innovációs és Technológiai Minisztérium (Ministry of Innovation and Technology of Hungary), December 2018

Ernst & Young Baltic (2019), youth Guarantee Targeted - robust youth Return of the labor market. Final evaluation of the "European Social Fund support and a special grant of youth employment initiatives, including youth guarantees implementation of the scheme, investment efficiency, effectiveness and impact", February 2019

Farvaque N. (2018), Note de synthèse – Lot 6 – Assistance à maîtrise d'ouvrage sur l'évaluation d'impact de l'IEJ 2018, Orseu Recherche & études

Havet N. (2019), Evaluation 2018 de l'impact de l'Initiative pour l'emploi des jeunes (IEJ) en France. Rapport final - Lot N°3: Evaluation contrefactuelle de l'impact de 3 dispositifs nationaux co-financés par l'IEJ, Ministry of Labour, March 2019

Henar, L, Chahin A., Fresno J. M., Jiménez N. and Martínez-Cantos J. L. (2018), II Evaluación De La Iniciativa De Empleo Juvenil - Informe de Evaluación, Ministerio de Trabajo, Migraciones y Seguridad Social, December 2018

Ipsos (2019), Evaluacija Inicijative za zapošljavanje mladih u sklopu Operativnog programa Učinkoviti ljudski potencijali 2014. – 2020. za 2018. Godinu, Izvješće u tijeku provedbe, Ministarstvo Rada i mirovinskoga sustava, Zagreb, January 2019

Pastore, F, and Pompili, M (2018) Servizio di valutazione unitaria dei programmi co-finanziati con Fondi dell'Unione Europea nel periodo di programmazione 2014-2020. I Rapporto tematico di valutazione – I risultati di PIPOL

Quadrant Conseil, KPMG (2018). Évaluation nationale de l'impact de l'IEJ, dimensions stratégiques et qualitatives. Report prepared at the request of DGEFP

Wolf-Watz, O. and Öhlin, J. (2018), Utvärdering av Sysselsättningsinitiativet, Strategirådet, December 2018

4 Conclusions

In its three successive chapters, this work assessed the way in which EU support to human capital development is monitored and evaluated from different perspectives.

First, it reviewed the EU standards for monitoring and evaluation of EU and benchmarked them with the literature and the guidelines of other international institutions. In doing so, it highlighted that despite a much improved result orientation there is a lack of dedicated tools for the design and assessment of M&E systems for human capital and that some ambiguity persists in differentiating the roles of monitoring and evaluation.

Second, it assessed the resulting implications of improvable standards on two major EU initiatives for the development of human capital, suggesting that there are gaps in the monitoring and evaluation practice which hamper a full understanding of what really works in the field.

Third, it focused on one specific general equilibrium modelling approach which is being used to estimate the overall aggregated effects of cohesion policy in the EU, and particularly of the European Social Fund and the Youth Employment Initiative. It suggested that some of the identified shortcomings are due to the lack of detailed micro-level information and that some work could be done to improve upon them. It also clarified that other micro and macro-econometric approaches have several limitations particularly when it comes to the evaluation of human capital development through the European Structural Funds at the macro level, also stressing the complete lack of such evaluations, to the best of the author's knowledge, with respect to the current 2014-2020 programming period.

In general, it transpires from the project that it is the lack of systematic and holistic approaches in the assessment of the M&E of human capital policies which creates difficulties. Not having a dedicated guidance for their design means that there aren't clear benchmark against which one could assess the fitness for purpose of an M&E system. Not having a clear framework on the approach to measure human capital in particular, risks reducing the comprehensiveness of the data collected and its relevance to highlight the programmes' contribution to the overarching goal of human capital development. Finally, a certain distance between the micro-level and macro-level evaluation means that neither of the two is considered fully reliable by policy makers and hence used inconsistently in informing policy making.

However, identifying problems is much easier than developing adequate solutions. It is in fact fully acknowledged that improving a M&E system comes at a high cost, and it goes beyond this research an assessment of whether this outweighs the benefits of having better evidence about the programmes funded. In fact, the guiding principle for the new 2021-2027 programming period in the EU has been, as discussed in chapter two, that of reducing administrative burden for all stakeholders in defining the new monitoring arrangements for the future of the funds. But this project suggests that there exist more than a few incremental improvements to be taken into consideration which could highly increase the cost effectiveness of the M&E systems for the development of human capital, with specific recommendations included in each chapter. After all, M&E is very costly already and anyways. So if some further efforts might strongly

increase its fitness for purpose, its “effectiveness” element might rise much faster than the related “costs”.

Lastly, one rather striking aspect is the low level of debate on such issues in the academic literature. In the view of the author, it is paramount that more attention is paid from the academia to policy evaluation “in practice”, and thus to the studies which actually lead to policy recommendations being delivered to policy makers. Despite several limitations, this research aims to make a first step into bridging the gap between the work of practitioners carrying out and discussing policy evaluation in the surroundings of the EU and the evidence stemming from the academic debate, in line with the aim of this hybrid doctoral project. But it calls for additional studies dealing with the subject matter with a view to overcoming the limitations of the current research, eventually contributing to better policy making..

5 Annex I - EU policies for human capital development – an overview

5.1 Introduction and an operative definition of human capital

The purpose of this analysis is to set the scene on EU policies for human capital development, and specifically:

- to discuss its legal basis
- to identify the overarching strategies and understand how are these monitored
- to identify the main policies initiatives and supporting financial instruments (programmes)

A number of documents have been consulted to draw a comprehensive picture on what the EU institution are doing with a view to fostering human capital across the EU-28, including:

- Europe 2020, and their related framework strategies in relevant fields (e.g. ET2020, European Agenda for Culture);
- The EC policy priorities (Junker's Agenda) and the European Pillar of social rights;
- The 2014-2020 Multiannual Financial Framework, including the aims of each programme;
- The European Social Pillar and related social scoreboard

This multi-pronged approach appears as necessary if a study is to identify first, the EU right to act in the human capital field and, second, all relevant policy interventions, initiatives and programmes that aim at developing human capital in the rather fragmented EU policy framework.

Finally, some relevant sources of data and indexes on skills and skills related are mapped, which can serve for inspiration to understand which data does the EU already collect in this field and how it approaches HC measurement-related issues.

Two key introductory remarks are necessary to understand the following paragraphs.

1. A distinction should be drawn between:
 - EU framework strategies and flagship initiatives
 - EU single market, EU external action and, especially, the EU programmes (financial lever)

EU strategies and flagship initiatives normally set the overall aims and objectives the EU pledges to pursue through a number of tools, that can be broadly categorised into the EU single market, EU external action and, especially, EU programmes. Therefore, they do not come with a specific budget attached.

The EU single market is both the cornerstone of the European project and a means to pursue through dedicated rules and regulations the aims of the EU. The EU external action is rather aimed to make EU foreign policy more coherent and effective, thus increasing Europe's global influence. Lastly, but most importantly, a number of funded programmes, whose reference budget is set through the multiannual financial framework, then substantiate the EU strategies in various fields, providing resources

aimed at achieving, in the first place, each programme's specific objectives, but, ultimately, the EU framework objectives.

In principle, monitoring goes where money is spent. This reflects one of the key purposes of monitoring, that is, to ensure that resources are spent in a way that is efficient towards the results that are anticipated.

EU framework strategies have however a broader remit than the related programmes, and although are not accompanied with specific resources, they do influence the way money is spent at the EU level.

This is the reason why if one is to take into account the monitoring needs of the EU in a given policy domain, it is interesting to look at both strategies (with their related initiatives and monitoring frameworks) as well as programmes, to understand how the EU policy works in practice and where one can argue it is necessary to monitor and evaluate its results.

2. The concept of human capital adopted for this overview follows the is as follows:

“The knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic well-being”, as per the OECD publication (2001) *The Well-being of Nations, the role of human capital*.

In discussing how human capital development is monitored and measured, reference will be made to the concepts highlighted in section 1.2. This means, in particular, that non-monetary measures of HC will be taken into account, especially those related to health, self-esteem and confidence, participation in the society etc. As per the framework developed by the WEF, also the deployment of human capital will be factored in, which means that access to work, employment and gender balance in employment, to name but a few – are also objectives which are considered relevant for the development of human capital.

What this means in practice is that the following broad typologies of policies will be observed:

- General education and research
- VET
- Lifelong learning
- Access to employment
- Culture and the creativity industry
- Health
- Poverty prevention/ Social protection
- Social inclusion/ Social Cohesion
- Gender equality and anti-discrimination

Nevertheless, for tractability purposes, the area of Research and Development will not be discussed in details (and in particular Horizon 2020, the smart specialisation strategies etc).

5.2 The scope for EU action in the field of human capital development

First and foremost, the EU mandate is set out clearly in the Lisbon Treaties,⁸² which represent the legal basis upon which the EU can take decisions that will be – to differing degrees – enforced in EU Member States.

According to such legal basis, “[r]esponsibility for employment and social policy lies primarily with national governments. EU funding supports and complements their efforts”.⁸³

In fact, according to the TEU, art. 3(3), in establishing the EU internal market it will work for the sustainable development “*aiming at full employment and social progress*” and “*combat social exclusion and discrimination, and shall promote social justice and protection, equality between women and men, solidarity between generations and protection of the rights of the child.*” In addition, the same article clarifies that the EU shall “*promote economic, social and territorial cohesion, and solidarity among Member States.*” In addition, art. 19 of the TEU reaffirms the right of EU institutions to take action to combat discrimination.

The Treaty on the Functioning of the European Union goes on to clarify objectives and procedures that the EU must follow in implementing its social policy. From the outset, article 4(2) clarifies it that the EU and MS have “*shared competence in the field social policy, for the aspects defined by this treaties*”. It later specifies which are these aspects and how this shared competence may be enacted in practice.

In particular, articles 145 to 150 detail the role and rules of EU’s shared competence in the field of employment, Articles 151 to 161 establish those for the social policy and 162-169 cover the European Social fund, the education and training policy, the health and culture policies, as well as the consumers’ protection.

As regards the field of employment, article 145 establishes that “*Member States and the Union shall, in accordance with this Title, work towards developing a coordinated strategy for employment and particularly for promoting a skilled, trained and adaptable workforce and labour markets responsive to economic change with a view to achieving the objectives defined in Article 3 of the Treaty on European Union.*” It thus introduces the concept of a coordinated strategy. This is followed on by article 146, which recalls another key provision of the TFEU, namely article 121(2), setting out that in the field of economic policy MS should see their policies as a matter of common concern. Known as multilateral surveillance, this implies that the EU, and particularly the EC, shall initiate a process that involves also the European Council and results into the Council of the EU issuing “*broad guidelines of economic policy*” to the MS. The Council, together with the EC, is also appointed to the monitoring of the compliance of MS’ economic policies with these broad guidelines. A warning can be issued through qualified majority if it is ascertained that a given policy may jeopardise the proper functioning of the EU market. Specific legislative initiatives can be taken to better detail what can happen after the

⁸² Treaty of the European Union and Treaty on the Functioning of the European Union, however after the Treaty of Lisbon also the Charter of Fundamental Rights is legally binding.

⁸³ https://europa.eu/european-union/topics/employment-social-affairs_en

warning. In the vein of this multilateral surveillance, the aforementioned article 146 and subsequent 148 foresee that recommendations may be issued by the Council (after a specific procedure involving different EU committees and institutions) that are consistent with the broad guidelines of article 121(2) and aim to better coordinate MS policies towards the employment objectives of the EU.

Within the Title X, social policy, detailed rules are defined by article 153 that clarify how the EU shared competence in the field of social policy can be translated into practice in order to achieve the objectives enshrined in article 151.⁸⁴ Article 153 lists the domains in which the EU can support and complement the activities of MS and then specifies how this may be done.⁸⁵ The first way, that applies to the entire list of domains, is once more for the EU to adopt measures that encourage coordination, mutual learning, exchange of best practices and harmonisation of laws. In all fields but that of combating social exclusion and the modernisation of social protection system, the EU may also set out minimum requirements by way of directives. It is provided, however, that these should not impose administrative or legal constraints which may work to the disadvantage of Small and Medium enterprises. Even more importantly, art. 153 sets out that in principle, the Council can act based on the ordinary legislative procedure. However, in the field of, a.o., social security and protection of workers when the contract is terminated the Council shall act unanimously. In addition, and critically, art. 153(4) specifies it that “*The provisions adopted pursuant to this Article: (i) shall not affect the right of Member States to define the fundamental principles of their social security systems and must not significantly affect the financial equilibrium thereof*”. This is a key provision which significantly constraints the room for manoeuvre at the EU level when advancing proposals on social protection (be that security or assistance) systems, given the usually severe budgetary implications.

So, overall, it is apparent how in the field of employment and social policy, the EU has a rather limited remit and the legislative tools it can deploy to the end of setting minimum requirements are often subject to the unanimity principle – which heavily hampers their use. Nevertheless, the Treaties confer to the EU suitable means to exert the role of harmoniser, which takes different forms as it will be shown further below.

Interestingly, in the field of gender equality, the TFEU art. 157 already mandates MS to apply the principle of equal pay between men and women. Finally, in addition to spurring cooperation and mutual learning in various fields of social policy, culture, education and training the TFEU

⁸⁴ “the promotion of employment, improved living and working conditions, so as to make possible their harmonisation while the improvement is being maintained, proper social protection, dialogue between management and labour, the development of human resources with a view to lasting high employment and the combating of exclusion”

⁸⁵ “(a) improvement in particular of the working environment to protect workers' health and safety; (b) working conditions; (c) social security and social protection of workers; (d) protection of workers where their employment contract is terminated; (e) the information and consultation of workers; (f) representation and collective defence of the interests of workers and employers, including co-determination, subject to paragraph 5; (g) conditions of employment for third-country nationals legally residing in Union territory; (h) the integration of persons excluded from the labour market, without prejudice to Article 166; (i) equality between men and women with regard to labour market opportunities and treatment at work; (j) the combating of social exclusion; (k) the modernisation of social protection systems without prejudice to point (c).”

also sets the basic rules for the European Social Fund and entrust the EU with the right to adopt legislation in the domains of health and consumer protection, under certain circumstances.

More importantly, when it comes to the fields of education and training, as well as that of social cohesion (articles 162-167), the mandate of the EU appears to be somewhat broader.

In particular, in according to Article 165 TFEU “*The Union shall contribute to the development of quality education by encouraging cooperation between Member States and, if necessary, by supporting and supplementing their action [...]*” pursuing a range of objectives, including mutual learning and bench learning practices, exchanges between students and learning mobility in general, comparability of qualifications etc. In order to do so, it can either act to create incentives (this is the case of the Erasmus + programme, for instance) through the ordinary legislative procedure, or act through the council to promote recommendations.

In the field of VET, according to Article 166 TFEU “*The Union shall implement a vocational training policy which shall support and supplement the action of the Member States, while fully respecting the responsibility of the Member States for the content and organisation of vocational training*” again to increase quality of and access to C-VET and I-VET systems, increasing learning mobility, facilitating industrial restructurings share good practices. The ordinary legislative procedure is foreseen for the EP and the EC to adopt measures that could support such goals and, just like in the field of educational policy, the Council can adopt recommendations on proposal from the EC. A similar approach can be seen in the field of culture (article 167 TFEU).

Lastly, with respect to the TITLE XIV, on Public Health, article 168 seems to approach the matter from a health protection standing point. Health is not described in terms of ensuring prosperity and growth, but rather there’s an emphasis on ensuring proper treatment of cross border issues and sanitary hazard. Just like in other cases, the EU can promote measures to improve general health conditions, but with specific regard to “*to combat the major cross-border health scourges, measures concerning monitoring, early warning of and combating serious cross-border threats to health, and measures which have as their direct objective the protection of public health regarding tobacco and the abuse of alcohol*” as well as adopt recommendations but should cannot act to mandate harmonisation of health systems and the like.

So overall, the EU can act in a number of fields related to human capital development, including:

- Education (primary, secondary and tertiary), including mobility of students and staff and the harmonisation of qualification, acknowledgement of formal and informal learning etc.
- Vocational and Educational Training (C-VET and I-VET), including mobility of workers and informal learning, adaptability of workers and aiding industrial restructuring
- Social cohesion, notably through the European Social Fund and the Fund for European Aid to the Most Deprived
- Employment and social protection, though mainly through the Open Method of Coordination, by sharing good practices and mutual learning and with respect to setting minimum standards and under given conditions

- Gender equality and anti-discrimination, especially with respect to the equal pay principle
- Health, although in the treaties this is discussed more from a health security perspective

5.3 Europe 2020

5.3.1 EU 2020 main targets and flagship initiatives

The most relevant single document that describes the overall EU policy strategy is arguably Europe 2020.

Adopted in 2010 by the Commission led by Mr Barroso, Europe 2020 is the overarching strategy informing EU policy choices, a reference framework for activities both at EU and national/regional levels. It sets a number of targets that are intended as a reference of where the EU wants to be by 2020. It follows the former Lisbon strategy 2000-2010.

Europe 2020 targets are as follows:

- Employment: 75% of people aged 20–64 to be in work
- Research and development (R&D): 3% of the EU's GDP to be invested in R&D
- Climate change and energy (EU 20-20-20 targets):
 - greenhouse gas emissions 20% lower than 1990 levels
 - 20% of energy coming from renewables
 - 20% increase in energy efficiency
- Education
 - rates of early school leavers below 10%
 - at least 40% of people aged 30–34 having completed higher education
- Poverty and social exclusion
 - at least 20 million fewer people in – or at risk of – poverty/social exclusion

Already from the phrasing of the targets it is clear that they are mutually reinforcing. Especially investment in human capital, that is, in this case, in education, can help boost employment and reduce poverty. Human capital is arguably also a key driver for mitigating climate change, and boost research and development. However, the latter two can surely be viewed as more strictly related to investments in physical capital.

The targets were then enacted through 7 flagship initiatives, as per the following list:

- "Innovation Union" to improve framework conditions and access to finance for research and innovation so as to ensure that innovative ideas can be turned into products and services that create growth and jobs.
- "Youth on the move" to enhance the performance of education systems and to facilitate the entry of young people to the labour market.
- "A digital agenda for Europe" to speed up the roll-out of high-speed internet and reap the benefits of a digital single market for households and firms.
- "Resource efficient Europe" to help decouple economic growth from the use of resources, support the shift towards a low carbon economy, increase the use of renewable energy sources, modernise our transport sector and promote energy efficiency.

- "An industrial policy for the globalisation era" to improve the business environment, notably for SMEs, and to support the development of a strong and sustainable industrial base able to compete globally.
- "An agenda for new skills and jobs" to modernise labour markets and empower people by developing their skills throughout the lifecycle with a view to increase labour participation and better match labour supply and demand, including through labour mobility.
- "European platform against poverty" to ensure social and territorial cohesion such that the benefits of growth and jobs are widely shared and people experiencing poverty

Again, human capital appears as rather cross-cutting. It is clearly the cornerstone of "An agenda for new skills and jobs", but its relevance is well apparent also in "youth on the move" and in the European platform against poverty.

5.3.2 Additional targets within the frame of the EU 2020 strategy

5.3.2.1 Education and Training 2020

Within the framework of Europe 2020, the Council also approved its conclusions on European cooperation in education and training, which provide the EU strategic framework for 2020 in this domain.

The priorities identified were the following:

- Making lifelong learning and mobility a reality
- Improving the quality and efficiency of education and training
- Promoting equity, social cohesion, and active citizenship
- Enhancing creativity and innovation, including entrepreneurship, at all levels of education and training

A number of benchmark targets have been set in order track progress towards the objectives of this framework strategy, and notably:

- **Participation in lifelong learning:** An average of at least 15 % of adults should participate in lifelong learning.
- **Low skilled youths:** The share of low-achieving 15-year-olds in reading, mathematics and science should be less than 15 %.
- **Tertiary level attainment:** The share of 30-34 year-olds with tertiary educational attainment should be at least 40 %.
- **Early school leaving:** The share of early leavers from education and training should be less than 10 %.
- At least 95 % of children between 4 years old and the age for starting compulsory primary education should participate in **early childhood education**.
- **Employment of graduates:** The share of employed graduates (20-34 year-olds) having left education and training 1-3 years before the reference year should be at least 82 %.

- **Graduates mobility:** An EU average of at least 20 % of higher education graduates should have had a period of higher education-related study or training (including work placements) abroad, representing a minimum of 15 ECTS credits or lasting a minimum of three months.
- **VET student mobility:** An EU average of at least 6 % of 18-34 year-olds with an initial vocational education and training (IVET) qualification should have had an IVET-related study or training period (including work placements) abroad lasting a minimum of two weeks, or less if documented by Europass.

Progress towards these targets is monitored through several initiatives of the EC and primarily through the Education and Training Monitor, which is an annual publication that captures the evolution of education and training in the EU. A signifying point is that data for the last two targets is not yet available, although it should become so by 2019.

In 2012 and 2014 the EC decided to engage in a stock-taking exercise which included evaluating the progress made towards the objectives of the framework.

This has later led to a joint report from the Commission and Member States, adopted in late 2015 by the council, which specifies six new priorities, as follows:

- relevant and high-quality skills and competences for employability, innovation, active citizenship and well-being (e.g. creativity, sense of initiative and critical thinking);
- inclusive education (i.e. including the increasing diversity of learners), equality, non-discrimination and the promotion of civic competences (e.g. mutual understanding and democratic values);
- open and innovative education and training, including fully embracing the digital era;
- strong support for educators (e.g. improved recruitment, selection and training processes as well as continuing professional development);
- transparency and recognition of skills and qualifications to facilitate learning and labour mobility (e.g. by means of the European quality reference framework);
- sustainable investment (including exploring the potential of the investment plan for Europe), performance and efficiency of education and training systems.

It is also based on this framework that the New Skills Agenda for Europe (2016) was later approved.

Ultimately, the ET 2020 is relevant for the development of human capital in its entirety, and can be seen as the real reference framework for human capital development in the EU. Yet the dimension of human development, there including also health, social capital and other factors of well-being, are arguably not captured by this framework. Such dimensions should be rather sought for in the strategic plans for health and food safety 2016-2020, by DG SANTE, but no framework targets are therein envisaged.

5.3.2.2 European Agenda for Culture

The European Agenda for Culture is a reference framework launched in 2007 by a council resolution that sets out the strategical objectives of the EU in the field of culture. These are, at broad:

- promotion of cultural diversity and intercultural dialogue;

- promotion of culture as a catalyst for creativity in the framework of the Lisbon Strategy for growth, employment, innovation and competitiveness;
- promotion of culture as a vital element in the Union's international relations;

The Agenda is then implemented through successive workplans, the latest being the 2015-2018 Work Plan for Culture, that further specify the focus of EU's efforts in the field.

- Accessible and inclusive culture
- Cultural heritage
- Cultural and creative sectors: creative economy and innovation
- Promotion of cultural diversity, culture in EU external relations, and mobility

Both are broken down into concrete actions and objectives. They provide a reference for EU spending in promoting culture, especially through the Creative Europe programme, as specified later on.

Human capital is clearly the backbone of this policy.

5.4 The current political guidelines for the EC: 'A New Start for Europe: My Agenda for Jobs, Growth Fairness and Democratic Change'

As the new Commission settled in in 2014, another extensive list of priorities has been brought about. The ten policy priorities of the EC lead by Jean Claude Juncker are as follows:

1. A new boost for jobs, growth and investment
2. A connected digital single market
3. A resilient Energy Union with a forward-looking climate change policy
4. A deeper and fairer internal market with a strengthened industrial base
5. A deeper and fairer Economic and Monetary Union (EMU)
6. A reasonable and balanced free trade agreement with the United States
7. An area of Justice and Fundamental Rights based on mutual trust
8. Towards a new policy on migration
9. Europe as a stronger global actor
10. A Union of democratic change.

Again, a number of EC priorities are tightly linked to human capital development. This is particularly evident in the first point of the EC agenda. Importantly, as part of the new boost for jobs, growth and investment priority, in June 2016 the EC launched A new skills agenda for Europe, which updates the “agenda for new skills and jobs, one of the 7 flagship initiatives promoted within the frame of Europe 2020. The New Skills Agenda for Europe is also tightly matched with the revision of the ET 2020 framework, and focuses attention on those benchmarks where Europe is lingering more clearly behind, and will be discussed here below.

5.5 A new skills agenda for Europe and the European Social Fund

5.5.1 A New Skills Agenda for Europe - introduction

Within a rather fragmented framework of EU funds and policies for human capital development, in June 2016 the European Commission launched A New Skills Agenda for Europe (NSA), which is a broad and comprehensive plan involving Member States and stakeholders aimed at **improving the quality and relevance of skills formation** spanning across all education and training levels, as well as lifelong learning. It also aims at making skills more visible and comparable within and across countries, and at improving data on skills so as to make it easier to anticipate Labour Market (LM) developments and help make better career choices. The NSA supports a shared commitment and works towards a common vision about the strategic importance of skills for sustaining jobs, growth and competitiveness. By adopting a needs-based approach the NSA aims at bridging widening skills mismatches.

The NSA is not an entirely new initiative and builds upon the experience of, e.g., the New Skills for New Jobs initiative of 2008 and the Agenda for New Skills and Jobs, by strengthening and, in some cases, streamlining existing policies to better assist Member States in their national reforms as well as to trigger a change of mindsets in both individuals and organisations.

It is centred around three key work strands or general **objectives**, broken down into 10 concrete **actions**:

A. Improving the quality and relevance of skills formation

1. **Skills guarantee**, to help low-skilled adults acquire a minimum level of literacy, numeracy and digital skills and progress towards an upper secondary qualification
2. **Review of Key competence** framework for a better understanding of qualifications and to make better use of all available skills in the European labour market
3. **Making VET a first choice** by enhancing opportunities for VET learners to undertake a work based learning experience and promoting greater visibility of good labour market outcomes of VET
4. **Digital Skills and Job Coalition**, bringing together Member States and education, employment and industry stakeholders to develop a large digital talent pool and ensure that individuals and the labour force in Europe are equipped with adequate digital skills

B. Making skills and qualifications more visible and comparable

5. **Review of the EQF** to help more people acquire the core set of skills necessary to work and live in the 21st century with a special focus on promoting entrepreneurial and innovation-oriented mind-sets and skills.

6. **Skills Profile Tool for Third Country Nationals** to support early identification and profiling of skills and qualifications of asylum seekers, refugees and other migrants.

C. Improving skills intelligence and information for better career choices

7. **Revision of the Europass Framework** offering people better and easier-to-use tools to present their skills and get useful real-time information on skills needs and trends which can help with career and learning choices.
8. **Analysis and Best Practices** on the issue of brain drain
9. **Blueprint for Sectoral Cooperation on Skills** to improve skills intelligence and address skills shortages in specific economic sectors.
10. **Initiative on graduate tracking** to improve information on how graduates progress in the labour market.

These actions are rather diverse in that they cover a broad range of target groups and tackle different challenges; furthermore, some are geared towards the support of individuals, while others operate at a rather systemic level; they also have different implementation timelines, with actions 1, 4, 5 being launched in 2016 and the remaining in 2017 and 2018.

By ensuring that European citizens have maximum access to skills (from basic, horizontal and entrepreneurial skills, up to highly qualified skills driving competitiveness and innovation), the NSA represents an important initiative in the European Semester process which every year produces country Specific Recommendations and towards reaching the Europe 2020 goals.

The NSA seems particularly suitable to spur progress towards some of the targets of the ET2020⁸⁶ framework, and particularly participation in education and training (LLL), the share of employed graduates with at least upper secondary education attainment as well as early school leaving through VET.

In order to do so, it leverages on different funding facilities provided by the EU, as well as national funds, in that the NSA comes with no budget attached.

The ESF, together with the other ESI funds (ERDF, CF, EAFRD, EMFF) and EC programmes and initiatives (AMIF, Horizon 2020 and Erasmus +), is expected to play a key role in supporting the **objectives** and **actions** of the NSA. Through the ESF, substantial investment in skills development and relevance for the labour market will be provided by Member States. The NSA Communication,⁸⁷ anticipates that at least 27 million euros (corresponding to the ESF allocation to the thematic objective 10 of the ESIF) from the ESF can be considered as highly relevant for the implementation of the NSA, and goes on to suggest that additional ESF resources aimed at promoting quality employment (TO8) and social inclusion (TO9) could

⁸⁶ https://ec.europa.eu/education/policies/european-policy-cooperation/et2020-framework_en

⁸⁷ European Commission (2016) [A New Skills Agenda for Europe: Working together to strengthen human capital, employability and competitiveness](#) {SWD(2016) 195 final}

serve the same purpose. Given the relevance of the ESF for the NSA, the next paragraph turns to describe its main purpose and structure.

5.6 The European Pillar of Social Rights and the Social Scoreboard

5.6.1 Key principles and human capital development

The European Pillar of Social Rights is arguably one of the most relevant policy initiatives promoted by the European Commission under Juncker's presidency and adopted jointly, after extensive consultation and dialogue with stakeholders, by the European Parliament and Council of the European Union in November 2017. Initially thought only for the Eurozone, the EPSR was eventually endorsed by all EU MS, including the UK. It *“consists of a set of social rights and principles and is accompanied by a package of proposals, comprising pre-existing initiatives, new legislation and soft law measures”*.⁸⁸

The EPSRs is structured around three main chapters, namely:

- Equal opportunities and access to the labour market;
- Fair working conditions; and
- Social protection and inclusion.

Its overarching purpose is to deliver “new and more effective rights for citizens”.

As the EPSRs touches upon a number of fields in which the EU competences are at best modest, as clarified in section 1.3.2, many wonder about what could its actual contribution be over coming years. It is believed that the EU own resources, and notably the ESF and other funds in the field of human capital promotion, will have a key role to play here.⁸⁹

Nevertheless, the EPSRs states a number of principles which are key to the domain of skills development. In the first area “equal-opportunities and access to the labour market” it establishes:

- Education, training and life-long learning “people's right to high quality and inclusive education, training and life-long learning so that they can maintain and acquire skills that enable them to participate fully in society and to successfully manage transitions in the labour market.” (relevant from a HC capacity, development and deployment perspective in the WEF framework)
- Gender equality, in all fields including in employment related matters (relevant from a HC deployment and know how perspective)
- Equal opportunity and non-discrimination in all fields (relevant from an HC deployment perspective, and to create a culture of tolerance which can further attract HC)

⁸⁸ https://www.cambridge.org/core/services/aop-cambridge-core/content/view/162B548D38A7D7385F4C8FF0A19C61CF/S1574019618000093a.pdf/european_pillar_of_social_rights_effectively_addressing_displacement.pdf

⁸⁹ CECHIN-CRISTA, P. et al (2013), The Social Policy of the European Union, International Journal of Business and Social Science Vol. 4 No. 10 [Special Issue – August 2013]

- Active support to employment (relevant from an HC deployment and know-how perspective)

In the area of Fair Working condition, it includes principles which are relevant from health and gender balance perspective, so again it is about HC deployment and know-how.

Finally, in the area of Social Protection and inclusion, again it is the element of adult accumulating competences and avoiding skills obsolescence (deployment and know-how) which seems more germane to the concept of human capital, but also health protection, housing and assistance for the homeless and inclusion of people with disabilities should form part of it in its extended definition.

5.6.2 The JRC's social scoreboard

The social scoreboard is a key monitoring tool to track MS progress towards the objectives defined by the EPSRs and, as a result, human capital development in Europe. It includes a range of indicators, which are divided in 12 areas. Below the list is reported (from which a range of disaggregated measures have been removed).

Table 25 – main indicators in the social scoreboard

Equal opportunities and access to the labour market	1. Education, skills and lifelong learning	Early leavers from education and training
		Adult participation in learning
		Underachievement in education
		Tertiary educational attainment, age group 30-34
	2. Gender equality in the labour market	Gender employment gap
		Gender pay gap in unadjusted form
	3. Inequality and upward mobility	Income inequality (quintile share ratio S80/S20)
		Variation in performance explained by students' socio-economic status
	4. Living conditions and poverty	At-risk-of-poverty or social exclusion rate (AROPE)
		Severe material deprivation rate (SMD)
Persons living in a household with a very low work intensity		
5. Youth	Severe housing deprivation (% of owners, with mortgage or loan)	
	Young people neither in employment nor in education and training, age group 15-24 (NEET rate)	
Dynamic labour markets and fair working conditions	6. Labour force structure	Employment rate (20-64)
		Unemployment rate (15-74)
		Activity rate
		Youth unemployment rate
	7. Labour market dynamics	Share of long-term unemployment
		Activation measures – labour market policies participants per 100 persons wanting to work (total)
		Employment in current job by duration (from 0 to 11 months)
	8. Income, including employment-related	Transition rates from temporary to permanent contracts (3-year average)
		Real unadjusted gross disposable income of households per capita: Index 2008=100
		Net earnings of a full-time single worker without children earning an average wage (levels in Purchasing Power Standards (PPS), three year average)
Public support / Social protection and inclusion	9. Impact of public policies on reducing poverty	In-work at-risk-of-poverty rate
		Impact of social transfers (other than pensions) on poverty reduction
		General government expenditure in social protection
		General government expenditure in health
	10. Early childhood care	General government expenditure in education
		Aggregate replacement ratio for pensions
	11. Healthcare	Children aged less than 3 years in formal childcare
		Self-reported unmet need for medical care (% of male population)
		Healthy life years at the age of 65 (males)
	12. Digital access	Out-of-pocket expenditure on health care
Digital skills (% of individuals with basic or above basic overall digital skills)		
		Connectivity dimension of the Digital Economy and Society Index (DESI)

Source: author's elaboration based on the JRC's social scoreboard

One can see that most of these indicators are HC relevant, but also that:

- The costs of education, training and lifelong learning activities are not considered
- The quality of education is taken into account only for 15 y.o. (PISA)
- There are some elements on disposable income and earnings
- There is nothing on soft skills, non-cognitive skills etc.

5.7 Additional policy level arrangements for the monitoring of HC initiatives

In addition to the Social Scoreboard and the ET2020, some additional indicators are worth highlighting, especially from CEDEFOP and Eurofound, on HC capacity and deployment

- Cedefop:
 - Education for hiring
 - Employment (and employment growth) in high-tech economy & occupations
 - European Skills Index
 - Future annual employment growth and needs, job openings and prospects
 - Importance of foundation, job specific and transversal skills
 - Job satisfaction
 - Learning intensity at work
 - Skills development at work, in online vacancies, obsolescence and under-utilisation
 - Training for work
 - Job-specific skills gaps, Foundation skills gaps, transversal skills gaps, under-skilling and under-skilling at hiring
 - Workplace learning
- Eurofound skills in occupations, that is, typologies of skills required in different occupations

Table 26 – skills in occupations as per Eurofound

Intellectual	Creativity and resolution
	Gather and evaluate information
	Literacy
	Numeracy
Physical	Dexterity
	Strength
Social	Manage and coordinate
	Service and attend
	Sell and influence
	Teach, train and coach
Use of methods	Autonomy
	Routine
	Teamwork
Use of technology	Use of ICT
	Use of machine

Source: Eurofound

These datasets show some good potential to track also additional elements of Human capital development, such as skills capacity, development, deployment, relevance of non-cognitive skills and future skills trends. These data, or similar approaches to data collection, could arguably be used in the assessment of some of the programmes funded by the MFF. So now attention is turned to these programmes, to see which ones should be considered relevant for human capital development.

5.8 The Multiannual Financial Framework and its programmes: how is human capital development funded in the 2014-2020 programming period

The multiannual financial framework (MFF) is a budgeting tool of the EU which sets out 'ceilings', i.e. the maximum amount of money the EU can spend, by policy priority and programme. It is a tool that serves the EU itself, as well as its stakeholders, to have a better idea of where money is being spent by the EU.

The framework has a 7-year timescale, while the actual spending is approved on a yearly basis and does not normally allocate the same level of funding envisaged by the MFF in order to leave some room for manoeuvre. The MFF has many specificities that, however, fall outside of the scope of this contribution. Amongst other things, it is worth mentioning that it has been revised in 2016, to allow more flexibility and additional funding on some specific programmes, and that it is divided into 6 categories of expense, the so called headings, as per the below:⁹⁰

- **Smart and Inclusive Growth**
 - **Competitiveness for growth and jobs:** includes research and innovation; education and training; trans-European networks in energy, transport and telecommunications; social policy; development of enterprises etc.
 - **Economic, social and territorial cohesion:** covers regional policy which aims at helping the least developed EU countries and regions to catch up with the rest, strengthening all regions' competitiveness and developing inter-regional cooperation.
- **Sustainable Growth: Natural Resources:** includes the common agricultural policy, common fisheries policy, rural development and environmental measures.
- **Security and citizenship:** includes justice and home affairs, border protection, immigration and asylum policy, public health, consumer protection, culture, youth, information and dialogue with citizens.
- **Global Europe:** covers all external action ('foreign policy') by the EU such as development assistance or humanitarian aid with the exception of the European Development Fund (EDF) which provides aid for development cooperation with African, Caribbean and Pacific countries, as well as overseas countries and territories. As it is not funded from the EU budget but from direct contributions from EU Member States, the EDF does not fall under the MFF.
- **Administration:** covers the administrative expenditure of all the European institutions, pensions and European Schools.

⁹⁰ http://ec.europa.eu/budget/mff/introduction/index_en.cfm

- **Compensations:** temporary payments designed to ensure that Croatia, who joined the EU in July 2013, does not contribute more to the EU budget than it benefits from it in the first year following its accession.

Each of these is then broken down by the actual programmes funded by the EU. It is already visible as some of these headings are less relevant (administration, compensation) when it comes to human capital development, whereas for the rest, there is a need to further investigate.

The analysis of the MFF is key within this contribution, as it is a mandatory step to appraise how the EU allocates money for human capital development, and, in turn, which are the most relevant EU programmes that should be analysed if one is to discuss how the EU is monitoring and evaluating its investments in human capital development.

What follows is the result of an overview of each programme's aims and areas of intervention/activities. The programmes have been categorised by relevance, where relevance is intended as whether a given fund supports the carrying out of activities that promote human capital development and, more specifically, its monitoring and evaluation practice its worth assessing within the context of this research project.

The criteria employed to suggest a given level of relevance of the programmes from this perspective have been:

- Whether the given programme is concerned with the promotion of human capital, and, if so, to what extent (fully, only partly, indirectly)
- Whether the promotion of human capital envisaged by the programme falls within a standard definition of human capital (skills, competences) or a broader one (health, citizenship, human rights)
- Whether the fund is all-encompassing with respect to human capital development or it covers only one or more specific aspects (e.g. only skills/only health/only human rights etc.) or only in a given sector (agriculture, maritime, energy etc.)

The table below, provides an overview of those programmes that have been considered having medium to full relevance, accompanied with an explanation and the allocation according to the MFF as a share of the total EU yearly budget (2014). In the table below, not all EU programmes have been reported, but only those whose relevance was considered at least medium, which implies that some form of human capital development activity is promoted by the fund.

Funds such as the ERDF, the large infrastructure programmes (ENGOS, Galileo, Copernicus, Iter) or the Connecting Europe Facility – to name but a few – are therefore not reported in the table. Another key programme which is entirely unrelated to human capital development, and is the single largest programme of the EU budget, is the “European Agricultural Guarantee Fund (EAGF) — Market related expenditure and direct payments”, which represents the cornerstone of the EU Common Agricultural Policy. A full overview of this programmes is however provided in Annex I.

Table 27 – overview of programmes in the MFF 2014-2020 and their relevance when it comes to human capital development

Programme	Aim	relevance	Explanation	Est. Allocation (% of tot MFF)
1a: Smart and inclusive growth: Competitiveness for growth and jobs				
CFS Research and innovation: Horizon 2020	<i>Horizon 2020 is the financial instrument implementing the Innovation Union, a Europe 2020 flagship initiative aimed at securing Europe's global competitiveness. Seen as a means to drive economic growth and create jobs, Horizon 2020 has the political backing of Europe's leaders and the Members of the European Parliament. They agreed that research is an investment in our future and so put it at the heart of the EU's blueprint for smart, sustainable and inclusive growth and jobs .</i>		Flagship programme for research and innovation. Finances both physical and human capital. It fosters higher skills and scientific excellence to support the EU 2020 strategy.	6.17%
CFS Research and innovation: Euratom Research and Training Programme	<i>Euratom aims to pursue nuclear research and training activities with an emphasis on continually improving nuclear safety, security and radiation protection, notably to contribute to the long-term decarbonisation of the energy system in a safe, efficient and secure way. By contributing to these objectives, the Euratom Programme will reinforce outcomes under the three priorities of Horizon 2020: Excellent science, Industrial leadership and Societal challenges.</i>		Complementary to Horizon 2020 and ITER, this programme finances research and training in the field. Skills and competences are a key priority of this fund	0.20%
Competitiveness of enterprises and SMEs (COSME)	<i>COSME will support:</i> - better access to finance for SMEs - access to markets for SMEs - entrepreneurship - more favourable conditions for business creation and growth		Most of the funding is spent on priorities which are not directly linked with human capital development. However, under the entrepreneurship objective (3% of total funding) COSME supports entrepreneurs by strengthening entrepreneurship education, mentoring, guidance and other support services. Actions support specific groups who may find it difficult to reach their full potential, such as young people, women and senior entrepreneurs. The programme also aims to help businesses access opportunities offered by digital technologies.	0.19%
Education, Training, Youth and Sport (Erasmus+)	<i>Erasmus+ is the EU Programme in the fields of education, training, youth and sport for the period 2014-2020. Education, training, youth and sport can make a major contribution to help tackle socio-economic changes, the key challenges that Europe will be facing until the end of the decade and to support the implementation of the European policy agenda for growth, jobs, equity and social inclusion.</i>		erasmus + invest directly in skills, competences and knowledge (in a broad sense) that should be provided to individuals to improve social inclusion, equity, jobs, growth etc.	1.06%
Social change and Innovation (PSCI/EaSI)	<i>The Programme supports:</i> - the Union's policies and laws, promoting evidence-based policy-making, raising employment participation, enabling social policy innovation and social progress, in partnership with the social partners, civil society organisations and public and private bodies; - voluntary geographical and occupational mobility for workers on a fair basis and contributing to a high level of quality and sustainable employment information via information exchanges and dissemination and other forms of cooperation, such as cross-border partnerships; - access to, and the availability of, financing for persons in vulnerable situation and supporting the development of the social investment market and facilitating access to finance for social enterprises.		despite this fund being fully concerned with activities in the field of human capital development, services are mostly not provided at the individualised level. EaSI in fact funds mainly analysis, mutual learning and dissemination, operating cost of NGOs or other actors in the field, other than microfinancing and labour mobility. Monitoring needs therefore differ from those of, e.g., the ESF	0.08%

Programme	Aim	Relevance	Explanation	Est. Allocation (% of tot MFF)
1b: (Smart and inclusive growth) Economic, social and territorial cohesion				
Youth Employment initiative (specific top-up allocation)	<i>The Youth Employment Initiative is one of the main EU financial resources to support the implementation of Youth Guarantee schemes. It was launched to provide support to young people living in the regions where youth unemployment was higher than 25% in 2012. It exclusively supports young people who are not in education, employment or training (NEETs), including long-term unemployed youngsters or those not registered as job-seekers.</i>		This is a key instrument to support the youth guarantee, which is fully devoted at providing upskilling through trainings and work experience.	1.23%
European Social Fund (ESF)	<i>the ESF covers TO 8, 9, 10 and 11, namely access to employment, social inclusion, education and training and capacity building</i>		It is the main EU instrument for the promotion of human capital. Its relevance is highest as it covers human capital under many relevant dimensions (thus including social inclusion, non discrimination, citizenship etc.) generally through the provision of better skills and competences (both basic skills and more advanced ones)	8.11%
European territorial cooperation	<i>made of three further branches, namely cross-border, transnational and interregional cooperation, it is aimed at improving cooperation between EU regions</i>		Funded activities are comprised of a mix of infrastructure and human capital investments, the latter being particularly relevant in the context of interregional cooperation, where the sharing of best practices and mutual learning is key. Yet activities normally address systems and organisations rather than individuals.	0.35%
European Aid to the Most Deprived (FEAD)	<i>Supports EU countries' actions to provide material assistance to the most deprived. This includes food, clothing and other essential items for personal use, e.g. shoes, soap and shampoo. Material assistance needs to go hand in hand with social inclusion measures, such as guidance and support to help people out of poverty. National authorities may also support non-material assistance to the most deprived people, to help them integrate better into society.</i>		FEAD is an enabling fund aimed at providing immediate relief to the most deprived. With the ultimate goal of contributing to lifting 20 million people out of poverty, it provides a range of services, from food support and basic material assistance (type I OPs) to	0.34%

Programme	Aim	Relevance	Explanation	Est. Allocation (% of tot MFF)
2: Sustainable Growth: Natural Resources				
European Agricultural Fund for Rural Development (EAFRD)	<i>The EU's rural development policy helps the rural areas of the EU to meet the wide range of economic, environmental and social challenges of the 21st century. Frequently called "the second pillar" of the Common Agricultural Policy (CAP), it complements the system of direct payments to farmers and measures to manage agricultural markets (the so-called "first pillar"). Rural Development policy shares a number of objectives with other European Structural and Investment Funds (ESIF).</i>		<p>The EAFRD funds a mix of infrastructural and human capital interventions. Among its priorities (6 overall) are, notably</p> <ul style="list-style-type: none"> - fostering knowledge transfer and innovation in agriculture, forestry and rural areas - enhancing the viability and competitiveness of all types of agriculture, and promoting innovative farm technologies and sustainable forest management - promoting food chain organisation, animal welfare and risk management in agriculture - promoting social inclusion, poverty reduction and economic development in rural areas <p>Particularly fostering knowledge transfer and promoting social inclusion are key in the human capital development perspective</p>	9.55%
European Maritime and Fisheries Fund (EMFF)	<i>The EMFF shall contribute to the achievement of the following objectives: (a) promoting competitive, environmentally sustainable, economically viable and socially responsible fisheries and aquaculture; (b) fostering the implementation of the CFP; (c) promoting a balanced and inclusive territorial development of fisheries and aquaculture areas; (d) fostering the development and implementation of the Union's IMP in a manner complementary to cohesion policy and to the CFP. The pursuit of those objectives shall not result in an increase in fishing capacity</i>		<p>also in this case the promotion of human capital forms part of a wide range of interventions pursuing the objectives of the EMFF</p>	0.69%
Environment and climate action (Life+)	<i>The LIFE programme is the EU's funding instrument for the environment and climate action. The general objective of LIFE is to contribute to the implementation, updating and development of EU environmental and climate policy and legislation by co-financing projects with European added value.</i>		<p>typical LIFE + actions include best-practice, demonstration, pilot, information, awareness-raising or dissemination projects. Their is clearly a component related to supporting knowledge creation and sharing, thus developing human capital, although pretty project based (delivering concrete project results rather than general upskilling).</p>	0.28%

Programme	Aim	Relevance	Explanation	Est. Allocation (% of tot MFF)
3: Security and citizenship				
Asylum and Migration Fund (AMF)	The Asylum, Migration and Integration Fund (AMIF) was set up for the period 2014-20, with a total of EUR 3.137 billion for the seven years. It will promote the efficient management of migration flows and the implementation, strengthening and development of a common Union approach to asylum and immigration.		Concrete actions to be funded through the AMF can include a wide range of initiatives, such as the improvement of accommodation and reception services for asylum seekers, information measures and campaigns in non-EU countries on legal migration channels, education and language training for non-EU nationals, assistance to vulnerable persons belonging to the target groups of AMIF, information exchange and cooperation between EU States and training for staff on relevant topics of AMIF. thus, the AMF directly funds upskilling initiatives as well as the sharing of best practices, international cooperation etc.	0.28%
Justice	This programme shall contribute to the further development of a European area of justice based on mutual recognition and mutual trust. It promotes: judicial cooperation in civil matters, including civil and commercial matters, insolvencies, family matters and successions, etc. judicial cooperation in criminal matters judicial training, including language training on legal terminology, with a view to fostering a common legal and judicial culture effective access to justice in Europe, including rights of victims of crime and procedural rights in criminal proceedings initiatives in the field of drugs policy (judicial cooperation and crime prevention aspects)		Activities supported by this programme are quite clearly related to human capital development in as much as they concern trainings, mutual learning and awareness raising amongst other things. Part of the funding is absorbed through support provided to organisation.	0.03%
Rights and Citizenship	This programme shall contribute to the further development of an area where equality and the rights of persons, as enshrined in the Treaty, the Charter and international human rights conventions, are promoted and protected. Its nine specific objectives are to: Promote non-discrimination; Combat racism, xenophobia, homophobia and other forms of intolerance; Promote rights of persons with disabilities; Promote equality between women and men and gender mainstreaming; Prevent violence against children, young people, women and other groups at risk (Daphne); Promote the rights of the child; Ensure the highest level of data protection; Promote the rights deriving from Union citizenship; Enforce consumer rights		Activities supported by this programme are quite clearly related to human capital development in as much as they concern trainings, mutual learning and awareness raising amongst other things. Part of the funding is absorbed through support provided to organisation.	0.04%

Programme	Aim	Relevance	Explanation	Est. Allocation (% of tot MFF)
3: Security and citizenship				
Europe for Citizens	<p>The aim of this programme is:</p> <ul style="list-style-type: none"> - To contribute to citizens' understanding of the EU, its history and diversity - To foster European citizenship and to improve conditions for civic and democratic participation at EU level - To raise awareness of remembrance, common history and values - To encourage democratic participation of citizens at EU level, by developing citizens' understanding of the EU policy making-process and, by promoting opportunities for societal and intercultural engagement and volunteering at EU level 		a modern definition of human capital necessarily encompasses also citizenship, active participation to the social life and the like. This is why activities funded under the "Europe for citizens" programme, are fully relevant when speaking about human capital development in Europe	0.02%
Health for Growth	<p>it seeks to strengthen and emphasise the link between a healthy population as a key driver for economic growth.</p> <p>The general objective of the Health for Growth Programme is to work with the Member States to encourage innovation in healthcare, increase the sustainability of health systems, to improve the health of the EU citizens and to protect them from cross-border health threats. This programme focuses on fewer but more specific actions of proven EU added value that will deliver concrete results and respond to identified needs or gaps. The programme seeks to improve the way Member States cooperate in the area of health and to provide leverage for the reform of national health policies.</p>		health being considered as a key driver for growth and human development means that this fund falls within the scope of any exercises concerned with human capital development	0.04%
Creative Europe	<p>to help the cultural and creative sectors seize the opportunities of the digital age and globalisation;</p> <p>Enable the sectors to reach their economic potential, contributing to sustainable growth, jobs, and social cohesion;</p> <p>Give Europe's culture and media sectors access to new international opportunities, markets, and audiences.</p>		<p>THIS WILL, SPECIFICALLY, INVOLVE PROVIDING FUNDING FOR: 2,500 artists and cultural professionals, 2,000 cinemas, 800 films, 4,500 book translations</p> <p>A financial guarantee facility of up to €750 million for small businesses active in the sector will also be established in 2016.</p> <p>this fund is thus relevant for human capital development</p>	0.12%

Programme	Aim	Relevance	Explanation	Est. Allocation (% of tot MFF)
4: Global Europe				
Instrument for Pre-accession assistance (IPA)	The Instrument for Pre-accession Assistance (IPA) is the means by which the EU supports reforms in the 'enlargement countries' with financial and technical help. The IPA funds build up the capacities of the countries throughout the accession process, resulting in progressive, positive developments in the region.		Under the IPA a wide range of interventions are financed. Among these are some which directly target human capital development, which is also one of the priorities of the Instrument.	1.08%
European Neighbourhood Instrument (ENI)	Since it was launched, in 2004, the European Neighbourhood Policy (ENP) has been strengthening relations, bringing tangible benefits to both the EU and its Neighbourhood partners, including the introduction of regional initiatives and support to democratisation.		Again this is a rather composite fund, which comprises initiatives that seek to improve public services and promote mutual sharing, as well as boost investment in infrastructures in human capital. Thus human capital, especially through young and women employability, is well present within the framework of ENI, but the instrument is not relevant in its entirety	1.50%
Development Cooperation Instrument (DCI)	Its prime objective is the reduction of poverty. It contributes also to the achievement of other goals of EU external action, in particular fostering sustainable economic, social and environmental development as well as promoting democracy, the rule of law, good governance and respect for human rights.		Although the theme covered by this instrument are rather germane to the concept of human capital development (human rights, democracy, sustainable development) and human development is one of the objective of this instrument, upskilling seems to be playing a minor role.	1.60%
European Instrument for Democracy and Human Rights (EIDHR)	The EIDHR is designed to help civil society to become an effective force for political reform and defence of human rights. Building on its key strength, which is the ability to operate without the need for host government consent, the EIDHR is able to focus on sensitive political issues and innovative approaches and to cooperate directly with local civil society organisations which need to preserve independence from public authorities, providing for great flexibility and increased capacity to respond to changing circumstances.		This instrument finances one issue, notably the respect of human rights, which falls well in the scope of human capital development. As many other cooperative instruments which belong to external services, it is aimed to support organisations which will foster human rights protection outside the EU. Thus, under its mandate, only specific support to NGOs and other relevant actors is envisaged, rather than 360 degrees promotion of human capital	0.13%
Humanitarian aid	the EU provides needs-based humanitarian assistance to the people hit by man-made and natural disasters with particular attention to the most vulnerable victims. Aid is channelled impartially to the affected populations, regardless of their race, ethnic group, religion, gender, age, nationality or political affiliation.		this instrument is an emergency one, thus the aim is to provide immediate relief to those in need. This is relevant to human capital development however: 1. it is rather an enabler than human capital in itself 2. it is provided outside of the EU, as per other external policy funds/instruments	0.63%
European Voluntary Humanitarian Aid Corps EU Aid Volunteers (EUAV)	EU Aid Volunteers brings together volunteers and organisations from different countries, providing practical support to humanitarian aid projects and contributing to strengthening the local capacity and resilience of disaster-affected communities.		this programme contributes to the development of human capital in the form of providing training and capacity building activities, amongst other things.	0.01%

Source: author's elaboration based on Multiannual Financial Framework data and info available on the European Commission's website on the aims of each programme

As it can be seen, a fairly significant number of programmes is concerned with human capital development in and outside of the EU.

Overall, around 9% of the EU budget has been considered fully relevant for monitoring purpose on the subject of human capital development. This is comprised of the ESF, Erasmus + and the EaSI. Other highly relevant programmes appear to be the YEI, Horizon 2020, the FEAD but also programmes such as Europe for Citizens, Health for Growth, as well as funding for the European external policy such as the Instrument for pre-accession assistance (IPA).

Overall, funding considered at least medium relevant accounts for 35% of the total EU budget.

With a view to the criteria employed above, there are significant differences among them. Some, such as the ESF or Erasmus +, cover a wide range of upskilling and social inclusion interventions which are cross-sector and are offered to large shares of the populations. Others are thematic, or concerned with very specific objectives.

While the former will be of highest interest for this research project, in order to have a broad overview of how is human capital monitored and evaluated in the EU, there will be merit also in analysis also the monitoring practices of more targeted programmes that are concerned with human capital development in the EU.

6 Annex II - Monitoring of the EU human capital development policy: A case study on the New Skills Agenda for Europe and the ESF's monitoring system

6.1 Introduction

This section aims to discuss the recent initiative from the European Commission 'A New Skills Agenda for Europe', with a view to understanding how the progress towards its objectives is and will be monitored, as well as how and whether the monitoring system of the European Social Fund (ESF) 2014-2020, the main European tool to support the development of human capital, can help so do.

It moves from an overview of the recent New Skills Agenda for Europe (hereafter also NSA) and the way in which the ESF 2014-2020 can contribute to it, whilst trying to identify NSA's actions that may be particularly worth monitoring by means of the ESF monitoring system. It ultimately seeks to assess whether the ESF monitoring system may somehow be suitable to track progress towards (some of) the objectives of the NSA. In doing this, use is made of QSR NVivo, a software for the so called 'content analysis', which helps appraise to what extent the way Specific Indicators are phrased in ESF is suitable for monitoring purposes under the NSA.

6.1.1 Where the NSA and the ESF aims align

According to the NSA Communication and its accompanying documents,⁹¹ ESF funding that may be most relevant for the ESF is that allocated under **Thematic Objective 10** "Investing in education, training and vocational training for skills and lifelong learning" and its ESF-relevant Investment Priorities **Additional ESF-relevant investment priorities** such as those related to access to employment, integration of young people in the labour market and adaptability of workers are also bound to support education and training (e.g. for young early school leavers or adults without the necessary qualification levels allowing for direct labour market integration support) and dual learning, notably through apprenticeships. These are considered of direct importance to the NSA, in particular those linked to improving basic skills and upgrade existing skills of employed (so that they remain and progress on the labour market) and unemployed alike (in order to promote their access to the LM).

Concerning the typologies of actions supported by the ESF it has to be expected that some actions supported by the NSA will have already been embedded into ESF programming while other, more innovative, might be integrated in the course of future implementation.

In order to see in greater detail what can be the synergies between the ESF and the NSA, the next paragraph will discuss separately each NSA action and discuss the potential for ESF support as well as specific monitoring needs.

⁹¹ Ibid. and European Commission (2016) Analytical underpinning for a New Skills Agenda for Europe. SWD(2016) 195 final

6.1.2 Structure of the NSA and relevant actions⁹²

As the NSA is the focus of this paper, what follows is a summary of the New Skills Agenda for Europe with a view to identifying types of activities/interventions⁹³ envisaged by the NSA that may also fall under the scope/funding of the ESF. The reason for this is that the ESF is considered its main supporting instrument.

As anticipated, the three work strands/general objectives of the NSA are:

1. Improving the Quality and Relevance of Skills Formation
2. Making Skills and Qualifications more Visible and Comparable
3. Improving skills intelligence and information for better career choices

In what follows, a brief summary is provided of the actions of which the three strands are composed, with a view to identifying scope for ESF support.

6.1.2.1 Work strand 1: Improving the quality and relevance of skills formation

6.1.2.1.1 EC proposal on the Recommendation on the establishment of a Skills Guarantee

This action, later recast in the “Upskilling Pathways”, chiefly foresaw the building blocks presented in the table below.

Table 28 – Activities/Interventions promoted under the Skills Guarantee (level 1)

Activities/Interventions encouraged	Target groups	Likely funded under the ESF
Skills assessments/skills audits ⁹⁴		Yes
Modular/ad hoc trainings		Yes
Validation and recognition of skills acquired in accordance to existing frameworks	Low skilled adults, both in-work and out of work, with below EQF 4 level of education, who are not eligible for the Youth Guarantee.	Yes
Designation of a body, or small number of bodies, responsible for implementing the Skills Guarantee		No
Outreach measures and strategies to sensitise people to the benefits of long-term upskilling, including guidance on existing opportunities and incentives.		Yes
Support for staff in the adult learning field	Staff in the adult learning field	Yes

⁹² Based on FGB (2018b), which was written by the author and the FGB team

⁹³ Please note that within this document, the terms activity and intervention are used interchangeably. These refer to the forms of support that are envisaged in the ESF and/or encouraged by the NSA, to be read in contrast to the term action, which refers to the 10 NSA initiatives set out by the NSA Communication. In particular, activities/interventions can be seen as a sub-set of actions of each NSA action, i.e. each action encourages a set of activities/interventions.

⁹⁴ As proposed in the 2012 Council recommendation on validation of non-formal and informal learning; this usually results in a statement of skills of the individual which can be the basis for a plan for the next steps in training and the support offered

Activities/Interventions encouraged	Target groups	Likely funded under the ESF
Sharing of best practices through existing Electronic Platform for Adult Learning in Europe	Those in charge of the design, implementation and evaluation of policies in the field of adult learning	Yes

Source: author's elaboration on the basis of the text of the NSA recommendation and included also in FGB (2018b)

Many of the actions of which the Upskilling Pathways is made of are very well aligned to what supported by the ESF. In addition, these are individualised actions, hence the need for dedicated monitoring arrangements seems high.

6.1.2.1.2 EC proposal for a review of the Key Competences for Lifelong Learning

That's an EU-level action which should lead to a new recommendation on lifelong learning. No need for individualised monitoring.

6.1.2.1.3 EC support to the implementation of the Riga conclusions on VET

This action, subsumed under the heading 'Making VET a first choice', emphasises the need to support the 'Riga conclusions for quality and labour market relevant vocational skills and qualifications'⁹⁵ through specific actions, as per the table below.

Table 29 - Activities/Interventions promoted under the EC support to the implementation of the Riga conclusions (level 1)

Activities/Interventions encouraged	Target groups	Likely funded under the ESF
Work-based learning experiences for learners (internships and especially apprenticeships, etc.)	VET learners	Yes
Provision of opportunities for VET learners to combine experiences acquired in different settings	VET learners	Yes
Promotion of partnerships between learning providers, research and business, with a particular focus on needs for higher level skills at sectoral level	Learning Providers, Research and business, VET Learners	Yes
Improving data availability on labour market outcomes of VET	VET Learners	Yes
To streamline the existing EU level governance of the VET sector, including a more explicit coordination role for the Advisory Committee on Vocational Training.	VET sector governing bodies, VET learners	No
Launching a first European VET Skills Week in 2016 and increasing co-operation with World Skills Organisation to showcase vocational studies as a first class option	VET Learners and providers	No

⁹⁵ RIGA CONCLUSIONS 2015 ON A NEW SET OF MEDIUM-TERM DELIVERABLES IN THE FIELD OF VET FOR THE PERIOD 2015-2020, AS A RESULT OF THE REVIEW OF SHORT-TERM DELIVERABLES DEFINED IN THE 2010 BRUGES COMMUNIQUÉ, available at https://www.izm.gov.lv/images/RigaConclusions_2015.pdf

Activities/Interventions encouraged	Target groups	Likely funded under the ESF
Provision of efficient and integrated guidance services to enhance access to VET	VET Learners and the general population	Yes
Making available validation of non-formal and informal learning	VET Learners	Yes
Further strengthen key competences in VET curricula or develop these through I-VET and C-VET	VET Learners	Yes
Introduction of systematic approaches and more opportunities for initial and continuous professional development of VET teachers, trainers and mentors in both school and work based settings.	VET Teachers, trainers and mentors in both school and work based settings	Yes
Further development of quality assurance mechanisms in VET in line with EQAVET recommendation	VET learners and providers	Yes
Establishment of continuous information and feedback loops in I-VET and C-VET systems based on learning outcomes, as part of quality assurance	VET learners and providers	Yes

Source: author's elaboration on the basis of the text of the NSA recommendation and included also in FGB (2018b)

Many of these are in line with the individualised as well as systemic support provided under the ESF/YEI.

6.1.2.1.4 Digital Skills and Jobs Coalition

This is mainly about social partners pledging to bridge the digital skills gap. It does not directly concern ESF support.

6.1.2.2 Work strand 2: Making Skills and Qualifications more Visible and Comparable

6.1.2.2.1 EC proposal for the revision of European Qualification Framework

Also in this case, the focus is the revision of a recommendation. No direct involvement of the ESF nor need for individualised monitoring.

6.1.2.2.2 Launch of a 'Skills Profile Tool for Third Country Nationals'

The below is essentially a tool which should be used in employment services to speed up the recognition and validation of skills of third country nationals. The main building blocks of this action are listed below.

Table 30 - Activities/Interventions promoted under the 'Skills Profile Tool for Third Country Nationals'

NSA Action	Activities/Interventions encouraged	Target groups	Likely funded under the ESF
Launch of a 'Skills Profile Tool for Third Country Nationals'	Support the training of staff in reception facilities to speed up recognition procedures	Training staff in reception facilities	Yes
	Promote the sharing of information and best practices on understanding and recognition of skills and qualifications	Relevant national/regional	Yes

NSA Action	Activities/Interventions encouraged	Target groups	Likely funded under the ESF
		authorities, reception facilities, NGOs	
	Make available online language learning for newly arrived migrants, including refugees	Third nationals	country No ⁹⁶
	Launch of a tool to assist services in receiving and host countries to identify and document skills, qualifications and experience of newly-arrived third country nationals.	National receiving third country nationals	services third- No ⁹⁷

Source: author's elaboration on the basis of the text of the NSA recommendation and included also in FGB (2018b)

The nature of this action is both systemic but also individualised in the sense that the roll out of the skills recognition tool will then involve end-beneficiaries. However, it should be mainly funded through Erasmus +.

6.1.2.3 Work strand 3: Improving skills intelligence and information for better career choices

6.1.2.3.1 EC proposal for the revision of the Europass Framework for the provision of better services for skills and qualifications

Again, this mainly concerns the revision of a EU-level tool. In view of the above, these activities are not likely to be funded under ESF.

6.1.2.3.2 Best practices on brain drain

This action was never fully initiated but would mainly concern the sharing of best practices on how to deal with brain drain.

6.1.2.3.3 Launch of Blueprint for sectoral cooperation on skills

This is about developing partnerships which include the academia, social partners, policy makers in given industry sectors to improve the way skill is formed and reduce skills mismatch. Mainly funded by the Erasmus+ and COSME funds, once the strategies will be developed their roll out might be supported by the ESF (e.g. through training courses which embed the programmes or training modules devised in the context of the blueprints). No need for individualised monitoring though at this stage.

6.1.2.3.4 Initiative on graduate tracking

The aim was to improve the tracking of graduates, to have a better feedback loop towards VET as well as universities and inform both pupils, policy makers and educational institutions on

⁹⁶ Through Erasmus + online linguistic support (100.000 licences for online language courses will be made available to refugees over three years)

⁹⁷ The creation, hosting and maintenance of the tool will be financed through Erasmus+

what skill is needed in the labour market. Again, this is a rather systemic action which does not need to be monitored through individualised arrangements.

6.2 Analysis of ESF Indicators and their suitability to monitor the NSA.

6.2.1 Approach to the analysis and keywords selected

6.2.1.1 Rationale and scope of the analysis

As clarified in the previous sections, the NSA comes with no monitoring system attached. It is in fact a document that collates initiatives which places stronger emphasis on existing policies, rather than bringing about substantive change or innovation. In some cases, NSA actions also encourage MS to develop new actions or amend existing ones, including by identifying specific target groups, e.g. low-skilled adults, that should be better addressed.

In any event, and this is key from an accountability perspective, it has no dedicated budget.

Therefore, in principle, some may argue that there is no need to see how money is being spent, because no money is spent with the sole purpose of pursuing the objectives of the NSA; all that it is done is done with money belonging to other funds and facilities, under their given framework of monitoring and evaluation.

Yet, if we assume that the NSA is to steer such funding towards some specific initiatives that, at the same time, fall within the broad array of actions financed by, e.g., the ESIF, but also are devised to support the aims of the NSA, then it is sensible to contend that there is merit in gauging whether such aims are being addressed, and, if so, to what extent.

In this vein, this analysis argues that the best way to measure progress towards the objectives of the NSA appears to be making use of the existing monitoring system in place for its principal supporting fund, i.e. the ESF, given the broad scope of such support.

Thus, this analysis goes on to check whether existing ESF Common and Specific indicators are fit for such purpose.

However, the broad remit of the NSA means that the analysis of the ESF monitoring of relevant initiatives and action will have to be limited to a sub-set of actions /activities promoted by the NSA. As described above, a rather intuitive reason for this is that systemic actions such as the revision of the EQF or EUROPASS framework do not need to be tracked at the individualised service level, and their progress will be otherwise observed and analysed.

In brief, and according to the analysis in section 5.1.2, interventions in higher need of monitoring boil down to two main NSA actions:

1. The Upskilling Pathways, and thus the provision of upskilling to low skilled adults (below ISECD for and over 25 years, not eligible for the Youth Guarantee) which can be further detailed as:
 - a. Skills assessments/Skills audits
 - b. Modular ad-hoc training
 - c. Validation and recognition of skills acquired

- d. Outreach measures on the benefits of upskilling and guidance
 - e. Support to staff in the adult learning field
2. 'Making VET a first choice', and thus the interventions aimed at enhancing the quality of VET systems, which can be further detailed in:
- a. internships and especially apprenticeships for VET learners, also in different settings
 - b. Promotion of partnerships between learning providers, research and business
 - c. Improving data availability on labour market outcomes of VET
 - d. Provision of efficient and integrated guidance services to enhance access to VET
 - e. Making available validation of non-formal and informal learning
 - f. Further strengthen key competences in VET curricula or develop these through I-VET and C-VET
 - g. Training of trainers/mentors/ VET teachers
 - h. Further development of quality assurance mechanisms in VET in line with EQAVET recommendation and feedback loops in I-VET and C-VET systems

For the analysis of common indicators, given their very limited number, no specific tool nor software has been employed. A simple reading of the indicators suffices to cherry-pick those relevant. However, their formulation immediately strikes as too broad to provide conclusive information on whether ESF actions are effectively spurring progress towards the objectives of the NSA. It is so as common indicators, be them output, immediate or longer-term results, take into account only a single socio-economic feature of a participant (age, skills level, social condition etc.) and do not cover for the type of support granted.

Because common output indicators are the single element within the monitoring system that is meant for aggregation, this analysis has nevertheless started from them.

Conversely, 'programme-specific indicators can be very specific in order to highlight certain aspects of the ESF support which are of particular importance for the Member States/regions. A programme-specific indicator consisting of more than one dimension also allows the managing authority and the monitoring committee to monitor more specifically than the common indicators selected specific aspects of programme implementation'.⁹⁸

Thus, the analysis moved on to Programme-Specific Output Indicators, much larger in number (3256) and far more varied and detailed in their formulation. In assessing the results, one should in any event bear in mind that, even where specific indicators are found that can measure progress towards the achievement of the NSA goals, aggregability issues may affect the estimates.

⁹⁸ http://www.gren.pt/np4/np4/?newsId=1334&fileName=Draft_ESF_guidance_monitoring_evaluation.pdf

Another important caveat should be stressed: this analysis focused on output indicators in that results indicators (both specific and common) should be logically linked to the former – though somewhat less specific in their formulation. A one-to-one relationship is in most cases expected, and in any event the array of output indicators is comprehensive of the smaller and more focused one of results. Hence the choice of analysing output indicators for proxying how NSA-relevant themes are covered by the ESF monitoring system.

6.2.1.2 Steps, tools and methods for the analysis

Summing up, this analysis is divided into two distinct steps:

1. Analysis of ESF common output indicators
2. Analysis of ESF specific output indicators

6.2.1.2.1 First step, analysis of common output indicators

For the first part, no specific tool has been selected for the analysis. The analytical approach for this section was as follows:

1. Look at the existing list of common output and result indicators
2. Filter out on a qualitative basis those that observe progress on dimensions that are unrelated to the NSA
3. Determine how those that are deemed relevant are distributed across thematic objectives and investment priorities, to gain a better understanding of whether they actually capture progress that may fall under the remit of the NSA.
4. Discuss findings and limitations

6.2.1.2.2 Second step, analysis of specific output indicators

The second step of the analysis was concerned with steering the focus towards specific output indicators.

NVivo has been selected as a tool for carrying out this analysis.

The first step consisted in importing a database containing the wording of each Specific Output Indicator as well as its relevant attributes (MS, Operational Programme, IP, Target set).

NVivo allows to automatically code open-ended questions in a survey, such as the phrasing of Specific Output indicators, as well as its attributes. Open-ended questions are placed into a single ‘node’ which is the unit NVivo uses to for a collection of information on a specific theme/dimension of analysis.

After having auto-coded nodes and attributes, the ensuing step of the analysis was to fine-grain the coding of specific elements within the phrasing of specific output indicators. In order to do so, NVivo functionalities related to the area of ‘query’ have been used. Commands such as word search and word frequency have proved quite useful to the end of identifying sub-themes within the given set of specific indicators.

A number of keywords, also in combination, have been employed to this end; amongst other:

- Basic skills (basic skills, low-skilled, no skill, ISCED 1, 2, 3 etc.)
- Adults (adult, over 25, over 29, > 29, >34 etc.)
- Guidance and counselling
- Modular training
- Migrants
- Brain Drain
- Support to teachers (teacher, instructor, mentor, staff etc.)
- Validation, Recognition, Certification
- VET (also vocational, training, internship, apprenticeship etc.)

Once identified through the queries, content has been firstly assigned to the relevant node (theme) and then fine-grained manually to ensure that all information in each node was fully relevant to it.

Finally, cross-node queries have been carried out in order to identify content that was relevant under two or more themes/dimensions. This was needed as, e.g., the Upskilling Pathways is not to be intended simply as offer of basic skills training to those lacking them, but rather to adults (not eligible for the Youth Guarantee) under defined conditions of insufficient educational attainment (below ISCED 4). The same applies to other NSA actions; hence also the need to turn on to analysing the choice of more specific indicators.

The coverage and distribution of relevant indicators has finally been observed under relevant dimensions (IPs/MS/typology of region).

The whole analysis has been carried out with a view to understanding how and to what extent may ESF indicators track progress towards the objectives of the NSA.

6.2.2 Findings

6.2.2.1 Step 1 - Common Indicators

As anticipated, Member States and regions (through their Managing Authorities) have selected and defined targets for a number of Common Indicators, both in terms of outputs and results. The first step of the analysis has been to look at common indicators to check whether these may already help assess ESF support to the objectives of the NSA, particularly with respect to relevant NSA actions.

Common output indicators are listed in the ESF regulation as per Table 31 and Table 32, referring to participants and entities, in order.

Table 31 – ESF Common Output Indicators on participants

ID	Name
1	unemployed, including long-term unemployed
2	long-term unemployed
3	inactive
4	inactive, not in education or training
5	employed, including self-employed
6	below 25 years of age
7	above 54 years of age

8	above 54 years of age who are unemployed, including long-term unemployed, or inactive not in education or training
9	with primary (ISCED 1) or lower secondary education (ISCED 2)
10	with upper secondary (ISCED 3) or post-secondary education (ISCED 4)
11	with tertiary education (ISCED 5 to 8),
12	participants who live in jobless households
13	participants who live in jobless households with dependent children
14	participants who live in a single adult household with dependent children
15	migrants, participants with a foreign background, minorities (including marginalised communities such as the Roma)
16	participants with disabilities
17	other disadvantaged
18	homeless or affected by housing exclusion
19	from rural areas

Table 32 – ESF common output indicators on entities

ID	Name
1	number of projects fully or partially implemented by social partners or non-governmental organisations
2	number of projects dedicated at sustainable participation and progress of women in employment
3	number of projects targeting public administrations or public services at national, regional or local level
4	number of supported micro, small and medium-sized enterprises (including cooperative enterprises, enterprises of the social economy)

The first point is that common indicators on entities cannot, by definition, serve the purpose of understanding to what extent actions directed at individuals are progressing towards their targets.

Therefore, there is a need to focus on the 19 output indicators which measure interventions on participants.

Indicator 1-5 focus on the status of participants in the labour market. Because the main actions of the agenda (VET and UP, skills profile for migrants or brain drain) do not select participants based on their employment status, these are not quite meaningful within the context of this study. On a side note, these 5 are logically linked as follows: CO1 + CO3 + CO15 = total number of participants. COI2 and 4 cover a sub section of CO1 and CO3 participants. One point worth mentioning, however, is that Member States are indeed encouraged by the council regulation on the Upskilling Pathways to identify a given target group for the provision of upskilling. This means that, in the future, this possibility should be explored, still bearing in mind that the multiple criteria to access the UP (age + level of skills) will continue to affect estimates.

Next, COI 6 to 8 refer to the age of the participants. The only measure which is contingent upon an age requirement, which can somehow help estimate whether those values count for the objectives of the given action, is the Upskilling Pathway. However, the UP is geared towards individuals which are older than 25-29 years old (depending on national regulation for the provision of the youth guarantee). No common indicator can match this UP requirement. COI 9 to 11 are the most relevant for the NSA. This can be said only for the UP, again, and again with some reservations. In fact, the identification of low skilled as per COI9 and COI10, namely

participants with below ISCED2 educational attainment and participants with below ISCED4, could in principle help proxy the amount people involved in training activities for the low skilled, which is the very purpose of the UP. However:

- the type of support is not specified (could as well be social inclusion measures other than training, even financial support measures);
- participants can be of any age, also in receipt of the Youth Guarantee; and
- ISCED 4 participants are also covered (the UP would allow up to ISCED 3 level)

Thus, this combination is somewhat relevant for the Upskilling Pathways but may substantially overestimate the number of interventions that fall under its specific scope.

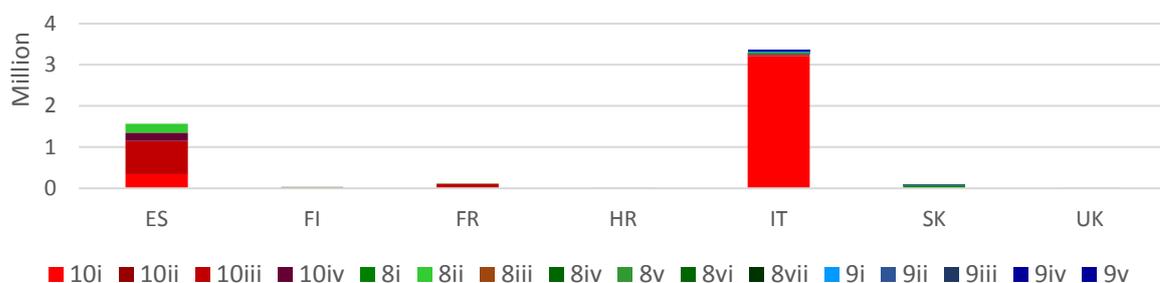
In order to better appraise the extent to which this can affect estimates, an analysis of the distribution of the target values for COI9 and COI10 is provided.

The first step is to assess how many participants with below level 2 ISCED qualification are targeted across all Member States, and under which investment priorities.

A key premise to this analysis is that, differing from specific indicators, common indicators are computed even where these are not selected by the MA, i.e., there is no target set for that indicator. As a result, looking at the selection of Common Indicators is relevant to assess the MAs’ programming choice, while aggregated values of progress may go beyond the selected targets.

Figure 14 displays the 2023 COI targets by MS and IP, thus highlighting those MS that have specifically selected such indicator.

Figure 14 – Cross-IPs distribution of 2023 COI targets on below ISCED 2 participants addressed, by MS



Source: own elaborations based on AIRs2016

The first point is that not all MS appear in this overview. This is because many MS would not have targeted the low skilled with their common output indicators. This is likely due to:

- A number of MS making use of specific rather than common indicators
- A number of MS focusing on other policy priorities.

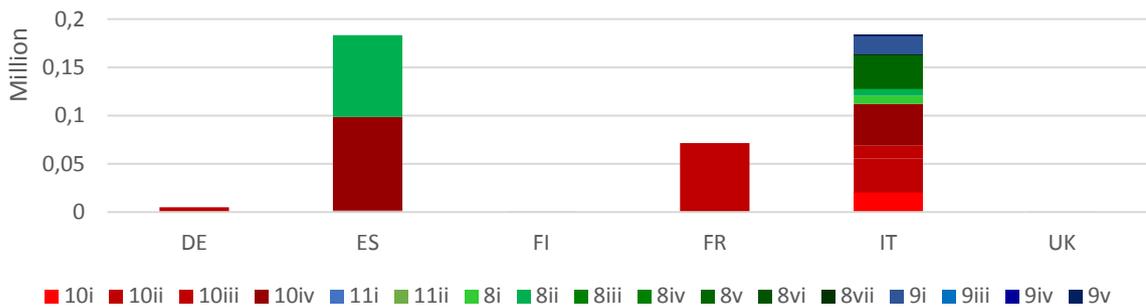
The second finding from this graph is that Italy and Spain show very high targets of the relevant COI, reaching to nearly 5 million participants overall.

However, possibly the most meaningful feature showed by Figure 15 is the cross-IPs distribution of 2023 targets for COI9. In fact, while the largest share of targets is represented

by TO10 investment priorities and this may confirm the focus on upskilling, it is IP 10i to account for the largest share of 2023 target. Because IP 10i is aimed at preventing early school leaving, this clearly points to the targeting of young early school leavers, likely eligible for the Youth Guarantee. This means that adults in need of basic skills are quite likely not the main focus of these interventions, which, in turn, means that much of this distribution is not relevant of the UP.

The second assessment turns on to ESF targets for participants with ISCED level 3 or 4.

Figure 15 – Cross-IPs distribution of 2023 COI targets on ISCED 3 or 4 participants addressed, by MS



Source: own elaborations based on AIRs2016

Although the magnitude of these targets appears significantly reduced, it is still for Italy and Spain to account for the majority of the 2023 targets. It is possible to also observe how non-TO10 targets tend to grow in comparison with targets on below ISCED 2 participants. It is also notable how 10i nearly disappears from this aggregate. Thus, this second representation appears more accurate in terms of themes covered according to the distribution of investment priorities in the graph. However, it is also apparent how ISCED 4 participants cannot be filtered out, although they fall out of scope of the Upskilling Pathways.

If we turn attention to the remaining Common output indicators, similar problems arise with the lack of overlap between their scope and that of the UP.

Finally, because Common Indicators do not distinguish by the type of intervention produced, measures for the promotion of VET could not be proxied for.

Ultimately, if COI are to track progress towards the achievement of the aims of the NSA, it will be so only to a very limited extent and by allowing a significant degree of approximation.

This is particularly as the eligibility for the UP is made up of several conditions (not being eligible for the Youth Guarantee, having below ISCED4 as educational attainment and, at the MS level, belonging to a given priority group), while Common Indicators can observe one dimension at a time. Estimates could be retrieved to some extent by looking at aggregates of COI by investment priority, but these will fall short of being accurate.

6.2.2.2 Step 2 – Programme-Specific indicators through NVivo

Because the main limitation of COI was indeed that they observe only at one feature of the end-beneficiary of the support at a time, as well as they do not distinguish by the typology of support provided, the second step looked into programme specific indicators, which, conversely, may be able to do all of this.

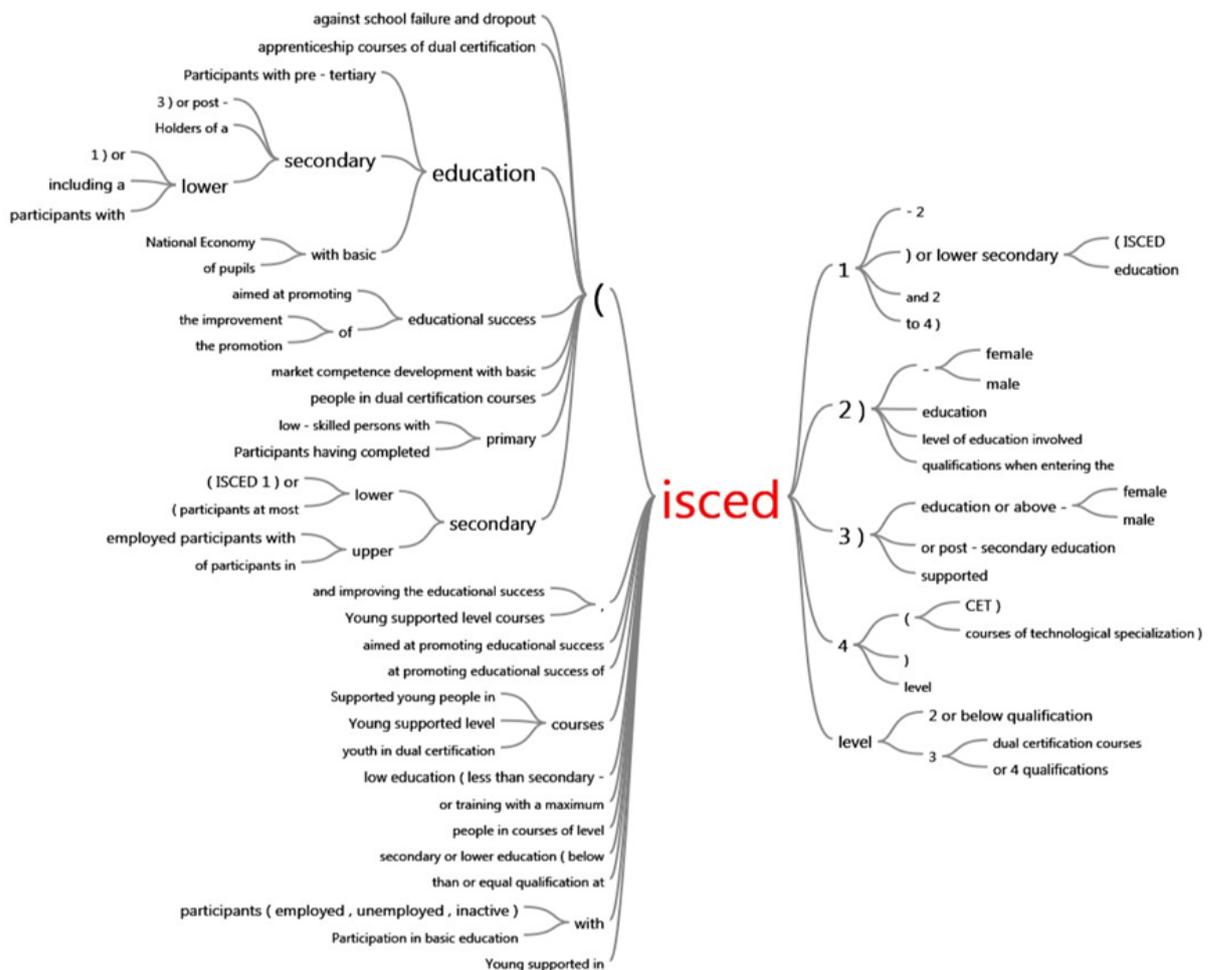
Thus, the second step of the analysis, after having properly instructed NVivo to import the list of all programme-specific indicators translated to English, was to carry out text queries on this data set.

As anticipated, text queries are useful to scan through vast collection of contents, and namely through the over 3200 Programme-Specific Output Indicators under observation.

One interesting feature that comes with text queries is the possibility to create word trees. Word trees are useful to quickly appraise *‘the context surrounding words and phrases from across your data’* and to *‘discover recurring themes and phrases that surround a word of interest that can take you down new paths of investigation and directions for your research’*⁹⁹.

The example below, showed by Figure 16 presents the word tree for the word “ISCED”. The relative query was cleaned of ISCED 5, 6, 7 and 8 retries, in order to focus only on ISCED 1-4 levels, thus on participants with low skills.

Figure 16 – word tree for the word ‘ISCED’



Source: author’s elaboration through NVIVO based on the text of the programme specific indicators as reported in FGB et al, 2016

⁹⁹ QSR NVivo 11 software description

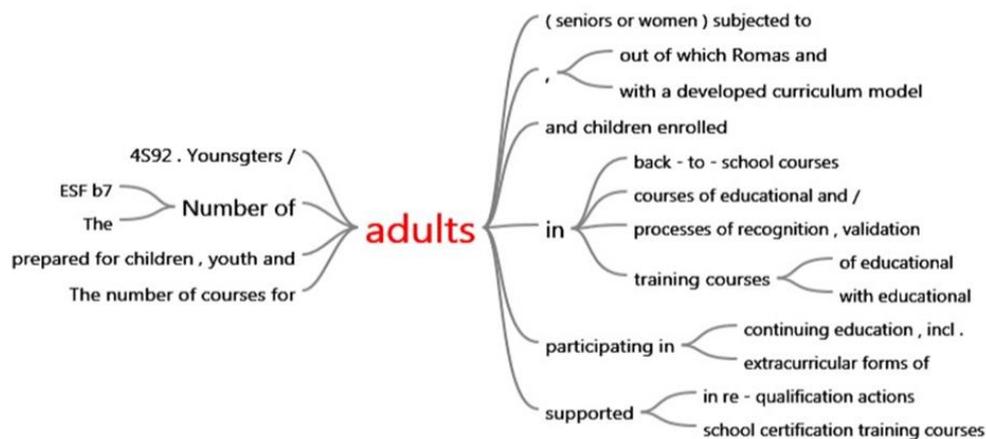
This query resulted in around 60 occurrences. The wording on the left-hand side of the word tree helps have a quick look through them. One surrounding word which presence is rather widespread is clearly ‘young’, followed by ‘dual courses’ which are typical of VET and education. Sometimes indicators refer to the nature of the intervention (e.g. apprenticeship) some other times to the aim of the intervention (e.g. improving/promoting educational success), or to the employment status of the participant. Wording on the right-hand side of the figure refers only to the specific ISCED level addressed. The specification ‘adults’ is entirely missing, which is rather suggestive of the extent to which specific indicators can actually track progress towards the objectives of the Upskilling Pathways. As shown by Figure 18 results with the word query ‘low skills’ confirm this trend.

Figure 17 – word tree for ‘basic skills’

Participants without — basic skills

When turning to the word “adults”, the word tree shows consistent results.

Figure 18 – word tree for ‘adults’



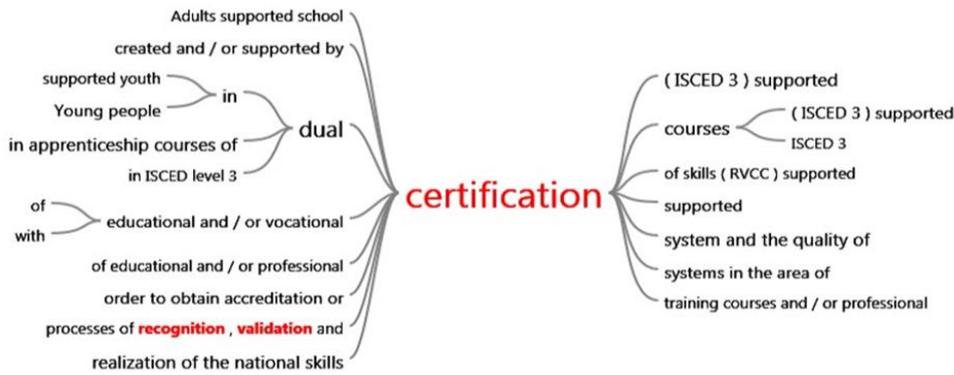
Source: author’s elaboration through NVIVO based on the text of the programme specific indicators as reported in FGB et al, 2016

On the left-hand side of the tree 2 out of 4 recurrent formulation do not single out adults (which are considered alongside other age classes). The right-hand side of the graph shows the typologies of interventions monitored, especially from the centre to the bottom right-hand corner. A couple of formulations worth mentioning are those relating to processes of recognition and validation, fully in line with the UP, as well as re–qualification actions.

As to the specific characteristics of the Upskilling pathways, the query for ‘skills assessment’ and ‘skills audit’ results in no items. Modular training appears four times, 3 in 4 linked to VET but the remaining one - Participants from disadvantaged groups, including LTU with low qualifications in certified modular training’ has a decent correlation with the structure of the UP.

Interestingly, the issue of certification is linked quite frequently to low levels of skills (e.g. ISCED 3); nevertheless, the audience targeted appears rather wide, as shown by Figure 19 below.

Figure 19 – word tree for ‘Validation, Recognition and Certification’



Source: author’s elaboration through NVIVO based on the text of the programme specific indicators as reported in FGB et al, 2016

As to the accompanying measure of the UP, such as outreach and guidance, the relevant queries result in a quite high number of occurrences (44).

Figure 20 – word tree for ‘Counseling and guidance’

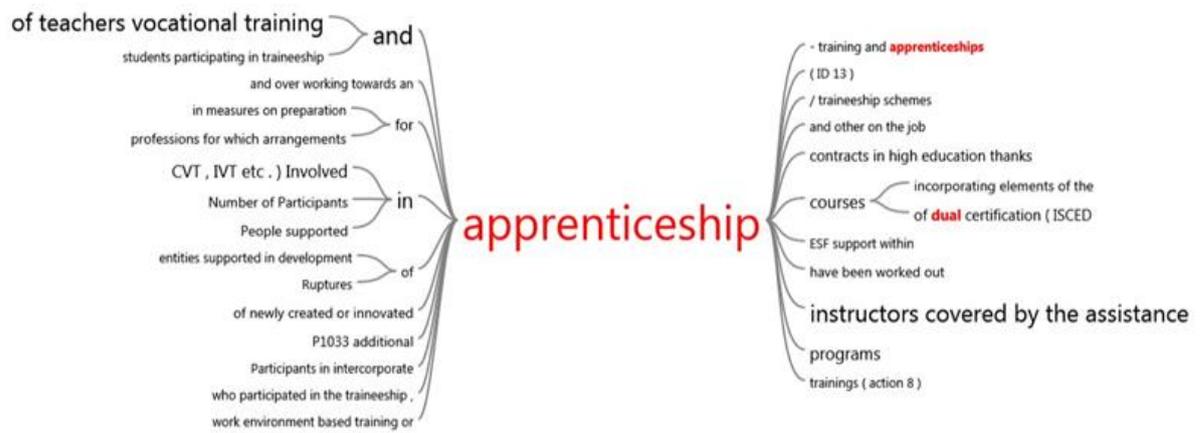


Source: author’s elaboration through NVIVO based on the text of the programme specific indicators as reported in FGB et al, 2016

However, the word tree displays quite clearly how the appearance of adults and adult learning in this context is seldom at best. Conversely, VET accounts for a large share of indicators linked to the issue of counselling. In addition, the issue of certification appears to be often couple with that of apprenticeship and dual learning. This emphasises how the NSA action for the enhancement of VET systems ties to and overlaps with the Upskilling Pathways to a large extent. Other than belonging to the same work strand, it is indeed apparent that basic skills training should be offered by different means, including VET and dual learning experiences, with a view to progressing towards a qualification. Likewise, VET should make use of the same guidance system that is advocated for in the UP.

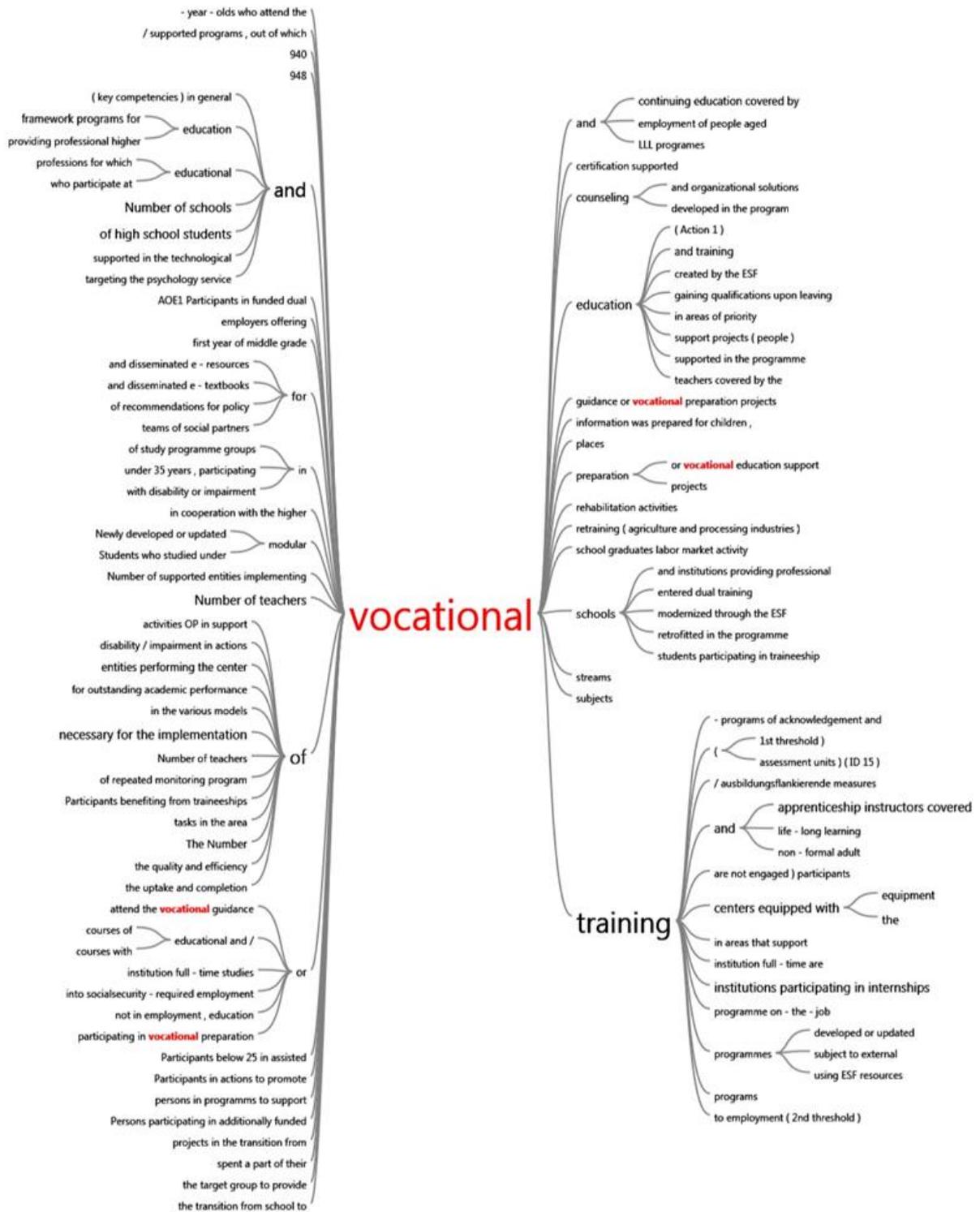
The area of apprenticeships and dual learning (89 retriees) and that of VET (126), are particularly well overlapping, as shown by their respective word trees (in Figure 21 and Figure 22 below).

Figure 21 – word tree for ‘Apprenticeship, traineeship, internship’



Source: author's elaboration through NVIVO based on the text of the programme specific indicators as reported in FGB et al, 2016

Figure 22 – word tree for ‘VET’



Source: author's elaboration through NVIVO based on the text of the programme specific indicators as reported in FGB et al, 2016

More generally, the issue of VET is – somewhat unsurprisingly – very well covered by ESF Programme Specific Indicators. In fact, the NSA action ‘making VET a first choice’ is a rather a quest to re-state the importance of VET, as for the Riga Conclusion on the same issue, then

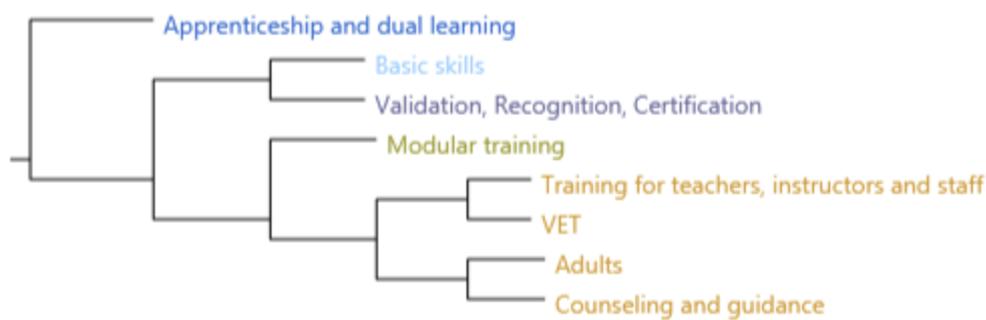
an entirely new initiative. Being a slightly less innovative measure, it is easy to see how the 2014-2020 ESF programming (and thus monitoring) is quite aligned to its objectives.

VET seems specifically geared towards youths; however, the left-hand side of the word tree displays a number of additional categories of individuals, including teachers, and also entities (schools, social parts and partnerships) and interestingly projects and programmes to enhance the quality of VET. Issues such as counselling are mentioned, as well as certification, training of instructors and other, consistently with what described above.

That of support for teachers, instructors and staff is another theme which expectedly cuts across VET and Upskilling. Similarly, also the issue of counselling and guidance have overlaps with the two main NSA actions. This is mirrored in their respective word trees (omitted for the sake of synthesis).

As to the overlaps of contents for indicators relating to either the UP or VET, NVivo allows also to compare the wording of the specific nodes/themes where relevant content has been coded, to measure their similarity.

Figure 23 – themes clustered by word similarity



Source: author's elaboration through NVIVO based on the text of the programme specific indicators as reported in FGB et al, 2016

There appears indeed a bit of a mix of interventions that would be specific for VET and other that are more germane to the Upskilling Pathways. While apprenticeship and dual learning remain on a dedicated branch of the tree, basic skills is coupled with validation, recognition and certification, to confirm the impression gathered from their respective word trees.

These similarities stress how interlinked these two NSA actions are in achieving their objectives.

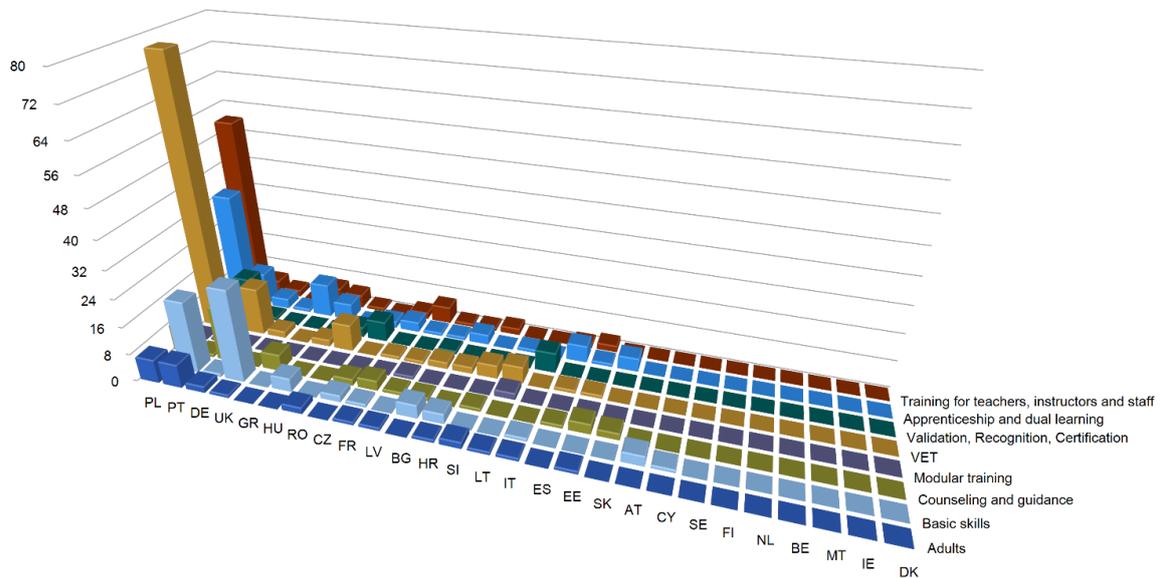
Finally, as opposed to the main action here discussed – the UP and VET -, queries on the sharing of best practices on brain drain and on the tool for recognising skills for third country nationals returned no relevant results.

This reinforces the impression that the former actions have a distinctive magnitude through the ESF. They are, however, covered by its specific indicators to a different extent, the UP appearing to be only marginally matched by the wording of ESF Specific Output Indicators, while VET looking more in line with them.

At any rate, the former considerations cannot go further than observing how the wording of specific indicators may mirror that of relevant NSA intervention. In what follows, an attempt is made at estimating the relative frequency of relevant programme-specific indicators.

The first step is to look at the larger collection of coded indicators, that is, all indicators identified through the queries presented above.

Figure 24 – number of relevant indicators by theme (node) and MS

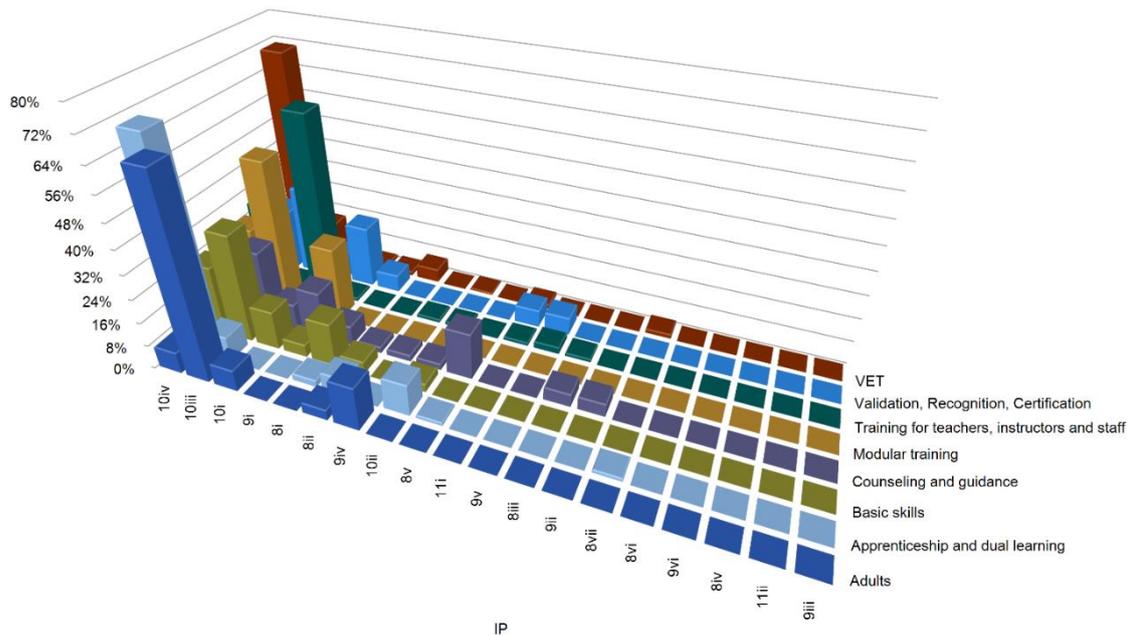


Source: author's elaboration through NVIVO based on the text of the programme specific indicators as reported in FGB et al, 2016

The first graph shows the distribution of indicators aimed at tracking a number of outputs that are to differing degrees relevant to the NSA, by MS. Already at a first glance, the coverage of indicators referring to the themes listed along the z axis appears quite modest. Almost 10 countries have nearly zero relevant programme-specific indicators. At the opposite end of the graph, countries like PL, PT, DE and UK have in some cases a rather widespread presence of relevant programme-specific indicators. This is also due to the fact that, as aforementioned, some countries, such as Poland, have made a large use of Programme-Specific indicators and have selected more limitedly common indicators.

Figure 25 goes on to check on which investment priorities relevant programme-specific indicators are set.

Figure 25 – percentage of relevant programme-specific indicators by theme (node) and IPs



Source: author's elaboration through NVIVO based on the text of the programme specific indicators as reported in FGB et al, 2016

Relevant IPs appear to be those anticipated in the review of the NSA, and namely those of TO10, accompanied by relevant TO8 (ip 8i, 8iii and 8v) and TO9 (9i, 9iv, 9v) investment priorities.

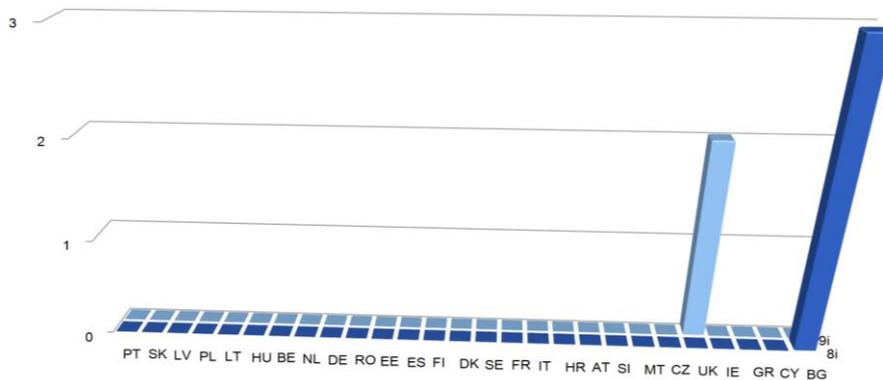
While the above graphs provide for a much more accurate picture of ESF indicators relevant for the NSA than that at the level of Common Output Indicator (Figure 24 and Figure 25 their wording being more akin to the aims of the NSA, they are not yet quite in line with the specific requirements set by the Upskilling Pathways, which scope is rather articulated and calls for a use in combination of more than one of the themes taken into account thus far.

Accordingly, the final step of this analysis is to cross-code relevant content that is common to two or more dimensions necessary to define the scope of the UP, e.g., 1. adults, 2. with low skills, 3. in basic skills training that received a support shaped around a three step approach (skills assessment, modular training and then validation of skills) with accompanying measures such as guidance and outreach.

Given the very high level of detail brought about by this fine-tuning operation, the number of relevant indicators under such strict conditions drops drastically.

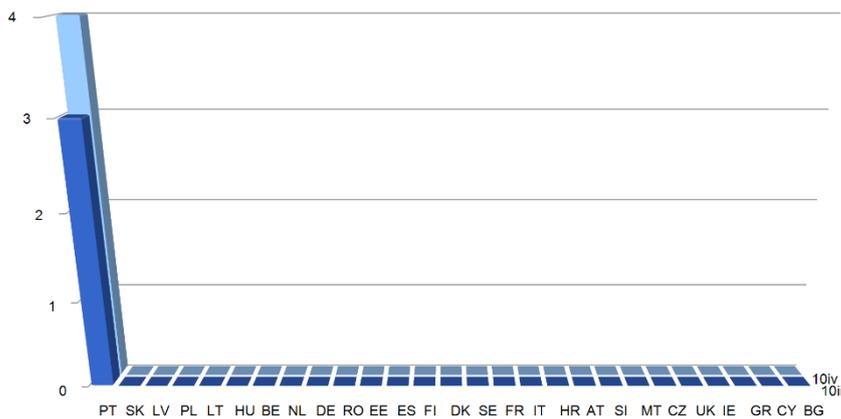
Figure 26 and Figure 27 below show the MS distribution of some combination of criteria particularly suitable for the UP, 'adults and basic skills' and 'basic skills and certification' respectively.

Figure 26 – number of programme-specific indicators cross coded on two themes, namely *adults and basic skills* by MS and IPs



Source: author’s elaboration through NVIVO based on the text of the programme specific indicators as reported in FGB et al, 2016

Figure 27 – number of programme-specific indicators cross coded on two themes, namely *basic skills and certification* by MS and IPs



Source: author’s elaboration through NVIVO based on the text of the programme specific indicators as reported in FGB et al, 2016

It is apparent how figures plummet and how only a small number of countries under a small number of investment priorities can provide fine-grained information to monitor progress relevant for the UP. If the same cross-coding/ cross-thematic operation is replicated with more than two criteria (e.g. adults, basic skills AND certification), then the queries end up finding no relevant results.

This fine-graining operation is less compelling when it comes to the analysis of ESF indicators relevant for VET in that its aims are much broader in scope and wind up encompassing nearly all actions carried out for the enhancement of VET under the ESF.

6.3 Conclusions

In light of the importance placed upon human capital development in the EU2020 strategy, underpinned by the various benefits thereof, and bearing in mind the increasing and widespread emphasis on monitoring and evaluation of public policies and expenditure, the main goal of this

analysis has been to check to what extent the NSA, arguably the most significant and recent initiative for skills development in the EU, is currently monitored (or could be monitored in the future).

The starting point of the analysis was that the NSA comes with no dedicated monitoring system attached, in that it is not endowed with a dedicated budget. However, the NSA does aim at steering funding from other EU instruments towards the achievement of its objectives. Thus, the analysis went on to check whether NSA's most relevant supporting fund, the ESF, may allow tracking such progress.

Another key element brought to the fore by this analysis is that not all NSA actions could be deemed as needing a monitoring system at an individualized level, in that progress of more 'systemic' actions, such as the revision of the EQF or of the Europass Framework – to name but a few – is suitable to be observed on an ad-hoc basis. Thus, this analysis focused on two main rather individualized actions of the NSA, namely the Upskilling Pathways (former EC's skills guarantee) and 'making VET a first choice'. Another two actions were considered partly relevant for the monitoring of individualized support, notably the 'sharing of best practices on brain drain' as well as the 'skills profile tool for third country national'. However, these two are much more limited in scope and could be considered only to some extent, as they will also rely on other EU funding.

The ensuing analysis on ESF Common output indicators drew attention to their very broad formulation, which fails to disentangle figures on ESF interventions that can be reliably deemed relevant for the NSA and aligned to its priorities.

Only one Common Output Indicator was found to be relevant, to some degree, to give an indication on how interventions targeted at the low-skilled are unfolding. Yet this can only provide a broad idea of how interventions relevant for the Upskilling Pathways, the flagship initiative of the NSA, are progressing towards their targets.

Because MS have made vast use of Programme-specific indicators, this analysis sought to ascertain whether their more specific formulation could provide better support for monitoring relevant NSA actions.

The resulting analysis first clarified that there is some scope for tracking progress towards the aims of the NSA action on VET, as well as to the Upskilling Pathways, while nothing can be done for the remaining two NSA actions considered relevant for a monitoring at individualised level.

Nevertheless, the analysis pointed quite clearly to the lack of appropriate coverage of the ESF monitoring system based on Programme-specific indicators, first of all because some countries make a much wider use of these and their formulation seems quite aligned to the priorities of the NSA, but for many others that is not the case.

This is particularly true for the 'Upskilling Pathways', which scope is rather articulated. When checking for programme-specific indicators taking into account all the dimensions of the UP (age, level of skills, type of support) only a few indicators were found.

The picture is relatively more favourable when it comes to VET intervention, probably due to the fact that ‘make VET a first choice’ is slightly less innovative than the Upskilling Pathways, and as such its main aims were already reflected through the ESF 2014-2020. In any event, programme-specific indicators could not serve the purpose of providing an overall picture of how the NSA is unfolding, given their acknowledged limitations when it comes to aggregability.

On a side, though relevant, note, this analysis shed light on the strong interlinkages between VET and the UP, which is an element to bear in mind also when considering the monitoring for these two actions.

All in all, this paper provide some evidence to argue that, although the themes of the NSA are matched, at large, by the monitoring system of the ESF, no organic monitoring system can track progress towards the objectives of the NSA, and some efforts should be fostered at MS level to set up appropriate frameworks to collect relevant data.

These may leverage on existing ESF monitoring systems to some extent and in some countries more than in others; in particular, micro-data may, if made available and aggregated, be used with conditional distributions of participants (i.e., age AND ISCED level, including checking for given typologies of intervention). However, these are not reported in Annual Implementation Reports, which is the main monitoring document of the ESF and the main source of information for this analysis. In addition, micro-data is expectedly not fully comparable, as there is no fixed form set by the regulations. Thus, this issue would require further research in order to identify:

- the coverage and quality of microdata; and
- any issues related to data aggregation/standardisation, including administrative burdens thereof.

7 Annex III – Additional details on the experimentation and its background

7.1 A brief description of the RHOMOLO model

7.1.1 Background: a brief taxonomy of General Equilibrium models

Economists – but also policy makers and the society at large alike – have long been confronted with the quest for understanding “*what policy/economic model/strategy works, and for whom*”.

Amongst the vast array of tools that exist to answer to such challenging questions lie General Equilibrium models, which can be further broken down into different branches. The main two are:

- Computational General Equilibrium (CGE) models
- Dynamic Stochastic General Equilibrium (DSGE) models

One should probably start by saying that there is no strict definition of what “General Equilibrium model” means in practice. However, criteria have been identified (see for example Rutherford, 2000) that help describe recurring features of a general equilibrium model.

The starting point is that of multiple interactive agents, whose behaviour is rooted to the extent possible in microeconomics – thus they act to optimize their decisions via maximizing their utility functions. Next, their interactions do not happen in void but are mediated by markets and prices. Often these models are disaggregated (by sector, agent, market) but feature a necessarily limited number of parameters and data to mimic real-life interactions and dynamics. This entails assumptions being made that simplify the computational intricacies but may, at the same time, be conducive to unrealistic conclusions.

Typically, a trade-off exists between the complexity, and thus granularity, of these models and the extent to which they can remain tractable, not only from a computational standpoint but also from that of practitioners having to interpret their results. This limitation implies that GE models usually take different specifications and forms depending on their focus, be that a given theme, a given point in time, or a given territorial entity, and the actual questions they seek to answer. In other words, one model can be fitted with a more sophisticated, say, transport sector if its aim is to answer transport-related or spatially relevant questions; but will necessarily feature less articulated working mechanisms to replicate the dynamics under different dimensions of analysis.

In terms of the economic theory on which these models are based, usually general equilibrium models are neo-classic in spirit (based on the Real Business Cycle idea that economic agents will seek to efficiently adapt to exogenous shocks, and this will lead to long term equilibrium). However, they often allow for some “Neo-keynesian” specifications, e.g. with agents being confronted with frictions and constraints that hamper convergence towards market clearing values and leave the door open for a role for monetary policy or various degrees of imperfect competition in different markets.

As to the different branches they can belong to, the key difference between CGE models and DSGE models lie in the way they deal with expectations and dynamics, as well as are

“calibrated”; that is, in the way they factor in the “starting point”, as well as the way agents take decisions. While CGE models employ real data at a given point in time to define the baseline and then let the model work the way this can evolve over time/according to a set of policy shocks applied to it, DSGE generally take into account a range of parameters, based on past values, in a probabilistic manner (not just as points) and across a longer time-span, which allows them to entail uncertainty. This, in turn, means higher scientific credibility and explain their growing use. In a similar fashion, while CGE models struggle to feature forward-looking rational agents given to the fact that the dynamics they include are de facto iterative static equilibria, DSGE do normally include them.

Because of a somewhat simpler functioning mechanism, CGE can, on the other hand, include a larger number of markets/sectors and thus are often used in regional policy analysis. As anticipated, CGE models can be static, meaning that they would proxy only one step forward in time and confront the baseline with the estimated scenario, or dynamic, thus including expectations of economic agents, which however normally take simple forms (myopic/backward looking and so forth), sluggish adjustments towards market clearing/steady state values etc.

Again, the form and specification of a given GE model is affected to a large extent by its intended use.

7.1.2 The JRC’s RHOMOLO model¹⁰⁰ - main features

Below a range of features of the model are summarised, with additional detail dedicated to those that directly affect the modelling of human capital policies, thus particularly the labour market:

- Agents:
 - **Households** work along the lines of the representative agent, i.e., there is one single household per each region, that is equipped with 4 productive factors, namely high-skill, medium-skill and low-skill labour and private capital.
 - **Firms:**
 - a) *Production function*: the production function of firms within the model is that of a 3-level CES (Common Elasticity of Substitution). The total production is in fact generated by:
 - (i) Factor inputs (as a function of TFP - the total factor productivity), of which
 - Capital (TKP - Total capital productivity)
 - Private
 - Public
 - Labour (TLP – total labour productivity)
 - Low skill
 - Medium skill
 - High skill
 - (ii) Intermediate inputs

¹⁰⁰ Based on the various documents published by the JRC (see references), and particularly Mercenier et. Al (2016)

- Composite good 1
- Composite good ..
- Composite good N
- b) *Competitive behaviour*: while in the perfectly competitive markets prices equal marginal costs and no extra-profit can be made, in the imperfectly competitive markets, firms place a mark-up over their marginal costs. They optimize their production functions conditional on “*assumptions on the competitive game, as well as on the properties of the demand curve that these firms face*” (Mercenier et al., 2016:14). Importantly, markets are regionally segmented and firms have a given market power, which, both, affect mark-ups and optimum conditions.
- c) *Labour demand*: firms in regions and economic sectors create labour demand which depends on the price of capital and wage rate. Labour demand is disaggregated by skill level, with an elasticity parameter that determines the substitutability of labour by skill, together with the wage rate and shift and share parameters.
- **Government**: set at the regional level, it:
 - a) Earns from taxes applied to households’ income, firms’ production and input factors
 - b) Consumes local final goods, to a measure that can be endogenized depending on model’s specifications according to budget constraints (a fixed deficit)
 - c) Transfers money to households (set at a constant level in real terms)
 - d) Has fixed investment with no depreciation in public capital
- **Markets and sectors**
 - **Labour market**: in RHOMOLO v3, there are three different options for the labour market, namely a static wage curve,¹⁰¹ a dynamic wage curve and a Phillips curve. These can be used alternatively shifting the parameters of the baseline equation for the wage curve below

$$rw_{r,e,t} = a_e + \alpha rw_{r,e,t-1} - \beta u_{r,e,t} + \zeta \Delta p_{r,t} - \theta \Delta u_{r,e,t} + \omega \Gamma_t$$

Where

- “ rw ” is the real wage, negatively correlated with unemployment levels and positively correlated with the price of output and the productivity trend
- “ $u_{e,t}$ ” is the unemployment rate
- “ $\Delta p_{r,t}$ ” is the price of output

¹⁰¹ A wage curve is a reduced-form representation of structural models of imperfect labour markets, such as union wage bargaining models, efficiency wage models, or matching models labour markets

- “ Γ_t ” is the productivity trend
- “ e ” is the differentiation by skill level,
- “ r ” is the region
- “ a ” is an exogenously fixed productivity parameter

This means that the level of unemployment is endogenized and made dynamic through a wage curve. In other words, unemployment levels are linked to real wages but also to changes in output prices and past real wages. If the parameters α , ζ and θ are set to zero, then the curve becomes static, and real wage only depends upon unemployment rate. If then α is set to 1 the equation becomes a Phillip curve, with changes between rw in $t=0$ and rw in $t= -1$ are dependent from levels of unemployment in $t=0$ and unemployment and employment in absence of changes to exogenous parameters (i.e. the volume and composition of the labour force) come back to the steady state. Importantly, the β coefficient, which is basically the slope of the wage curve, is borrowed from the literature¹⁰² and is not differentiated across regions. The same applies to the dynamic parameters.

Labour demand is determined by the firms, by region, sector and skill level, and, at clearing conditions, the situation is as per the equation below:

$$\sum_j N_{r,j} ld_{r,j,e} = (1 - u_{r,e})L_{r,e}$$

Where the labour demand (ld , by region economic sector and skill level) for all firms (N , by region and economic sector) equals the labour force (by region and skill level, considering the unemployment rate). Importantly, labour supply is determined exogenously and assumed not to vary (no natural population change). Nevertheless, there is a module for labour mobility originally developed in Persyn et. al (2014), “*where workers' migration is governed by expected differences in the real incomes, and is also dependent on the probability to be employed in a given region.*” (Lecca et. Al 2018: 16). This means that the labour supply in $t+1$ equals the original labour supply plus the net migration rate ($m_{r,e}$), which is, in turn, contingent on a vector of features of the regional socio-economic context such as unemployment, real wage, distance etc.

Thus, the labour market supply evolves as in

¹⁰² The value of β is set at 0.1, based on Nikjamp and Poot (2005) and typical of a long-run wage curve. The dynamic parameters α , ζ and θ are set to 0.1, 0.25 and 0.03 respectively, based on Montuenga-Gómez, and Ramos-Parreño, 2005 and Nymoer and Rødseth, 2003.

$$L_{r,e,t} = L_{r,e,t-1}(1 + m_{r,e})$$

So, the model allows controlling for the effect of policies also in terms of how these affect skills mobility.

- **Goods/commodities:** There exist both perfectly competitive but also imperfectly competitive markets. The former is characterized by homogeneous goods (constant returns to scale, goods can only be differentiated by geographic area) while, in the latter, goods are differentiated (increasing returns to scale, differentiation not only by origin but also by the features of the product).
- **Transports:** transport costs are “iceberg type”¹⁰³ and are sector- and region-pair specific. This implies a 5 x 267 x 267 asymmetric trade cost matrix derived from the European Commission’s transport model TRANSTOOLS. It is asymmetric in the sense that costs from and to a given region may differ.
- **Spatial effects:** there are three notable effects which determine localization decisions of firms. These are largely determined by transport costs and, notably:
 - the market access effect: proximity to areas with high customers’ density attracts firms while heightening competition
 - the price index effect: proximity to providers lowers costs but also prices and implies higher competition
 - the market crowding effect: too much proximity compresses mark-ups due to competition effects.

As anticipated, skills mobility is also endogenized as a function of unemployment, real wage and distance between regions, by skill group.

- **Intertemporal issues and expectations:** “RHOMOLO assumes that the behaviour of inter-temporally optimising agents depends only on the current and past states of the economy, not on the inter-temporal optimisation.” It is in this sense myopic, with backward looking agents. Future versions of the code will allow for rational (forward looking) expectations (Lecca et al., 2018)

7.2 Detailed procedure for bringing ESF monitoring data at the NUTS2 level

As mentioned in the text, bringing ESF monitoring data at the NUTS2 level is not a trivial operation. A number of steps have been followed to this end, notably:

1. Reconstructing eligibility of EU NUTS2 regions for the cohesion policy 2014-2020
2. Retrieving AIR Table 7 data on the spend at the NUTS0 to 3 data

¹⁰³ Iceberg type costs imply that a fraction of the volume of goods shipped “melts” over the transfer. Thus, the total volume of goods that reaches its destination is diminished in comparison to that which leaves the production site.

3. Dealing with different NUTS classification issues
4. Identifying those Operational Programmes that showed expenditure at the NUTS0 or 1 level
5. Identifying criteria to breakdown spend at the NUTS2 level
6. Writing algorithms (excel macros, VBA) which could automatically distribute expenditure at the NUTS2 level based on the above criteria
7. Testing the resulting figures at the NUTS2 level (so that they matched the original values)
8. Re-aggregating expenditure by NUTS2 region (different OPs contribute to the same region)
9. Calculating output and results based on the relevant NUTS2 shares

7.2.1 Reconstructing eligibility of EU NUTS2 regions for the cohesion policy 2014-2020

The starting point is to check eligibility criteria of the 2014-2020 cohesion policy for all NUTS2 EU regions.

This was done based on the information included in the COMMISSION IMPLEMENTING DECISION of 18 February 2014 setting out the list of regions eligible for funding from the European Regional Development Fund and the European Social Fund and of Member States eligible for funding from the Cohesion Fund for the period 2014-2020 (notified under document C(2014) 974).¹⁰⁴ Information is included according to NUTS2006 and 2010 codes, as the first categorisation is done on the basis of the former but if a region has been merged, it should be ascribed to a given category of region based on the NUTS2010 codes.

The merged list was included in the dataset.

7.2.2 Dealing with different NUTS classification issues

As the data from AIR2019 can be a mixture of NUTS2006 and 2010, a bottom down approach was used, thus starting from the 2006 configuration some changes have been applied on a one to one basis so that it would match with the 2010 NUTS wherever necessary.

7.2.3 Retrieving AIR Table 7 data on the spend at the NUTS0 to 3 data

The regionalised financial data can be found in Table 7 of the Annual Implementation Reports drafted by Managing Authorities. It details, by operational programme, priority axis, investment priority and NUTS3 region (whenever available) the following info:

- total eligible cost of operations selected for support
- total eligible expenditure declared by beneficiaries to the managing authority
- number of operations selected

¹⁰⁴ Available at : <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32014D0099>

The focus is however on the second of those mentioned, that is, in simple terms, expenditure contracted out to beneficiaries and certified.

Data at the NUTS3 region was aggregated at the NUTS2 level. But not all programmes report financial progress at such a low level of regional breakdown.

7.2.4 Identifying those Operational Programmes that showed expenditure at the NUTS0 or 1 level

The following step has been to filter observations which showed expenditure at the NUTS0 or 1 level, to understand for which macro-regions and MS specific NUTS2 allocation criteria were necessary.

7.2.5 Identifying criteria to breakdown spend at the NUTS2 level

The rationale there has been to stay as proportional as possible to the original ESF/YEI allocation, based on the assumption that a sensible proxy of NUTS2 progress could be to ascribe progress to regions within the same Operational Programme, category of region and fund pro-quota allocation.¹⁰⁵

7.2.6 Writing algorithms (excel macros, VBA) which could split automatically distribute expenditure at the NUTS2 level based on the above criteria

As the SFC extraction with financial progress at the NUTS0 to 3 level included some 8.5 thousand non zero observations, which should have been split across all the relevant regions in the case of NUTS0 and 1 values, a VBA script became necessary to finalise the process. The rationale for the code was quite simple: for every NUTS0 and 1 observation, the row, after having been copied to a new worksheet, should be repeated for all the relevant regions (which are just a few in the case of NUTS1, or the entire list of a given MS in the case of NUTS0 values) with the value of financial progress multiplied by the relevant coefficient to be found in the categorisation of NUTS2 region sheet. No operation other than duplication is foreseen for NUTS2 or NUTS3 values.

7.2.7 Testing the resulting figures at the NUTS2 level (so that they matched the original values)

Once the values were produced and included in a new worksheet, these were tested against the original allocation by OP. The main idea is that the new allocation by OP should be equal to the old one reduced by that at the NUTS0 and 1.

A number of adjustment proved necessary so that the data could match, also due to the fact that sometimes information in AIR Table 7 is misclassified (e.g. regions that are assigned to

¹⁰⁵ In particular:

for the ESF, shares of allocation (within the same category of region) were based on the region's total population, which is a key calculation figure criteria for the financial envelopes according to articles 1-4 of Annex VII of the Common Provision Regulation Reg. (EU) 1304/2013

for the YEI, the simulations of allocated expenditure as per art. I annex VIII of the CPR (based on relative distances of 2011 and 2012 to the EU average youth unemployment data). In this case the data elaborated for the study supporting the Impact Assessment of Human Capital Investments shared with the JRC in February 2018 was used.

different categories of regions for different operations, wrong NUTS code, NUTS2 regions with just a 2 digit NUTS code, etc.)

Once checked, a clean version was produced.

7.2.8 Re-aggregating expenditure by NUTS2 region (different OPs contribute to the same region) and computing shares

As different OPs now contributed to the same NUTS2 region but the interest is into ultimately knowing what is the progress at the NUTS2 level, the data was re-aggregated through a pivot table. The final step was to calculate the shares by NUTS2 region. This is because, as already clarified, the data on outputs and results is at the level of OP and fund/category of region. So each observation there needs to be split into the relevant NUTS2 regions, but only within the relevant category of region and MS.

7.2.9 Calculating output and results based on the relevant NUTS2 shares

For the final step, that of transforming OP output and results data into NUTS2 data, a new worksheet was created, which included creating another VBA script that could do the split much like it had been for splitting YE financial progress at the NUTS2 level.

The allocation criteria, as anticipated, is the relative share of eligible expenditure declared by beneficiaries within an OP, by NUTS2 region.