



Regione Emilia Romagna

Servizio di valutazione indipendente del POR FSE 2014–2020

Valutazioni tematiche on-going relative alle priorità degli assi del Programma

**TEMA 2 - EFFICACIA DELLE POLITICHE FORMATIVE E DELLE
POLITICHE ATTIVE DEL LAVORO PER SOSTENERE L'INSERIMENTO E
IL REINSERIMENTO NEL MERCATO DEL LAVORO CONTRASTANDO
LA DISOCCUPAZIONE DI LUNGA DURATA –**

VALUTAZIONE DI IMPATTO CONTROFATTUALE

RTI Iris S.r.l. – Performer S.r.l.

(30 aprile 2022)



performer

INDICE

1	Introduzione.....	3
2	I destinatari degli interventi formativi del FSE e il profilo dei percorsi.....	5
3	Perché l'analisi controfattuale per la formazione e il lavoro.....	13
4	la tecnica del Propensity Score Matching.....	16
5	Primo trattamento dei dati amministrativi.....	18
6	Disegno dell'esperimento e operazioni per l'allineamento degli archivi.....	19
7	Stima dell'impatto della formazione FSE sull'occupazione.....	28
8	Considerazioni conclusive e prospettive applicative.....	42

In Italia continua a crescere la necessità di una formazione efficace. Si può vedere il problema da diversi punti di vista. Le imprese continuano a dichiarare la loro difficoltà a reperire candidature adeguate per le assunzioni. Negli ultimi anni, secondo le indagini di Excelsior la percentuale di posti potenziali che restano tali per carenza di candidati è cresciuta, fino al livello del 40,4% dei profili difficili da reperire secondo l'ultima rilevazione Union-camere Anpal, che vengono addirittura dichiarati introvabili. Non si tratta solo di ingegneri o di specialisti ICT: ormai l'elenco di figure difficili da reperire tocca un ampio bacino di profili, anche per occupazioni operaie e tecniche a media specializzazione e per lavori sui quali riemerge una scarsa disponibilità da parte dei giovani, che mostra un tratto culturale di distanza e scarsa conoscenza di molti contesti di lavoro oggi.

Sull'altro versante, i giovani manifestano evidenti disagi a vari livelli. L'Italia, secondo il rapporto BES appena uscito, è nella posizione peggiore tra i 27 paesi europei per quanto riguarda i NEET, giovani tra 15 e 29 anni che non sono inseriti nel lavoro e che non partecipano ad attività formative e che sono quindi ad elevato rischio di esclusione sociale. Dopo aver toccato nel 2014 – al culmine della crisi occupazionale – il livello del 26,2%, 10 punti percentuali al di sopra della media Ue27 – ora la percentuale di NEET è poco sotto il 25% dell'intera fascia di età e non si riduce. Inoltre, le indagini Pisa sulle competenze mostrano ampie aree di giovani che non riescono a raggiungere livelli accettabili nell'area dell'italiano e della matematica, competenze base per tutte le occupazioni secondo il nuovo modello europeo varato nel 2018.

Infine, l'Italia continua ad essere un paese dove le politiche attive sono poco praticate ed efficaci e per gli adulti che perdono il lavoro sono stati largamente utilizzati ammortizzatori passivi, che se sono efficaci nell'attenuare i problemi di reddito, soprattutto per la parte più tutelata del mercato del lavoro, sono tuttavia inefficaci a sostenere le persone nelle operazioni di riconversione e riqualificazione professionale, sempre più richieste nei mercati del lavoro detti transizionali, perché caratterizzati da rapidi cambiamenti ed elevata velocità nella trasformazione di ruoli professionali e competenze, dove le persone devono riuscire a muoversi da un lavoro all'altro in forza del continuo adeguamento delle loro competenze.

La pandemia di COVID-19, che ha ostacolato l'efficacia delle attività scolastiche e formative, soprattutto nella loro parte 'applicata' di tirocini e alternanza, e l'accelerazione della digitalizzazione e dell'orientamento verso la sostenibilità, che hanno ulteriormente arricchito e curvato la domanda di competenze verso nuovi mix e nuove ibridazioni, hanno aumentato l'urgenza di tutte le azioni che possono aumentare il rendimento della formazione, anche nei luoghi dove essa ha già una qualità elevata.

Giovani, adulti e imprese rappresentano quindi gli attori che traggono maggior beneficio dai livelli di efficacia della formazione finanziata innanzitutto dal Fondo Sociale Europeo, che per dimensione e missione rappresenta una leva strategica proprio sui punti macroscopici che abbiamo brevemente ricordato. Peraltro fin dagli anni '90 la strategia europea per l'occupazione aveva messo al centro della sua azione i pilastri dell'adattabilità, occupabilità, imprenditorialità e pari opportunità, che avevano dall'inizio individuato la funzione di fattore e motore di innovazione e sviluppo che la formazione può svolgere negli attuali contesti produttivi e che da allora è ulteriormente cresciuta.

Lo scenario regionale dell'Emilia Romagna è molto più favorevole di quello che abbiamo descritto a livello nazionale e si colloca in ottima posizione tra le regioni Italiane su tutti gli indicatori più importanti che possono rappresentare lo sviluppo, il lavoro e la relazione tra filiere dell'istruzione, formative e del lavoro.

Sui dati 2021, nonostante l'effetto depressivo della pandemia, il tasso di attività è assestato al 78,5% per i maschi e al 66,5% per le femmine, mentre il tasso di occupazione è al 75,3 e al 61,6 rispettivamente per i due generi. Si tratta di numeri che mettono l'Emilia Romagna tra le migliori regioni europee. Il tasso di disoccupazione è al 4,0 per i maschi e al 7,1 per le femmine.

La formazione e le attività di accoglienza e accompagnamento hanno sostenuto energicamente con notevoli risorse questo profilo dello sviluppo regionale, come mostrano le relazioni annuali di attuazione RAA del POR FSE 2014/2020 in Emilia Romagna. Inoltre, i riscontri di valutazione e soddisfazione ricevuti dai partecipanti e la loro relazione successiva con il lavoro sono positivi, come emerge dall'indagine campionaria svolta dal Servizio di valutazione indipendente del POR FSE 2014–2020, in particolare sul "Tema 1 - Efficacia delle politiche formative finalizzate alla buona e qualificata occupazione dei giovani".

Il quadro positivo sul mercato del lavoro regionale e i riscontri positivi sulla formazione, in realtà creano paradossalmente le condizioni per porre domande più impegnative alla regolazione regionale e agli attori impegnati su più versanti nella formazione, dai progettisti, ai docenti, ai tutor, alle imprese più o meno 'formative' che accolgono le persone transitate attraverso il sistema della formazione.

In primo luogo, si tratta di verificare se in una situazione di mercato molto favorevole, non si verifichi il paradosso che contraddistingue molti studi sull'efficacia del welfare, ovvero se le persone che vengono sostenute non siano in realtà le più deboli e quindi proprio quelle che conseguono meno risultati rispetto alla ricerca del lavoro. In secondo luogo, esiste una gamma di interventi di formazione e svariati abbinamenti con altri servizi, questo potrebbe rendere l'effetto finale sulla ricerca del lavoro differenziato e difficile da investigare. L'insieme di attività regionali mostra un ventaglio di temi, durate e presumibilmente differenze organizzative che pongono quesiti relativi al rendimento di questi investimenti e al loro potenziale di miglioramento.

Sappiamo che sono diffuse idee e rappresentazioni contrastanti sulla formazione. Da un lato si accetta il principio secondo cui la formazione non solo è utile ma è assolutamente necessaria nello scenario attuale. Dall'altro si registrano spesso segnali in contrasto con questo principio da persone e imprese che non sempre investono adeguate energie, impegno, continuità e sistematicità nelle attività formative. Dunque il rendimento della formazione rispetto al lavoro, per giovani e per imprese, come per tutti gli investimenti, più che un dato di fatto verificato è qualcosa che deve essere generato con azioni e relazioni appropriate, su cui bisogna interrogarsi, misurarne l'effetto e accrescerne l'efficacia. In questo breve studio ci poniamo l'obiettivo di lavorare su una semplice domanda: la formazione serve all'occupazione? Intorno a questa domanda possono essere sviluppate molte altre questioni rilevanti, dove la ricerca delle risposte dipende sostanzialmente dalla possibilità di rafforzare una logica di miglioramento dell'azione pubblica e utilizzare in modo pieno e sistematico le banche dati esistenti e le informazioni: in specifico, quale formazione serve? A chi serve? Per quanto tempo serve? A cosa serve?

Prima di entrare in questa analisi è utile riepilogare le informazioni disponibili sulla platea di persone coinvolte nelle iniziative formative finanziate dal FSE in Emilia Romagna.

2 I DESTINATARI DEGLI INTERVENTI FORMATIVI DEL FSE E IL PROFILO DEI PERCORSI

I destinatari degli interventi del FSE su cui si svilupperemo l'analisi controfattuale sono coloro che hanno usufruito di azioni di sostegno nell'ambito del **Tema 2** – “Efficacia delle politiche formative e delle politiche attive del lavoro per sostenere l'inserimento e il reinserimento nel mercato del lavoro contrastando la disoccupazione di lunga durata” e, nello specifico di almeno una delle seguenti offerte:

- Formazione per l'acquisizione di qualifiche (di accesso all'area professionale, di approfondimento tecnico e legate al Sistema Regionale delle qualifiche);
- Formazione permanente;
- Tirocini (di inserimento o reinserimento al lavoro o formativi e di orientamento).

In un periodo temporale che va da Settembre 2015 a Dicembre 2018, sono state coinvolte, quindi, 30.230 persone, di cui il 57,5% sono donne e il 42,5% uomini (grafico 1). La componente femminile prevale nei percorsi formazione permanente (83,3% F vs 75,5% M), mentre i percorsi di acquisizione di qualifiche professionali sono intrapresi in prevalenza da uomini (23,3% M vs 15,7% F – grafico 2).

Grafico 1 – Destinatari per genere

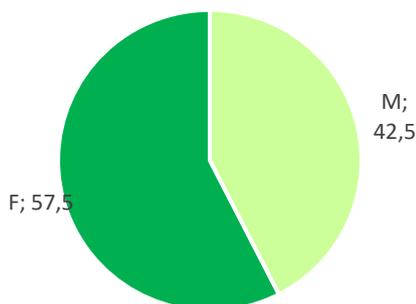
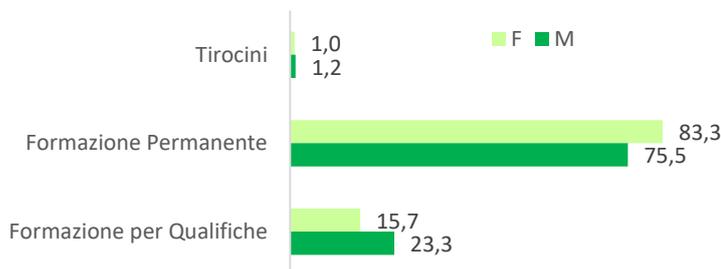


Grafico 2 – Destinatari per genere e macro tipologia FSE



Il 91% delle persone ha uno stato di disoccupazione ed un ulteriore 8,4% è in cerca di prima occupazione (tabella 3). Residuali le percentuali di occupati (0,1%), inattivi (0,5%) e studenti (solo 10 persone). All'interno dell'analisi di impatto occupazionale non si è tenuto conto delle posizioni di inattività, ovvero di coloro che pur essendo considerate nelle Forze di Lavoro, non sono classificate come occupate o in cerca di occupazione per specifiche condizioni quali l'essere ritirato/a dal lavoro, o inabile al lavoro, ecc.

Tabella 3 – Destinatari per condizione occupazionale dichiarata nel programma FSE

	N.	%
In cerca di prima occupazione	2529	8.4
Occupato (compreso chi ha un'occupazione saltuaria/atipica e chi è in CIG)	31	.1
Disoccupato alla ricerca di nuova occupazione (o iscritto alle liste di mobilità)	27520	91.0
Studente	10	.0
Inattivo diverso da studente (es. ritirato/a dal lavoro, inabile al lavoro, ecc.)	140	.5
Totale	30230	100.0

Rispetto all'età dei destinatari considerati nell'analisi, possiamo osservare una distribuzione omogenea per le varie fasce (23-24% - grafico 4), con una percentuale ridotta di over sessantenni (7,5%). La piramide dell'età si distribuisce, invece, in modo disomogeneo per genere (grafico 5): oltre il 30% degli uomini ha un'età al di sotto dei 30 anni, mentre le donne giovani sono il 17%; analogamente, anche tra i destinatari più anziani osserviamo una percentuale maggiore di uomini (9%) rispetto alle donne (6%). Per contro le donne sono più rappresentate nelle fasce di età centrali: 24% tra i 30 e i 39 anni (vs 21% M), 28% tra i 40 e i 49 anni (vs 20% M) e 25% tra i 50 e i 59 anni (vs 21% M).

Grafico 4 – Destinatari per fasce di età

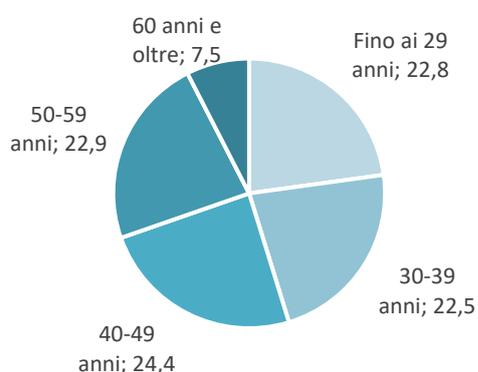
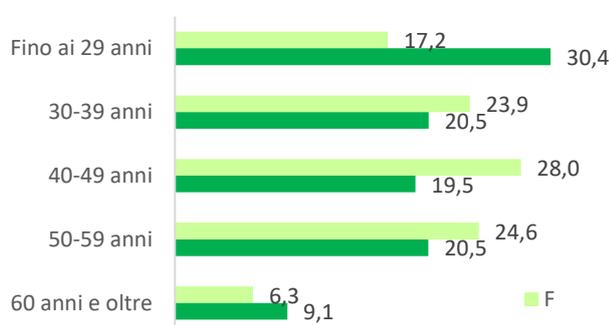


Grafico 5 – Destinatari per fasce di età e genere



Il 45% dei destinatari analizzati è nato all'interno dei confini della Regione Emilia Romagna (tabella 6); il 16% nelle regioni insulari e del sud Italia e il 7% in altre regioni del centro-nord. Complessivamente la componente di stranieri rappresenta, quindi, circa un terzo dei destinatari selezionati, con una significativa presenza di nativi africani (14%), europei di stati non appartenenti all'Unione Europea (7,4%), europei dei paesi UE (5,5%), sud e centro americani (2,8%) e asiatici (2,3%).

Tabella 6 – Destinatari per area di nascita

	N.	%
UE	1627	5,5
Europa non UE	2223	7,4
Asia	692	2,3
Africa	4180	14,0
America centro-sud	831	2,8
Oceania	5	,0
Stati Uniti d'America	11	,0
Canada	6	,0
Emilia-Romagna	13391	44,9
Nord e Centro	2092	7,0
Sud e Isole	4787	16,0
Totale	29845	100,0

Dalla scomposizione dell'area di nascita per genere (grafico 7) il dato più evidente riguarda la prevalenza maschile tra i nativi del continente africano (22% M vs 8% F) e lo sbilanciamento femminile tra le native di paesi europei, siano essi parte dell'UE (7% M vs 3% F) o extra EU (10% F vs 5% M).

Il 72% del campione analizzato ha la cittadinanza italiana. Le prime dieci cittadinanze straniere raccolgono un ulteriore 18% di destinatari (tabella 8). Le più numerose sono la cittadinanza rumena, quella nigeriana e quella marocchina.

La quasi totalità dei destinatari considerati ha residenza in Emilia Romagna (99,4% - tabella 9). Di questi il 30% nella sola area della Città metropolitana di Bologna. A seguire Reggio Emilia, Modena, Ferrara e Parma rappresentano tra il 9 e l'11% delle province di residenza del campione considerato nell'analisi.

Grafico 7 – Destinatari per area di nascita

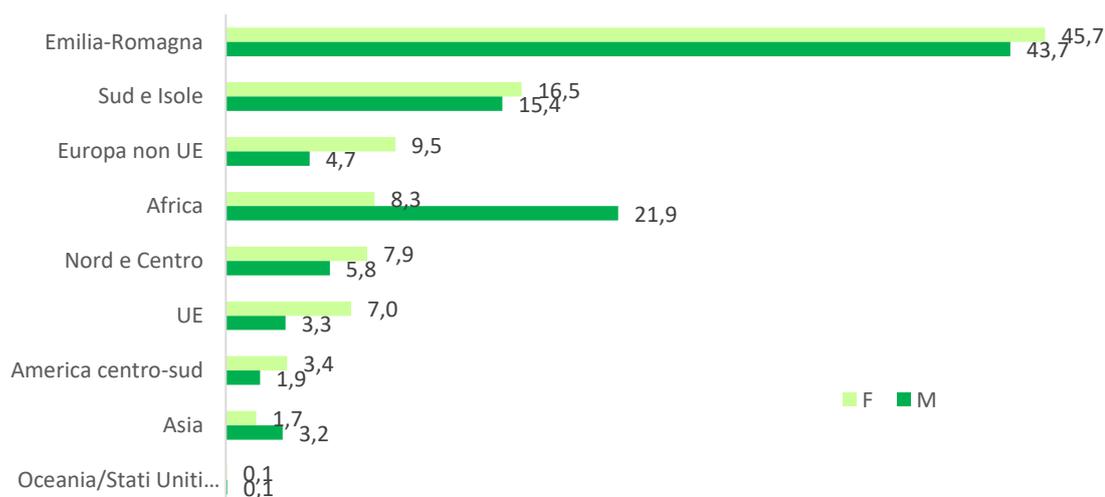


Tabella 8 – Le prime 10 cittadinanze straniere

	N.	%
Romania	911	3,6
Nigeria	800	3,2
Marocco	737	2,9
Albania	444	1,8
Senegal	359	1,4
Ucraina	332	1,3
Moldova	310	1,2
Tunisia	234	,9
Gambia	230	,9
Costa d'Avorio	202	,8
Totale	4559	18,2

Tabella 9 – Provincia di residenza

	N.	%
PC	1657	5,5
PR	2763	9,2
RE	3125	10,4
MO	3265	10,9
BO	9169	30,5
FE	3086	10,3
RA	2604	8,7
FC	2524	8,4
RN	1728	5,7
Altre province	169	0,6
Totale	30090	100

Il 40% del campione di destinatari considerato ha basse credenziali formative, avendo conseguito al massimo la licenza media o l'avviamento professionale (tabella 10), in questo gruppo si osserva una decisa prevalenza maschile. Le donne, come ben noto in letteratura, più frequentemente ottengono diplomi quinquennali e titoli accademici.

Tabella 10 – Titolo di studio

	F		M		Totale	
	N.	%	N.	%	N.	%
Nessun titolo	1178	6,8	1527	12,0	2705	9,0
Licenza elementare/attestato di valutazione finale	447	2,6	691	5,4	1138	3,8
Licenza media /avviamento professionale	4359	25,1	3891	30,5	8250	27,4
Titolo di istruzione secondaria di II grado che non permette l'accesso all'università	1482	8,5	1118	8,8	2600	8,6
Diploma di istruzione secondaria di II grado che permette l'accesso all'università	6459	37,3	3959	31,0	10418	34,6
Qualifica professionale regionale post-diploma	205	1,2	147	1,2	352	1,2
Diploma di tecnico superiore (ITS)	71	,4	63	,5	134	,4
Laurea di I livello (triennale)	1339	7,7	523	4,1	1862	6,2
Laurea magistrale/specialistica di II livello e vecchio ordinamento	1754	10,1	816	6,4	2570	8,5
Dottore di ricerca	43	,2	18	,1	61	,2
Totale	17337	100,0	12753	100,0	30090	100,0

La mappa 11 sintetizza i percorsi e le esperienze intraprese dalle persone selezionate all'interno delle proposte e dei servizi messi a disposizione attraverso il FSE (offerta formativa, orientativa e di accompagnamento).

Si tratta di percorsi spesso articolati (tabella 12), che per i due terzi delle persone, infatti, transitano attraverso corsi di formazione, dall'alfabetizzazione disciplinare all'acquisizione di qualifiche professionali, colloqui di orientamento, realizzazione di tirocini e processi di accompagnamento al lavoro.

Tabella 12 – Numero di attività intraprese dai destinatari

N. Attività	N.	%
1	9991	33.0
2	11002	36.4
3	5853	19.4
4	2043	6.8
5 e oltre	749	4.4
Totale	30230	100.0

I tre cardini sono appunto rappresentati dalla Formazione Permanente, dalla Formazione per l'acquisizione di Qualifiche e dai Tirocini.

L'offerta di **Formazione Permanente**, che ha coinvolto 25.164 persone, si caratterizza in Emilia Romagna per la proposta di corsi leggeri, di breve durata. Il 36% dei destinatari ha, infatti, usufruito di corsi di un massimo di 3 ore, il 16% tra le 3 e le 10 ore, il 21% tra 10 e 20 ore, il 23% tra le 20 e le 60 ore e solo il 4% oltre le 60 ore.

La **Formazione per l'acquisizione di qualifiche** ha interessato 5685 persone e si distingue in tre percorsi che sottendono ad un impegno formativo differenziato (tabella 13):

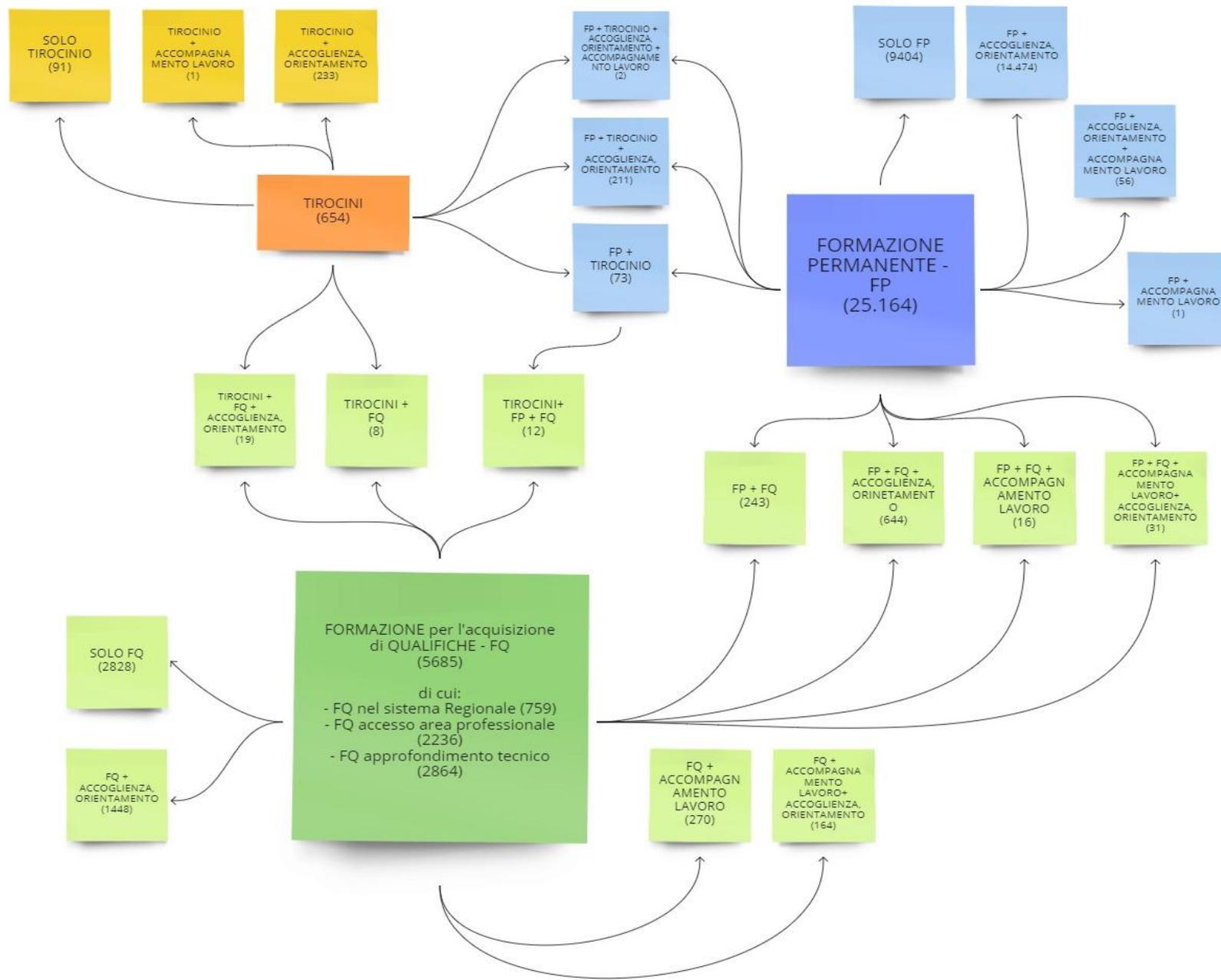
- La Formazione per l'acquisizione di qualifiche di approfondimento tecnico – specializzazione ha previsto un tempo di formazione massimo di 734 ore (corrispondente a 92 giornate a tempo pieno). Il 40% delle persone ha affrontato percorsi formativi al di sotto delle 160 ore, il 40% tra le 160 e le 204 ore e un ulteriore 20% oltre le 204 ore. Ad integrazione del percorso formativo, l'81% dei destinatari ha realizzato uno stage aziendale della durata media di un centinaio di ore (media 125 ore) e il 6,2% un project work che ha impegnato mediamente quasi 18 ore.
- La Formazione per l'acquisizione di qualifiche di accesso all'area professionale è quella a maggiore intensità di impegno, dal momento che il 40% delle persone ha affrontato percorsi con più di 280 ore d'aula, corredato nell'82% dei casi da uno stage della durata media di 179 ore e per un 6% circa di partecipanti dalla richiesta di definire un project work, con un impegno aggiuntivo medio di quasi 20 ore.
- La Formazione progettata con riferimento al Sistema regionale delle qualifiche ha previsto una presenza media in aula di 221 ore, con il 40% di destinatari inseriti in percorsi superiori alle 258 ore. La quota di persone per le quali è stato attivato un tirocinio supera l'86%, con un impegno medio di 166 ore.

Tabella 13 – Ore spese per le differenti tipologie formative per quintili di distribuzione

	Formazione progettata con riferimento al Sistema regionale delle qualifiche			Formazione per l'acquisizione di qualifiche di accesso all'area professionale			Formazione per l'acquisizione di qualifiche di approfondimento tecnico – specializzazione			
	Formazione	Stage	Project work	Formazione	Stage	Project work	Formazione	Stage	Project work	
N	759	655	5	2236	1833	107	2864	2337	144	
% su destinatari		86,3%	0,8%		82,0%	5,8%		81,6%	6,2%	
	Ore			Ore			Ore			
Media	220,6	165,6	16,0	223,7	170,6	19,5	164,2	125,1	17,8	
Deviazione std.	97,0	57,2	11,0	108,3	60,1	5,5	71,8	40,2	8,0	
Minimo	0,0	20,0	8,0	0,0	2,0	5,0	0,0	4,0	2,0	
Massimo	649,1	437,0	28,0	547,0	388,0	28,0	734,5	400,0	50,0	
Percentili	20	143,8	116,8	8,0	125,4	120,0	16,0	124,3	100,0	8,0
	40	180,0	136,0	8,0	180,0	152,0	16,0	159,5	116,0	16,0
	60	258,0	196,0	20,0	280,0	208,0	20,0	174,9	120,0	20,0
	80	312,0	229,0	28,0	331,0	236,0	28,0	204,0	171,0	25,0

Per 654 destinatari il progetto di **Tirocinio** è stato prevalente, nel 36% dei casi c'è stato anche un accompagnamento orientativo e per un ulteriore 32% della formazione permanente

Mapa 11 - Servizi e attività intraprese (tra parentesi il numero di persone)



Nota tecnica:

- 1) *L'area di nascita è una riaggregazione del **comune di nascita**, che nell'archivio FSE (casi) è espressa in modo codificato secondo la classificazione in uso presso l'Agenzia delle Entrate (come espressa nel Codice Fiscale). Nell'archivio delle Comunicazioni Obbligatorie (controlli) il Comune di nascita è rappresentato da una stringa di denominazione del Comune stesso. Per rendere confrontabili i due archivi si è reso, quindi, necessario un lavoro di ricodifica specifico, preliminare all'aggregazione per aree.
Inoltre, per procedere con l'appaiamento casi-controlli (Propensity Score Matching) i Comuni e le Province sono stati riportati alla classificazione Istat.*
- 2) *Analogamente, l'area di residenza e la **provincia di residenza** sono state costruite come riaggregazioni dei comuni di residenza. Nell'archivio FSE (casi) l'informazione di partenza era disponibile codificata secondo la classificazione dell'Istat, mentre nelle COB il comune di residenza era espresso con una stringa di denominazione, che è stata ricondotta alla stessa tipologia classificatoria utilizzata per i casi.*
- 3) *Anche per la **cittadinanza** si è reso necessario procedere con l'allineamento dei due differenti sistemi di classificazione utilizzati negli archivi di casi e controlli. Le denominazioni sono, quindi, state riportate alla classificazione Istat.*
- 4) *Il **titolo di studio** era presente per quasi tutti i destinatari del FSE. Nell'archivio dei controlli (COB) il 17,6% dei casi estratti ha titolo di studio mancante. Inoltre, la variabile resa disponibile nello scarico dei dati delle COB è una stringa descrittiva del titolo di studio e spesso dello specifico indirizzo di studi, senza tuttavia consentire una chiara corrispondenza con il livello di titolo (qualifica professionale, diploma, laurea I o II livello). Per procedere con l'appaiamento casi-controlli (Propensity Score Matching) è stato, quindi necessario ricondurre manualmente le descrizioni presenti nell'archivio dei controlli alla classificazione in uso per i casi FSE, selezionando i soli titoli la cui interpretazione è risultata inequivocabile. Nel complesso l'operazione ha ridotto il gruppo dei controlli a 231.410 osservazioni (persone) pari al 51,6% del bacino estratto.*

La valutazione d'impatto condotta con approccio controfattuale permette di rispondere a una domanda cruciale per le politiche, nel nostro caso le politiche attive del lavoro e la formazione: gli interventi formativi svolti nell'ambito del FSE hanno inciso sul livello di occupazione delle persone e quindi hanno fornito una risposta alla domanda di personale delle imprese? In particolare, in un contesto dinamico come quello della regione Emilia Romagna, i processi di reclutamento e di gestione del turn-over da parte delle imprese sono piuttosto fluidi, quindi le persone trovano lavoro sul mercato indipendentemente dal fatto che abbiano partecipato o meno ad attività formative. Per questo motivo e per lo storico orientamento della missione del FSE, le politiche attive sono di fatto un'offerta rivolta alle persone più in difficoltà, che faticano a trovare lavoro facendo conto esclusivamente sulle proprie risorse. In qualsiasi caso, se l'investimento formativo ha un rendimento, le persone formate dovrebbero essere facilitate nel reperimento di un lavoro soddisfacente. Sappiamo che il match sul mercato del lavoro è soggetto all'influenza di numerosi fattori che modificano aspettative e decisioni delle due parti, ma semplificando possiamo ritenere che l'investimento formativo, se funziona, dovrebbe aumentare la probabilità del match.

Dunque il problema strategico della valutazione, per rispondere alla domanda sull'efficacia della formazione, consiste nella possibilità di isolare e misurare l'effetto dell'intervento sul match rispetto a ciò che sarebbe avvenuto in assenza dell'intervento. Per fare questo in modo corretto sarebbero necessari tre tipi di informazioni.

La prima classe di informazioni si riferisce alle persone coinvolte nell'esperimento. Il loro livello di competenze, motivazioni e relazioni costituisce una dotazione essenziale per la performance sul mercato del lavoro, indipendentemente dalla successiva esposizione a interventi formativi, che possono incidere e migliorare il profilo di partenza.

La seconda classe di informazioni si riferisce al tipo di effetto che ci si attende dalla formazione, che dipende dai suoi obiettivi, durata, metodi ed efficacia interna. Più questo effetto è ampio sui diversi aspetti del profilo (conoscenze, competenze, relazioni, motivazione, apprendimento di pratiche, ecc.) più aumenta la probabilità di rilevare un effetto aggiuntivo sul lavoro.

La terza classe di informazioni riguarda il match e quindi il reperimento o meno di un lavoro da parte della persona esposta alla formazione. Anche questa informazione è complessa. Se il lavoro è stato reperito, la valutazione dovrebbe considerare la sua stabilità, il suo inquadramento e livello, il suo allineamento per settore, competenze e contenuti con le attitudini e gli orientamenti del candidato e con l'indirizzo della formazione svolta. Se il lavoro non è stato reperito in tempi rapidi, bisogna considerare un effetto messo in evidenza dalle indagini sulle carriere lavorative in Italia, che ha mostrato come a volte aspettare l'occasione giusta, il lavoro allineato alle proprie competenze e aspirazioni, soprattutto se si tratta dei primi ingressi sul mercato da parte dei giovani, anche a costo di aspettare per più tempo, può essere un indizio di forza personale e sul mercato, più che di debolezza. Talvolta le persone più in difficoltà accettano qualsiasi lavoro, anche se breve, precario e fuori dai propri interessi, mentre quelle più forti e preparate possono essere più selettive, anche se l'eventuale attesa dipende molto dal periodo e dallo specifico segmento di mercato che si sta osservando.

La domanda sugli effetti generati dalla formazione, quindi, è ampia, perché l'effetto dell'intervento è multidimensionale e si colloca in un percorso che vede il soggetto già presente sul mercato del lavoro prima e dopo l'intervento formativo.

L'analisi controfattuale può collocarsi in due scenari distinti, che si differenziano nella loro possibilità di gestire il problema chiave, che riguarda come vedremo la possibilità di controllare l'eterogeneità non osservata e/o non osservabile, ovvero quelle caratteristiche delle persone importanti nel lavoro ma non presenti nei dati disponibili per l'analisi.

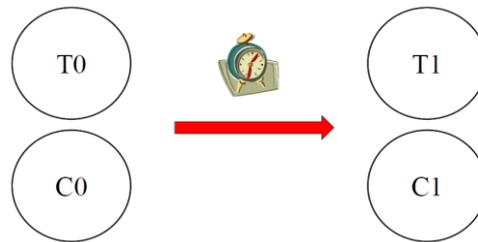
Il primo scenario è quello della progettazione dell'analisi controfattuale disponendo di informazioni su tutti i principali fattori che possono influenzare il risultato, in particolare l'effetto della formazione sul lavoro. In questo caso a partire dalle ipotesi teoriche e dalle dimensioni da misurare, che abbiamo brevemente riassunto sopra, si architetta l'analisi e si giunge al risultato, che deve poi essere calato nel contesto e interpretato alla luce delle connessioni causali ricostruibili nei match in Emilia Romagna.

Il secondo scenario, in cui si colloca anche questo lavoro, peraltro di gran lunga il più frequente nell'analisi del mercato del lavoro italiano, consiste nella realizzazione del percorso inverso, dove si mette a fuoco l'interesse per la valutazione d'impatto e si raccolgono i dati disponibili, che spesso consentono di misurare solo parte delle dimensioni in gioco. Questo percorso costringe a valutare con attenzione i limiti dell'analisi, per non attribuire ai risultati empirici, sempre interessanti, significati inappropriati rispetto alle premesse e agli obiettivi di valutazione.

In entrambi gli scenari i risultati del metodo quantitativo-controfattuale dovrebbero essere discussi in un contesto di valutazione in grado di attribuire significato agli effetti di breve e di medio-lungo periodo e dovrebbero servire a favorire da parte degli operatori attivi nell'ambito della formazione il miglioramento delle proprie performance. Ciò avviene riflettendo sulle catene causali che spiegano gli effetti quantitativi emersi, con metodologie di carattere qualitativo-realista (interviste, analisi dei documenti e del comportamento degli attori, ecc.), per comprendere i meccanismi sociali, culturali e organizzativi che guidano il comportamento delle persone e delle imprese in quello specifico contesto. Su queste trasformazioni indotte dall'intervento nella dotazione e nel comportamento delle persone si basano anche le aspettative di permanenza degli effetti nel tempo. Inoltre, la ricognizione qualitativa nello scenario potrebbe esplorare la presenza dei diversi effetti voluti dalla politica, ma anche di effetti contro-intuitivi ed effetti inattesi o non voluti relativi all'impatto, frequenti quando si utilizzano leve complesse come la formazione.

Entrando nel merito della metodologia controfattuale, la stima degli effetti non può essere condotta confrontando il gruppo delle persone trattate (T), beneficiarie dirette dell'intervento, al tempo 0, prima dell'intervento, e al tempo 1, dopo l'intervento. In questo caso il delta includerebbe anche effetti, positivi o negativi, che si sarebbero sviluppati indipendentemente dall'intervento stesso (figura 14). Analogamente, la stima non può essere basata confrontando le persone trattate dopo l'intervento, al tempo 1 (T1), con un gruppo di persone di controllo (C1), anch'esse selezionate dopo l'intervento. In questo caso sarebbe impossibile attribuire un eventuale delta riscontrato all'intervento, perché non si potrebbe essere sicuri che non esistano altre differenze tra le persone, soprattutto differenze non osservabili facilmente, che sono alla base delle differenze rilevate.

Figura 14



Il met, quindi, richiede che i due gruppi, persone trattate e persone di controllo, siano identificate prima dell'intervento e che siano attribuite per sorteggio casuale ai due gruppi, in modo da azzerare l'effetto di differenze non osservabili, ma rilevanti del modificare, in un senso o nell'altro, l'effetto della politica.

Come è noto, tuttavia, in molte politiche sociali come nel caso della formazione FSE, non è possibile la pratica alla lettera di questo metodo, perché non sarebbe giustificabile negare a una parte dei richiedenti (persone con necessità di qualificazione) l'intervento al solo scopo di condurre la valutazione in modo rigoroso. La metodologia ha dunque sviluppato la teoria dei quasi-esperimenti e degli esperimenti naturali, che consente di avvicinarsi alle condizioni sperimentali, senza creare condizioni difficilmente sostenibili nella relazione con i destinatari.

Una tecnica sperimentale può essere applicata quando un motivo indipendente da variabili che possono influenzare l'effetto divide i due gruppi di trattati e di controllo. Possono essere motivi di carattere geografico, temporale o amministrativo, che introducono una soglia o una cesura che priva una parte dei richiedenti dell'intervento. In questo caso essi possono rappresentare una base per la selezione di un gruppo di controllo, in quanto non si sono auto-selezionati ma sono stati esclusi per circostanze che potrebbero essere casuali rispetto alla teoria dell'intervento.

Quando un approccio sperimentale non è possibile occorre applicare tecniche di appaiamento iniziale dei due gruppi, trattati e di controllo, al tempo zero, in modo da simulare il sorteggio iniziale, anche se svolto successivamente. In questo caso con la tecnica del *Propensity Score Matching* ci si assicura che il gruppo dei controlli corrisponda il più possibile a quello dei casi, sulla base di un set di informazioni precedenti allo sviluppo dell'intervento. Sia nel caso dell'esperimento naturale, sia nel caso del quasi esperimento, la tecnica dell'appaiamento precedente è necessaria, perché consente di tenere sotto controllo l'appaiamento al tempo T0. L'appropriatezza e l'accuratezza delle variabili con cui si realizza l'appaiamento iniziale ha quindi importanza strategica nella costruzione del quasi esperimento, in particolare nel nostro caso presenza e qualità delle tre classi di informazioni elencate sopra, che consentirebbero di individuare dei sosia da confrontare alle persone formate che siano in possesso di analoghe caratteristiche prima dell'intervento. Trattati e sosia, appaiati prima, verrebbero poi seguiti nel loro percorso sul mercato del lavoro in modo da confrontare come i due gruppi si sono comportati sul mercato del lavoro.

Questo ha consentito di valutare l'efficacia della formazione con l'analisi controfattuale con la tecnica del quasi esperimento. Essa è stata progettata a posteriori (scenario 2 descritto sopra) e si è potuta valere dei seguenti archivi relativi alla formazione e al lavoro in Emilia Romagna.

- Registro delle attività formative e di accoglienza svolte nell'ambito del FSE. Universo sull'Emilia Romagna.
- È stato messo a disposizione per la simulazione l'archivio delle attività svolte nel periodo dal 1.9.2015 al 31.12.2018. Queste attività svolte (n. 310.861 eventi) sono catalogate per tipo di attività e data di svolgimento. Da questo archivio è stato tratto il database delle persone trattate, che include un totale di 30.230 persone, che hanno svolto da 1 a 23 attività, di cui almeno una tra formazione per l'acquisizione di qualifiche, formazione permanente o tirocini (come spiegato nel cap. 2). L'archivio contiene, per ciascuna di queste persone, le seguenti informazioni: sesso, titolo di studio, luogo di nascita, anno di nascita, cittadinanza, comune di domicilio e di residenza, condizione occupazionale, tipo di evento, data di inizio e data di fine dell'evento e, nel caso della formazione, anche il tempo effettivamente dedicato, con la specifica delle ore di tirocinio, formazione a distanza e realizzazione di project work.

L'analisi tecnica di questo archivio non ha messo in evidenza particolari criticità. Si conosce la tipologia di attività svolta, ma non sono presenti valutazioni sul completamento o sull'esito della formazione, dal punto di vista del livello di apprendimento. La presenza di chiavi individuali criptate ha consentito di associare a ciascuna di queste persone che hanno partecipato alla formazione gli eventi dell'archivio seguente relativi alle COB che li riguardavano.

- COB. Archivio delle Comunicazioni Obbligatorie di inizio di attività lavorativa. Universo sull'Emilia Romagna.

È stato messo a disposizione un archivio delle comunicazioni obbligatorie inoltrate delle imprese all'avvio di qualsiasi rapporto di lavoro e cambiamento nella relazione di lavoro secondo la normativa vigente. Si tratta delle comunicazioni registrate per tutte le persone che avevano una dichiarazione di immediata disponibilità (DID) nel periodo dal 1.1.2015 al 31.12.2019. L'archivio contiene, per ciascuna di queste persone, le seguenti informazioni: sesso, titolo di studio, luogo di nascita, anno di nascita, cittadinanza, comune di residenza, data DID, data COB, tipo COB, datore di lavoro e settore dell'impresa. L'analisi tecnica dell'archivio ha messo in evidenza la mancanza nell'archivio messo a disposizione di un certo numero di informazioni che esistono nel tracciato delle comunicazioni obbligatorie: 1) l'identificativo delle COB non sempre collega in modo univoco le diverse comunicazioni che si riferiscono allo stesso rapporto di lavoro (avviamento, proroga, passaggio, cessazione). Questa mancanza rende impossibile gestire i casi di molteplicità di comunicazioni come in presenza di part time differenti o di incongruenza tra date e comunicazioni. 2) il tipo di contratto cui si riferisce l'avviamento (tempo indeterminato, determinato, ecc. e la relativa scadenza programmata. Questa carenza ha reso impossibile utilizzare le informazioni sulle proroghe e sui passaggi di contratto, in quanto mancava la base su cui dovevano essere applicati e verificati. 3) la durata in giorni del contratto a tempo determinato, che ha reso impossibile costruire una ponderazione dei differenti avviamenti, distinguendo avviamenti ad esempio di un giorno, un mese o un anno. 4) la qualifica di avviamento, e quindi la catalogazione del rapporto riguardo alla qualifica legale. 5) una parte delle segnalazioni di cessazione, in quanto sono presenti avviamenti non intervallati da una cessazione. Questo insieme di carenze ha impedito di costruire indicatori relativi alla qualità e alla durata dell'attività svolta, sia prima della formazione, sia dopo la formazione. Ha inoltre reso più incerta la costruzione della cronologia dei percorsi individuali.

La costruzione della stima richiede due azioni preliminari per predisporre l'archivio delle persone che hanno seguito la formazione e dei sosia su cui condurre le analisi.

La prima consiste nel trattare i dati in modo da assegnare a tutte le informazioni il corretto ruolo nell'analisi, valorizzando quanto disponibile e registrando ciò che eventualmente manca, che avrà un ruolo fondamentale nella interpretazione dei risultati.

La seconda consiste nel disegno dell'esperimento, che richiede di prefigurare la simulazione creando corrette condizioni teoriche-concettuali. In particolare, occorre che trattati e sosia abbiano potenzialmente la stessa possibilità di successo nella ricerca del lavoro, in modo che il successo osservato possa essere interpretato come un effetto della formazione.

Classificazione delle informazioni ricavabili dagli archivi:

- 1) Prima classe di informazioni: profilo delle persone coinvolte nell'esperimento. Sono variabili di controllo e servono ad abbinare i sosia ai trattati.
Disponiamo sulle persone di sesso, anno di nascita, titolo di studio, luogo di nascita, cittadinanza e residenza. Si tratta di variabili molto importanti per definire un profilo lavorativo. Mancano tuttavia le variabili che a parità di profilo influenzano la performance sul mercato del lavoro come livello di competenze, motivazioni e tipo di relazioni personali e familiari. Queste vengono dedotte generalmente dal tipo di studi svolti, dalle votazioni conseguite e dalle esperienze e ruoli lavorativi occupati, che possono migliorare o peggiorare il profilo. Unica informazione di cui disponiamo in questo caso riguarda la presenza di comunicazioni di assunzione o di cessazioni nel periodo che intercorre tra l'avvio dell'archivio e la formazione, che indica una qualche esperienza di lavoro precedente, la cui natura tuttavia può essere molto differenziata.
- 2) Seconda classe di informazioni: tipo di effetto che ci si attende dalla formazione. Si tratta della variabile indipendente nell'esperimento. Disponiamo della informazione se ha partecipato alla formazione, a quale tipologia di corso e per quale durata. Mancano informazioni sull'esito della formazione.
- 3) Terza classe di informazioni: riguarda il fatto di aver lavorato dopo il trattamento. Si tratta della variabile dipendente, con cui si può misurare l'impatto del trattamento. Come abbiamo già osservato, mancano informazioni sulla qualità e sulla stabilità di questo lavoro. Possiamo utilizzare, oltre al fatto di aver lavorato o meno, anche il numero di occasioni e il settore di impiego, anche se in assenza di informazioni sul tipo e durata del contratto, non sono molto indicative circa la qualità del lavoro.

Nell'ambito di questo lavoro di classificazione sono anche state ridefinite, revisionate e riclassificate le informazioni e sono state costruite nuove variabili specifiche, finalizzate all'analisi. L'archivio è stato inoltre ripulito dai casi che avevano un'età fuori dall'età di lavoro ISTAT che fino al 2019 era di 15-65 anni ed è stata poi portata a 67 anni. Sono quindi state considerate le persone con anni di nascita dal 1950 al 2000.

6 DISEGNO DELL'ESPERIMENTO E OPERAZIONI PER L'ALLINEAMENTO DEGLI ARCHIVI

Riguardo al disegno dell'esperimento la strategia di analisi controfattuale è stata basata sul seguente percorso logico:

- Partire dall'archivio delle attività FSE:
 - l'archivio contiene 310.860 attività (267.555 persone),
 - trasformarlo in un archivio di persone trattate, mantenendo e riclassificando tutte le informazioni relative ai corsi. L'archivio contiene 30.230 persone, ovvero chi ha almeno un'attività di formazione permanente, per l'acquisizione di qualifiche o tirocinio
- Agganciare a ciascuna persona tutte le informazioni presenti nell'archivio delle comunicazioni obbligatorie, attraverso identificativi criptati, in modo da disporre per la stessa persona sia delle comunicazioni di lavoro precedenti sia di quelle successive alla formazione.
- Identificare la data di fine formazione come il momento in cui il soggetto si è presentato sul mercato, quindi il momento di apertura della finestra temporale dell'esperimento.
- Identificare nella data del 15.03.2020, data fino a cui erano disponibili le comunicazioni obbligatorie nei nostri dati, come data di fine dell'esperimento e quindi di chiusura della finestra temporale.

Questo archivio, con le necessarie ripuliture ed eliminazione di casi con dati mancanti è costituito da 24.376 persone che hanno una comunicazione precedente alla data di fine formazione nell'archivio delle comunicazioni obbligatorie e da 5.021 persone che non hanno comunicazioni precedenti.

Tabella 15 – Presenza di comunicazioni obbligatorie prima della realizzazione di attività FSE

COB precedente alla fine attività FSE

(cob_precedente)

	Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
No	5021	17,1	17,1	17,1
Sì	24376	82,9	82,9	100,0
Totale	29397	100,0	100,0	

Ai fini della progettazione dell'analisi controfattuale questi due gruppi devono essere trattati in modo differente.

Infatti, il primo gruppo può essere affiancato a sosia, anch'essi già presenti per lavoro precedente nell'archivio delle comunicazioni obbligatorie, definiti attraverso il PSM, e i due gruppi trattati FSE e non trattati possono essere confrontati rispetto alla probabilità di aver lavorato dopo.

Viceversa, il secondo gruppo non può avere dei sosia adeguati al confronto, perché coloro che sono stati trattati FSE e non hanno mai lavorato prima della data di fine formazione, ma non hanno lavorato neanche dopo, non hanno i loro sosia nell'archivio delle COB proprio perché non hanno mai lavorato. In altre parole, le persone in cerca di prima occupazione (ovvero che non hanno mai lavorato) non sono presenti nell'archivio delle COB e quindi non possono essere arruolate nel campione di sosia che dovrebbe avere questa caratteristica.

Si tratta di un problema che riguarda la struttura degli archivi che sono stati messi a disposizione. Dato che l'universo delle comunicazioni obbligatorie estratto contiene per definizione solo persone che hanno lavorato anche solo per un giorno e dato che le DID (date di dichiarazione di disponibilità al lavoro) sono state estratte in associazione a coloro che avevano una COB (anche se potevano essere estratte in modo indipendente), si è posto il problema della mancanza tra i sosia delle persone in cerca di prima occupazione e che successivamente alla DID non hanno svolto alcuna esperienza di lavoro, che possono essere invece presenti tra i partecipanti alle attività FSE, in quanto l'archivio FSE è stato strutturato a partire dalle persone che hanno svolto le attività di formazione e che possono essere assenti dall'archivio delle COB. Il confronto senza tenere conto di questa carenza strutturale porterebbe a sottostimare l'eventuale effetto positivo delle attività FSE nella ricerca del lavoro.

Alla luce di questa distinzione, per costruire l'archivio per l'esperimento sono stati cercati i sosia del primo gruppo (trattati FSE che avevano avuto almeno una occasione di lavoro prima della fine della formazione) tra coloro che avevano le caratteristiche analoghe.

Le informazioni disponibili in entrambi gli archivi utilizzabili per appaiare trattati e non trattati sono: sesso, anno di nascita, titolo di studio, luogo di nascita, data in cui si è aperta la finestra temporale di ricerca del lavoro.

Ciascuna di queste variabili è stata trattata e ha richiesto un lavoro di riclassificazione per renderla adeguata alla ricerca dei sosia, in particolare:

- sesso: F=1
- anno di nascita: variabile tagliata su età di lavoro (15-65)
- titolo di studio: anni di studio ricavati dal titolo acquisito su 8 modalità
- luogo di nascita: variabile con 9 modalità interpretabile come indice di distanza geografica e culturale dall'Emilia Romagna
- cittadinanza: 0 straniera, 2 italiana
- data fine formazione per i casi e data DID per i controlli: riclassificata in numero giorni (o mesi) trascorsi dall'apertura della finestra temporale per l'inizio della ricerca del lavoro, da zero a 1257. Questa variabile dovrebbe svolgere una funzione particolare nell'appaiamento, perché rende il più simili possibile trattati e sosia rispetto alla durata della finestra temporale per la ricerca del lavoro. In pratica la finestra segue la data della fine formazione per i casi appaiando ad essa la data DID dei controlli. Essa quindi è una finestra individuale e inizia dal 12 ottobre 2015 per i primi casi trattati che si presentano sul mercato e finisce il 21 dicembre 2018 quando gli ultimi casi trattati finiscono la formazione. Per i sosia è stata usata la stessa finestra all'interno della quale sono state cercate dalla procedura automatica di appaiamento le date più vicine possibile. Le occasioni di lavoro saranno poi cercate nell'archivio delle COB successivamente a tale data, tenendo conto che disponiamo delle estrazioni per tutti fino al 15.03.2020, anche se la durata delle estrazioni è differente per casi e controlli ed ha dovuto essere allineata a posteriori.

Tabella 16 - Sesso dei trattati FSE dell'universo finale

Sesso (sesso_PSM)				
	Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
0 M	12345	42,0	42,0	42,0
1 F	17052	58,0	58,0	100,0
Totale	29397	100,0	100,0	

Tabella 17 - Data di nascita dei trattati FSE dell'universo finale

	(anno_nascita_PSM)			Percentuale
	Frequenza	Percentuale	Percentuale valida	
1950	19	,1	,1	,1
1951	29	,1	,1	,2
1952	82	,3	,3	,4
1953	127	,4	,4	,9
1954	193	,7	,7	1,5
1955	235	,8	,8	2,3
1956	275	,9	,9	3,3
1957	332	1,1	1,1	4,4
1958	429	1,5	1,5	5,9
1959	491	1,7	1,7	7,5
1960	509	1,7	1,7	9,3
1961	626	2,1	2,1	11,4
1962	594	2,0	2,0	13,4
1963	658	2,2	2,2	15,6
1964	704	2,4	2,4	18,0
1965	694	2,4	2,4	20,4
1966	728	2,5	2,5	22,9
1967	744	2,5	2,5	25,4
1968	764	2,6	2,6	28,0
1969	786	2,7	2,7	30,7
1970	771	2,6	2,6	33,3
1971	770	2,6	2,6	35,9
1972	718	2,4	2,4	38,4
1973	764	2,6	2,6	41,0
1974	731	2,5	2,5	43,5
1975	748	2,5	2,5	46,0
1976	737	2,5	2,5	48,5
1977	721	2,5	2,5	51,0
1978	648	2,2	2,2	53,2
1979	617	2,1	2,1	55,3
1980	637	2,2	2,2	57,4
1981	658	2,2	2,2	59,7
1982	645	2,2	2,2	61,9
1983	658	2,2	2,2	64,1
1984	654	2,2	2,2	66,3
1985	706	2,4	2,4	68,7
1986	630	2,1	2,1	70,9
1987	679	2,3	2,3	73,2
1988	665	2,3	2,3	75,4
1989	675	2,3	2,3	77,7
1990	718	2,4	2,4	80,2
1991	701	2,4	2,4	82,6
1992	729	2,5	2,5	85,0
1993	695	2,4	2,4	87,4
1994	732	2,5	2,5	89,9
1995	774	2,6	2,6	92,5
1996	853	2,9	2,9	95,4
1997	698	2,4	2,4	97,8
1998	426	1,4	1,4	99,3
1999	176	,6	,6	99,9
2000	44	,1	,1	100,0
Totale	29397	100,0	100,0	

Tabella 18 - Titolo di studio dei trattati FSE dell'universo finale

(titolo_studio_PSM)				
	Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
0 Nessun titolo	2826	9,6	9,6	9,6
1 Licenza elementare	1066	3,6	3,6	13,2
2 Licenza media	8066	27,4	27,4	40,7
3 Titolo di istruzione secondaria II grado (no accesso università)	2671	9,1	9,1	49,8
4 Diploma di istruzione secondaria II grado (per accesso università)	10268	34,9	34,9	84,7
8 Laurea di I e II livello	4441	15,1	15,1	99,8
9 Dottore di ricerca	59	,2	,2	100,0
Totale	29397	100,0	100,0	

Tabella 19 - Provenienza dei trattati FSE dell'universo finale

(fasce_prov_nascita_PSM)				
	Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
1 Emilia-Romagna	13162	44,8	44,8	44,8
2 Nord e Centro	2061	7,0	7,0	51,8
3 Sud e Isole	4710	16,0	16,0	67,8
10 UE	1606	5,5	5,5	73,3
15 Europa non UE	2188	7,4	7,4	80,7
20 USA-Can-Australia	22	,1	,1	80,8
25 America centro sud	815	2,8	2,8	83,6
30 Asia	682	2,3	2,3	85,9
35 Africa	4151	14,1	14,1	100,0
Totale	29397	100,0	100,0	

Tabella 20 - Cittadinanza dei trattati FSE dell'universo finale

(newcittadinanza_PSM cittadinanza_PSM_sintesi)				
	Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
0 Stranieri	6740	22,9	22,9	22,9
1 Non rilevata	4996	17,0	17,0	39,9
2 Italiani	17661	60,1	60,1	100,0
Totale	29397	100,0	100,0	

Tabella 21 - Data fine formazione dei trattati FSE dell'universo finale

Periodo	Frequenza	Percentuale	Percentuale cumulativa	Periodo	Frequenza	Percentuale	Percentuale cumulativa	Periodo	Frequenza	Percentuale	Percentuale cumulativa	Periodo	Frequenza	Percentuale	Percentuale cumulativa
102 12-OCT-2015	1	0,0	0,0	301 19-JAN-2016	25	0,1	2,4	400 10-AUG-2016	5	0,0	2,6	602 03-MAR-2017	1	0,0	16,8
103 13-OCT-2015	2	0,0	0,0	202 20-JAN-2016	11	0,0	1,2	418 29-AUG-2016	3	0,0	3,6	607 08-MAR-2017	1	0,0	16,8
104 14-OCT-2015	1	0,0	0,0	203 21-JAN-2016	5	0,0	1,3	419 30-AUG-2016	1	0,0	3,6	615 16-MAR-2017	4	0,0	16,9
106 16-OCT-2015	2	0,0	0,0	204 22-JAN-2016	14	0,0	1,3	420 31-AUG-2016	1	0,0	3,6	619 21-MAR-2017	4	0,0	16,9
109 19-OCT-2015	2	0,0	0,0	207 25-MAR-2016	2	0,0	1,3	426 06-SEP-2016	2	0,0	3,6	620 22-MAR-2017	444	1,5	18,4
110 20-OCT-2015	1	0,0	0,0	207 25-MAR-2016	17	0,1	1,4	427 07-SEP-2016	3	0,0	3,6	625 27-MAR-2017	371	1,3	19,6
111 21-OCT-2015	1	0,0	0,0	208 26-JAN-2016	11	0,0	1,4	429 09-SEP-2016	8	0,0	3,6	628 28-MAR-2017	19	0,1	19,7
112 22-OCT-2015	1	0,0	0,0	209 27-JAN-2016	11	0,0	1,4	432 12-SEP-2016	16	0,1	9,4	627 29-MAR-2017	89	0,3	20,0
113 23-OCT-2015	2	0,0	0,0	210 28-JAN-2016	6	0,0	1,5	434 14-SEP-2016	11	0,0	9,4	628 30-MAR-2017	84	0,3	20,3
116 26-OCT-2015	1	0,0	0,0	211 29-JAN-2016	17	0,1	1,5	435 15-SEP-2016	6	0,0	9,4	629 31-MAR-2017	#####	10,1	30,4
117 27-OCT-2015	4	0,0	0,0	214 01-FEB-2016	8	0,0	1,5	440 20-SEP-2016	20	0,1	9,5	634 05-APR-2017	4	0,0	30,4
118 28-OCT-2015	2	0,0	0,0	217 03-FEB-2016	36	0,0	1,6	441 21-SEP-2016	1	0,0	1,5	635 06-APR-2017	0	0,0	30,5
119 29-OCT-2015	1	0,0	0,0	216 03-FEB-2016	3	0,0	1,6	443 23-SEP-2016	9	0,0	9,5	636 07-APR-2017	8	0,0	30,5
120 30-OCT-2015	2	0,0	0,0	217 04-FEB-2016	6	0,0	1,6	446 26-SEP-2016	1	0,0	9,5	640 11-APR-2017	69	0,2	30,7
124 03-NOV-2015	2	0,0	0,0	218 05-FEB-2016	19	0,1	1,7	447 27-SEP-2016	1	0,0	9,5	642 13-APR-2017	317	1,1	31,8
125 04-NOV-2015	2	0,0	0,0	219 06-FEB-2016	8	0,0	1,7	448 28-SEP-2016	1	0,0	9,5	643 14-APR-2017	2	0,0	31,8
126 05-NOV-2015	2	0,0	0,0	221 08-FEB-2016	4	0,0	1,7	449 29-SEP-2016	24	0,1	9,6	647 18-APR-2017	2	0,0	31,8
127 06-NOV-2015	2	0,0	0,0	222 09-FEB-2016	11	0,0	1,7	450 30-SEP-2016	7	0,0	9,6	648 19-APR-2017	5	0,0	31,8
128 07-NOV-2015	18	0,0	0,0	223 10-FEB-2016	5	0,0	1,8	451 01-OCT-2016	2	0,0	9,6	649 20-APR-2017	8	0,0	31,8
130 09-NOV-2015	5	0,0	0,0	224 11-FEB-2016	8	0,0	1,8	455 05-OCT-2016	5	0,0	9,7	655 26-APR-2017	2	0,0	31,8
131 10-NOV-2015	2	0,0	0,0	225 12-FEB-2016	14	0,0	1,8	457 07-OCT-2016	1	0,0	9,7	657 27-APR-2017	1	0,0	31,8
132 11-NOV-2015	1	0,0	0,0	228 15-FEB-2016	12	0,0	1,9	460 10-OCT-2016	2	0,0	9,7	657 28-APR-2017	4	0,0	31,8
133 12-NOV-2015	5	0,0	0,0	229 16-FEB-2016	11	0,0	1,9	463 11-OCT-2016	1	0,0	9,7	659 30-APR-2017	15	0,0	31,8
134 13-NOV-2015	6	0,0	0,0	230 17-FEB-2016	11	0,0	1,9	462 12-OCT-2016	2	0,0	9,7	661 02-MAY-2017	6	0,0	31,9
136 15-NOV-2015	1	0,0	0,0	231 18-FEB-2016	17	0,1	2,0	463 13-OCT-2016	2	0,0	9,7	662 03-MAY-2017	1	0,0	31,9
137 16-NOV-2015	11	0,0	0,0	232 19-FEB-2016	5	0,0	2,0	464 14-OCT-2016	6	0,0	9,8	663 04-MAY-2017	1	0,0	31,9
138 17-NOV-2015	7	0,0	0,0	235 22-FEB-2016	11	0,0	2,1	465 15-OCT-2016	12	0,0	9,8	665 05-MAY-2017	1	0,0	31,9
139 18-NOV-2015	6	0,0	0,0	236 23-FEB-2016	4	0,0	2,1	467 18-OCT-2016	1	0,0	9,8	666 08-MAY-2017	7	0,0	31,9
140 19-NOV-2015	5	0,0	0,0	237 24-FEB-2016	29	0,1	2,2	468 19-OCT-2016	1	0,0	9,8	667 09-MAY-2017	4	0,0	31,9
141 20-NOV-2015	11	0,0	0,0	238 25-FEB-2016	12	0,0	2,2	469 20-OCT-2016	1	0,0	9,8	668 11-MAY-2017	606	2,1	34,8
144 23-NOV-2015	12	0,0	0,0	239 26-FEB-2016	33	0,1	2,3	470 21-OCT-2016	1	0,0	9,8	673 16-MAY-2017	10	0,0	32,0
145 24-NOV-2015	11	0,0	0,0	240 27-FEB-2016	1	0,0	2,3	472 23-OCT-2016	10	0,0	9,8	674 17-MAY-2017	160	0,5	32,5
146 25-NOV-2015	18	0,1	0,0	242 29-FEB-2016	13	0,0	2,4	474 25-OCT-2016	2	0,0	9,8	675 18-MAY-2017	1	0,0	32,5
147 26-NOV-2015	4	0,0	0,0	243 01-MAR-2016	14	0,0	2,4	475 26-OCT-2016	1	0,0	9,8	676 19-MAY-2017	40	0,1	32,6
148 27-NOV-2015	4	0,0	0,0	244 02-MAR-2016	5	0,0	2,4	476 27-OCT-2016	9	0,0	9,8	679 22-MAY-2017	29	0,1	32,6
149 28-NOV-2015	1	0,0	0,0	245 03-MAR-2016	46	0,2	2,6	477 28-OCT-2016	10	0,0	9,9	680 23-MAY-2017	23	0,1	32,7
150 29-NOV-2015	1	0,0	0,0	246 04-MAR-2016	6	0,0	2,6	480 14-NOV-2016	6	0,0	9,9	681 24-MAY-2017	1	0,0	32,7
151 30-NOV-2015	7	0,0	0,0	249 07-MAR-2016	45	0,2	2,8	482 02-NOV-2016	4	0,0	9,9	682 25-MAY-2017	1	0,0	32,7
152 01-DEC-2015	6	0,0	0,0	250 08-MAR-2016	2	0,0	2,8	484 04-NOV-2016	3	0,0	9,9	683 26-MAY-2017	11	0,0	32,8
153 02-DEC-2015	11	0,0	0,0	251 09-MAR-2016	5	0,0	2,8	487 07-NOV-2016	1	0,0	9,9	685 28-MAY-2017	1	0,0	32,8
154 03-DEC-2015	2	0,0	0,0	252 10-MAR-2016	13	0,0	2,9	488 08-NOV-2016	19	0,1	10,0	686 29-MAY-2017	1	0,0	32,8
155 04-DEC-2015	8	0,0	0,0	253 11-MAR-2016	19	0,1	2,9	490 10-NOV-2016	245	0,8	10,8	687 30-MAY-2017	50	0,2	35,0
156 05-DEC-2015	2	0,0	0,0	256 14-MAR-2016	14	0,0	2,9	493 13-NOV-2016	5	0,0	10,8	688 31-MAY-2017	9	0,0	35,0
158 07-DEC-2015	1	0,0	0,0	257 15-MAR-2016	15	0,0	3,0	494 14-NOV-2016	19	0,1	10,9	689 01-JUN-2017	617	2,2	35,2
159 08-DEC-2015	1	0,0	0,0	258 16-MAR-2016	31	0,1	3,1	496 16-NOV-2016	2	0,0	10,9	689 05-JUN-2017	24	0,1	35,3
160 09-DEC-2015	5	0,0	0,0	259 17-MAR-2016	9	0,0	3,2	497 17-NOV-2016	7	0,0	10,9	691 06-JUN-2017	251	0,9	35,6
161 10-DEC-2015	3	0,0	0,0	260 18-MAR-2016	41	0,1	3,3	500 20-NOV-2016	4	0,0	10,9	695 07-JUN-2017	45	0,2	35,6
165 14-DEC-2015	19	0,1	0,0	261 19-MAR-2016	12	0,0	3,3	501 21-NOV-2016	6	0,0	10,9	696 08-JUN-2017	10	0,0	35,6
166 15-DEC-2015	5	0,0	0,0	262 21-MAR-2016	18	0,1	3,4	502 22-NOV-2016	1	0,0	10,9	697 09-JUN-2017	656	2,2	35,8
167 16-DEC-2015	12	0,0	0,0	263 22-MAR-2016	11	0,0	3,4	503 23-NOV-2016	1	0,0	10,9	700 12-JUN-2017	2	0,0	35,8
168 18-DEC-2015	4	0,0	0,0	264 23-MAR-2016	28	0,1	3,5	505 25-NOV-2016	6	0,0	10,9	702 14-JUN-2017	14	0,0	35,8
172 21-DEC-2015	5	0,0	0,0	265 24-MAR-2016	16	0,0	3,5	508 28-NOV-2016	1	0,0	10,9	704 16-JUN-2017	3	0,0	35,8
173 22-DEC-2015	1	0,0	0,0	270 29-MAR-2016	10	0,0	3,6	509 29-NOV-2016	1	0,0	10,9	707 19-JUN-2017	154	0,5	35,9
179 28-DEC-2015	1	0,0	0,0	271 30-MAR-2016	23	0,1	3,7	510 30-NOV-2016	16	0,1	11,0	708 20-JUN-2017	69	0,2	35,9
180 29-DEC-2015	2	0,0	0,0	272 31-MAR-2016	9	0,0	3,7	511 01-DEC-2016	9	0,0	11,0	709 21-JUN-2017	15	0,1	35,9
186 04-JAN-2016	7	0,0	0,0	273 01-APR-2016	12	0,0	3,8	512 02-DEC-2016	15	0,1	11,1	710 22-JUN-2017	5	0,0	35,9
187 05-JAN-2016	1	0,0	0,0	275 04-APR-2016	21	0,1	3,9	516 06-DEC-2016	11	0,0	11,1	711 23-JUN-2017	6	0,0	35,9
189 07-JAN-2016	5	0,0	0,0	276 05-APR-2016	11	0,0	4,0	517 07-DEC-2016	2	0,0	11,1	714 26-JUN-2017	72	0,2	35,9
190 08-JAN-2016	8	0,0	0,0	277 06-APR-2016	48	0,1	4,0	519 09-DEC-2016	15	0,1	11,2	715 27-JUN-2017	4	0,0	35,9
193 11-JAN-2016	12	0,0	0,0	278 07-APR-2016	21	0,1	4,1	524 14-DEC-2016	19	0,1	11,2	716 28-JUN-2017	325	1,1	40,9
194 12-JAN-2016	17	0,1	0,0	279 08-APR-2016	26	0,1	4,2	525 15-DEC-2016	13	0,0	11,3	717 29-JUN-2017	45	0,2	41,0
195 13-JAN-2016	18	0,1	0,0	281 10-APR-2016	17	0,1	4,3	526 16-DEC-2016	2	0,0	11,3	718 30-JUN-2017	#####	4,1	45,1
196 14-JAN-2016	7	0,0	0,0	281 11-APR-2016	17	0,1	4,3	530 20-DEC-2016	43	0,1	11,4	722 04-JUL-2017	9	0,0	45,1
197 15-JAN-2016	1	0,0	0,0	282 12-APR-2016	11	0,0	4,3	531 21-DEC-2016	10	0,0	11,4	723 05-JUL-2017	8	0,0	45,2
198 16-JAN-2016	5	0,0	0,0	283 13-APR-2016	3	0,0	4,3	532 22-DEC-2016	3	0,0	11,5	724 06-JUL-2017	402	1,4	45,5
200 18-JAN-2016	20	0,1	0,0	284 14-APR-2016	9	0,0	4,3	533 23-DEC-2016	7	0,0	11,5	728 10-JUL-2017	3	0,0	45,5
				285 15-APR-2016	78	0,3	4,6	537 28-DEC-2016	9	0,0	11,5	730 12-JUL-2017	3	0,0	45,5
				286 16-APR-2016	5	0,0	4,6	538 29-DEC-2016	7	0,0	11,5	731 13-JUL-2017	1	0,0	45,5
				287 18-APR-2016	27	0,1	4,7	539 30-DEC-2016	12	0,0	11,6	732 14-JUL-2017	2	0,0	45,5
				288 19-APR-2016	21	0,1	4,8	540 31-DEC-2016	72	0,2	11,8	735 17-JUL-2017	1	0,0	45,5
				289 20-APR-2016	26	0,1	4,9	549 09-JAN-2017	2	0,0	11,8	737 19-JUL-2017	14	0,0	45,6
				290 21-APR-2016	13	0,0	4,9	556 16-JAN-2017	5	0,0	11,9	738 20-JUL-2017	1	0,0	45,6
				291 22-APR-2016	11	0,0	4,9	559 19-JAN-2017	2	0,0	11,9				

L'appaiamento è stato realizzato con la procedura del *Propensity Score Matching* in SPSS utilizzando una metodologia a due stadi, che l'Istat applica nei casi particolarmente complessi¹ come quello dell'incrocio di due grandi indagini campionarie, quando è molto difficile l'appaiamento.

La metodologia è basata su due constatazioni:

- la procedura probabilistica è indispensabile per trattare contemporaneamente più variabili complesse che non sarebbe possibile incrociare in modo esatto,
- possono esserci categorie sulle quali, per la loro particolare importanza, si vuole invece un incrocio esatto, e cioè che ad ogni trattato corrisponda un sosia identico.

In questo caso l'Istat utilizza la **tecnica delle celle**. Il campione di trattati e di candidati sosia all'arruolamento viene ripartito in un numero adeguato di gruppi, corrispondenti alle celle dell'incrocio, e la procedura probabilistica viene avviata per ciascuna di queste celle. A procedura conclusa su tutti i gruppi l'archivio che contiene trattati e sosia viene ricomposto. Nel nostro caso sono stati ritenuti essenziali per valutare l'esito delle azioni sul mercato del lavoro il sesso e il livello di istruzione. Quindi le celle sono state costruite distinguendo maschi e femmine con istruzione bassa (senza titolo, elementare e media), media (qualifica e diploma) e alta (laurea o dottorato). La classificazione dei titoli di studio nelle COB, per gestire l'appaiamento è stata molto macchinosa in quanto nelle comunicazioni obbligatorie molti titoli di studio, disponendo nell'estrazione del solo campo alfabetico descrittivo del titolo, hanno una denominazione incerta, che in caso di forte rischio di errore è stata considerata mancante e sono stati utilizzati candidati sosia con titolo sicuro.

Tabella 22 - Gruppi di trattati e candidati sosia distinti per celle

Tavola di contingenza fasce_per_psm * vCHI

		(vCHI)			
		0 controllo	1 caso	Totale	
Celle per PSM (fasce_per_psm)	11 Maschi - basso titolo studio	N.	62531	5920	68451
		% in vCHI	37,4%	20,1%	34,8%
	12 Maschi - medio titolo studio	N.	13574	5059	18633
		% in vCHI	8,1%	17,2%	9,5%
	13 Maschi - alto titolo studio	N.	6566	1366	7932
		% in vCHI	3,9%	4,6%	4,0%
	21 Femmine - basso titolo studio	N.	53373	6038	59411
		% in vCHI	31,9%	20,5%	30,2%
	22 Femmine - medio titolo studio	N.	19381	7880	27261
		% in vCHI	11,6%	26,8%	13,9%
	23 Femmine - alto titolo studio	N.	11827	3134	14961
		% in vCHI	7,1%	10,7%	7,6%
	Totale	N.	167252	29397	196649
		% in vCHI	100,0%	100,0%	100,0%

¹ A. Sisto (2006), Propensity score matching: un'applicazione per la creazione di un database integrato ISTAT-Banca d'Italia, <http://polis.unipmn.it/pubbl/RePEc/uca/ucapdv/sisto63.pdf>.

Nel nostro caso, tuttavia, è presente un'ulteriore complessità, cui abbiamo già accennato.

I trattati devono essere distinti in tre condizioni, che sono presenti in ciascuna delle celle di cui sopra:

- Trattati che hanno lavorato negli anni precedenti la fine della formazione e che quindi possono avere dei sosia nelle COB tra coloro che hanno lavorato prima della medesima data, che per loro è la data della DID (fascia che abbiamo denominato TEST LAVORO).
- Trattati che NON hanno lavorato prima di finire la formazione ma hanno lavorato dopo e che quindi possono avere dei sosia nelle COB tra coloro che hanno lavorato dopo della medesima data (fascia TEST su cui possiamo confrontare SOLO il TIPO di LAVORO).
- Trattati che non hanno lavorato né prima né dopo la formazione e che non possono avere sosia nelle COB e quindi devono essere accantonati rispetto al confronto con altri soggetti non trattati (fascia NO TEST).

La procedura applicata ha quindi previsto che l'appaiamento sia elaborato all'interno di ciascuna delle 3 fasce, a loro volta all'interno delle 6 celle. Le fasce sono definite in base alla presenza del soggetto nelle COB prima, dopo, o mai. Le celle sono definite in base al sesso e al livello di istruzione.

Il punteggio di propensione all'interno delle 12 aree stabilite come sopra è stato calcolato utilizzando 5 variabili:

- 1) Numero anni di istruzione (che varia anche all'interno delle 3 fasce utilizzate per le celle),
- 2) cittadinanza italiana o straniera (1 italiana),
- 3) anno di nascita (l'anno più alto identifica i giovani),
- 4) numero giorni trascorsi nel range di apertura della finestra temporale tra inizio delle DID (12 ottobre 2015) e fine delle DID (21 dicembre 2018) (numero di giorni alto significa finestra più breve, dato che per tutti le occasioni di lavoro sono rilevate fino al 31 dicembre 2019),
- 5) Indice di distanza geografica e culturale del luogo di nascita dall'Emilia Romagna (massima Africa).

I modelli di appaiamento sono stati quindi guidati dalla regressione logistica all'interno di ciascun gruppo, è stato imposto un criterio di corrispondenza nel range tra 0 e 1 di 0,2 quindi piuttosto esigente e tutte le corrispondenze sono state trovate in numero esattamente corrispondente ai casi disponibili con metodo probabilistico senza sostituzione. I parametri dei modelli sono soddisfacenti con una capacità di previsione che varia tra il 75% e il 95%, anche se a volte all'interno di celle così definite il modello non pare aver migliorato le capacità di previsione già derivate dagli incroci. Il test WALD per i modelli è sempre altamente significativo come quello delle variabili utilizzate per il calcolo della propensione all'interno di ciascuna cella. Il range di varianza spiegata è stato anche controllato con i test R-quadrato di Cox e Snell e R-quadrato di Nagelkerke forniti di default dal sistema.

I gruppi contenenti i trattati e i sosia appaiati in ugual numero, derivati dalle 6 celle, sono stati poi ricomposti in un unico archivio per ciascuna delle due fasce su cui l'abbinamento è stato possibile sulle 3 fasce disponibili, che distinguono i gruppi proprio per le analisi che è possibile svolgere. Disponiamo quindi per l'analisi di tre archivi in base alle fasce di cui due con trattati e sosia abbinati.

Il primo archivio (tabella 21) di soggetti che hanno lavorato prima della data fine FSE o DID contiene 48.752 soggetti di cui la metà trattati nell'ambito delle iniziative FSE.

Tabella 23 - Gruppi di trattati e candidati sosia distinti per celle - soggetti che hanno lavorato prima della data fine FSE o DID

Tavola di contingenza fasce_per_psm * vCHI

		vCHI		Totale	
		0 controllo	1 caso		
Celle per PSM (fasce_per_psm)	11 Maschi - basso titolo studio	N.	4781	4781	9562
		% in vCHI	19,6%	19,6%	19,6%
	12 Maschi - medio titolo studio	N.	4273	4273	8546
		% in vCHI	17,5%	17,5%	17,5%
	13 Maschi - alto titolo studio	N.	1132	1132	2264
		% in vCHI	4,6%	4,6%	4,6%
	21 Femmine - basso titolo studio	N.	4847	4847	9694
		% in vCHI	19,9%	19,9%	19,9%
	22 Femmine - medio titolo studio	N.	6741	6741	13482
		% in vCHI	27,7%	27,7%	27,7%
	23 Femmine - alto titolo studio	N.	2602	2602	5204
		% in vCHI	10,7%	10,7%	10,7%
Totale	N.	24376	24376	48752	
	% in vCHI	100,0%	100,0%	100,0%	

Il secondo archivio (tabella 24), di soggetti che NON hanno lavorato prima della data fine FSE o DID ma solo DOPO, contiene 4.904 soggetti di cui la metà trattati nell'ambito delle iniziative FSE.

Tabella 24 - Gruppi di trattati e candidati sosia distinti per celle - oggetti che NON hanno lavorato prima della data fine FSE o DID ma solo DOPO

Tavola di contingenza fasce_per_psm * vCHI

		vCHI		Totale	
		,00 controllo	1,00 caso		
Celle per PSM (fasce_per_psm)	11 Maschi - basso titolo studio	N.	601	601	1202
		% in vCHI	24,5%	24,5%	24,5%
	12 Maschi - medio titolo studio	N.	478	478	956
		% in vCHI	19,5%	19,5%	19,5%
	13 Maschi - alto titolo studio	N.	133	133	266
		% in vCHI	5,4%	5,4%	5,4%
	21 Femmine - basso titolo studio	N.	407	407	814
		% in vCHI	16,6%	16,6%	16,6%
	22 Femmine - medio titolo studio	N.	559	559	1118
		% in vCHI	22,8%	22,8%	22,8%
	23 Femmine - alto titolo studio	N.	274	274	548
		% in vCHI	11,2%	11,2%	11,2%
Totale	Conteggio	2452	2452	4904	
	% in vCHI	100,0%	100,0%	100,0%	

Esaminando le caratteristiche dei gruppi che verranno usati per il test sul lavoro possiamo vedere come la procedura del PSM, associata al metodo delle celle, abbia assicurato una forte corrispondenza tra trattati e sosia. L'unica variabile che mostra uno scostamento, nonostante fosse tra quelle introdotte nell'algoritmo di calcolo, e quindi da guardare con attenzione è quella relativa alla distanza tra l'avvio dell'esperimento e la

dichiarazione di disponibilità della persona o la fine dell'attività formativa. Su questo i sosia mostrano un livello nettamente più basso di giorni trascorsi e quindi si sono mediamente presentati prima sul mercato. Dato che la variabile era presente e significativa in tutte le regressioni che hanno calcolato la propensione, ne deduciamo che il risultato rappresenti comunque il migliore match possibile, anche perché aumentando il rigore dell'abbinamento, ad esempio passando da 0,2 a 0,1, iniziavano a registrarsi casi di mancato abbinamento. Per evitare che questo dato abbia effetto sul risultato della stima, introdurremo la medesima variabile anche nella regressione logistica o lineare per valutare la probabilità di reperire un lavoro, in modo che il suo eventuale effetto differenziale possa essere azzerato.

Tabella 25 - Caratteristiche di trattati e sosia dopo PSM – Gruppo 1

vCHI		(sesso_PSM) Sesso per PSM	titolo_studio_P SM	anno_nascita_ PSM	(fasce_prov_na scita_PSM) Fasce luogo di nascita	(newcittadinan za2_PSM) cittadinanza_P SM_sintesi	ngiornidid_PS M
0 controllo	Media	,5821	11,0752	1975,31	8,4062	1,3988	254,33
	Mediana	1,0000	13,0000	1975,00	3,0000	2,0000	168,00
	Deviazione std.	,49322	4,46382	12,177	10,92370	,91708	185,722
	N	24376	24376	24376	24376	24376	24376
1 caso	Media	,5821	10,8405	1976,46	8,1452	1,4120	802,42
	Mediana	1,0000	13,0000	1976,00	2,0000	2,0000	856,00
	Deviazione std.	,49322	4,77313	11,962	11,64388	,80779	269,537
	N	24376	24376	24376	24376	24376	24376
Totale	Media	,5821	10,9579	1975,89	8,2757	1,4054	528,37
	Mediana	1,0000	13,0000	1975,00	2,0000	2,0000	505,00
	Deviazione std.	,49321	4,62250	12,083	11,29017	,86418	358,709
	N	48752	48752	48752	48752	48752	48752

Tabella 26 - Caratteristiche di trattati e sosia dopo PSM – Gruppo 2

vCHI		(sesso_PSM) Sesso per PSM	titolo_studio_P SM	anno_nascita_ PSM	(fasce_prov_na scita_PSM) Fasce luogo di nascita	(newcittadinan za2_PSM) cittadinanza_P SM_sintesi	ngiornidid_PS M
0 controllo	Media	,5057	10,5396	1984,66	10,7414	1,4184	178,97
	Mediana	1,0000	13,0000	1989,00	3,0000	2,0000	125,00
	Deviazione std.	,50007	5,35112	12,067	13,51995	,90843	154,443
	N	2452	2452	2452	2452	2452	2452
1 caso	Media	,5057	10,2916	1985,93	13,3789	1,1652	793,04
	Mediana	1,0000	13,0000	1990,00	3,0000	2,0000	856,00
	Deviazione std.	,50007	5,52449	11,592	14,68254	,92035	315,474
	N	2452	2452	2452	2452	2452	2452
Totale	Media	,5057	10,4156	1985,30	12,0602	1,2918	486,00
	Mediana	1,0000	13,0000	1990,00	3,0000	2,0000	302,50
	Deviazione std.	,50002	5,43936	11,847	14,17328	,92305	394,925
	N	4904	4904	4904	4904	4904	4904

La prima analisi che possiamo condurre è la più semplice e consiste nel confrontare trattati e sosia rispetto al fatto di avere avuto una occasione di lavoro successiva alla fine della formazione o alla DID, entro marzo 2020, seguendo l'ipotesi che il trattamento FSE possa favorire l'occupazione. Come si può osservare il 68,9% dei trattati ha lavorato mentre la percentuale arriva al 81,5% dei sosia. Tale risultato è statisticamente significativo, anche se risente in parte di una diversa esposizione temporale alla ricerca del lavoro, come abbiamo osservato discutendo i risultati del PSM, che compenseremo integralmente nelle analisi seguenti.

Tabella 27 - Trattati e sosia che hanno lavorato dopo la fine della formazione/DID

Tavola di contingenza cob_dopo_NEW * vCHI

			vCHI		Totale
			0 controllo	1 caso	
Lavorato dopo (cob_dopo_NEW)	0 No	Conteggio	4511	7589	12100
		% in vCHI	18,5%	31,1%	24,8%
	1 Sì	Conteggio	19865	16787	36652
		% in vCHI	81,5%	68,9%	75,2%
Totale	Conteggio	24376	24376	48752	
	% in vCHI	100,0%	100,0%	100,0%	

Questa differenza, che vogliamo approfondire, può essere dovuta a diversi ordini di spiegazioni, che possiamo riferire alle classi di informazioni di cui dovrebbe disporre l'analisi ideale:

- 1) Prima classe di informazioni: profilo delle persone coinvolte nell'esperimento. Sono variabili di controllo e servono ad abbinare i sosia ai trattati.
- 2) Seconda classe di informazioni: tipo di effetto che ci si attende dalla formazione.
- 3) Terza classe di informazioni: aver lavorato dopo il trattamento e qualità di questo lavoro.

Riguardo alla classe 1, le persone che sono state in trattamento con il FSE potrebbero essere più deboli e per questo incontrare maggiori difficoltà. Potrebbero anche essere meno interessate a lavorare immediatamente, dato che l'aver partecipato ad attività FSE, diversamente dalla manifestazione della DID, non rappresenta necessariamente la dichiarazione di immediata disponibilità. Per ridurre questo eventuale effetto i pochi casi che erano indicati come inattivi nell'archivio FSE non sono stati inseriti in questo database.

Riguardo alla classe 2, si tratta di valutare se l'attività svolta per contenuto e durata può effettivamente rappresentare un rafforzamento nella ricerca del lavoro.

Riguardo alla classe 3, si tratterebbe di valutare la qualità del lavoro, cosa che come abbiamo detto, è possibile solo marginalmente per le informazioni presenti nell'archivio estratto.

È utile comunque valutare l'esito in termini di probabilità di trovare un lavoro con una regressione logistica binaria multivariata, che consente di misurare l'effetto a parità delle altre informazioni di cui disponiamo.

Il modello (tabella 28) è significativo e stima correttamente tre quarti dei casi (75,3%). Il modello conferma il fatto che le persone trattate nell'ambito delle iniziative FSE hanno una minore probabilità di aver lavorato nel periodo successivo, a parità di tutte le altre condizioni presenti nel modello. In particolare, leggendo l'Exp(B), che indica lo stesso rapporto di probabilità dell'*odds ratio*, e il suo intervallo di confidenza, oltre che la sua significatività, possiamo constatare che, a parità di tutte le altre condizioni:

- Le donne restano sfavorite, con 0,87 probabilità di aver lavorato dopo contro 1 dei maschi.
- Il titolo di studio non è significativo, perché stiamo confrontando soggetti che abbiamo appaiato con sosia all'interno della medesima fascia di appartenenza del titolo di studio: basso (senza titolo, elementare o media), medio (qualifica o diploma) e alto (laurea o dottorato).
- L'anno di nascita è molto significativo, considerando che essere nati un anno dopo vuol dire avere mediamente il 4% di probabilità in più di aver lavorato (1,041 contro 1). Per i più vecchi la probabilità di aver lavorato è poco sopra il 40% mentre per i più giovani è nettamente sopra l'80%.
- Cittadinanza e distanza geografica sono significative con effetti non lineari. I cittadini italiani hanno lavorato un po' meno, con un effetto di qualche punto percentuale, ma anche le persone di provenienza molto lontana, come Africa e Asia paiono leggermente sfavorite. Stranieri o italiani che provengono da luoghi più vicini sembrano aver lavorato di più.
- La distanza della DID dalla chiusura dell'estrazione è significativa perché avere usato un mese in meno nella ricerca vuol dire circa lo 0,02% in meno di probabilità di aver trovato un lavoro. Ma per noi si tratta di un dato esclusivamente tecnico, che ci serve a fare il calcolo dei fattori che influiscono sul reperimento del lavoro a parità del periodo dedicato alla ricerca.
- Infine, a parità di queste condizioni, le persone presenti nell'archivio FSE risultano aver lavorato di meno, con una penalizzazione stimabile a parità di altre condizioni in 0,72 contro 1, e quindi tra il 25% e 30%.

Tabella 28 - Regressione logistica

		Variabili nell'equazione					
		B	S.E.	Wald	gl	Sign.	Exp(B)
Fase 0	Costante	1,108	,010	11173,158	1	,000	3,029

Riepilogo del modello

Fase	Logaritmo della verosimiglianza - 2	R-quadrato di Cox e Snell	R-quadrato di Nagelkerke
1	51103,072 ^a	,070	,104

a. Stima terminata all'iterazione numero 5 perché le stime dei parametri sono state modificate in misura inferiore a ,001.

Tabella di classificazione^a

Osservato		Previsto		Percentuale di correttezza	
		Lavorato dopo (cob_dopo_NEW) 0 No	1 Sì		
Fase 1	Lavorato dopo (cob_dopo_NEW)	0 No	643	11457	5,3
		1 Sì	419	36233	98,9
Percentuale globale					75,6

a. Il valore di divisione è ,500

		Variabili nell'equazione						95% C.I. per EXP(B)	
		B	S.E.	Wald	gl	Sign.	Exp(B)	Inferiore	Superiore
Fase	vCHI	-,327	,034	94,051	1	<,001	,721	,675	,770
1 ^a	Sesso per PSM	-,138	,023	36,028	1	<,001	,871	,833	,911
	titolo_studio_PSM	-,002	,003	,480	1	,488	,998	,993	1,003
	anno_nascita_PSM	,040	,001	1757,096	1	,000	1,041	1,039	1,043
	cittadinanza_PSM_sintesi	-,046	,017	7,606	1	,006	,955	,924	,987
	Fasce luogo di nascita	-,006	,001	21,304	1	<,001	,994	,991	,997
	mesidid	-,018	,001	317,503	1	<,001	,982	,980	,984
	Costante	-	1,883	1682,112	1	,000	,000		
		77,228							

a. Variabili inserite nella fase 1: vCHI, Sesso per PSM, titolo_studio_PSM, anno_nascita_PSM, cittadinanza_PSM_sintesi, Fasce luogo di nascita, mesidid.

Come abbiamo detto, questa minore probabilità di lavorare potrebbe indicare una debolezza, che nei quasi esperimenti è controllabile solo disponendo di ulteriori informazioni in grado di ridurre l'eterogeneità non osservabile, che avrebbero potuto anche essere ricavate con appropriate estrazioni da informazioni già presenti negli archivi amministrativi del lavoro e della formazione e quindi acquisibili e trattabili a basso costo. Certamente il fatto che la platea FSE sia più fragile della media potrebbe anche essere considerata una conferma di aderenza del FSE alla sua missione. La domanda che pone questa analisi riguarda piuttosto gli effetti del trattamento. Se è vero che parrebbe esagerato aspettarsi un cambiamento significativo di forza sul mercato del lavoro dopo aver frequentato solo un corso di breve durata, come molti di quelli erogati e descritti nel cap. 2, ci si potrebbe aspettare di più in presenza di interventi strutturati finalizzati al rilascio di qualifiche.

Date le caratteristiche degli archivi che abbiamo allestito, possiamo condurre rapidamente un secondo esperimento, utilizzando i sosia già selezionati. Per rispondere alla domanda posta, possiamo estrarre dall'archivio che abbiamo appena utilizzato i soli casi di trattati FSE che hanno svolto percorsi per l'acquisizione di qualifiche regionali, ma anche di accesso ad aree professionali e di approfondimento tecnico, lasciando appaiati i loro specifici sosia. Si tratta dei 5.685 soggetti rappresentati nella mappa dedicata ai percorsi con il riquadro verde in basso nella pagina della mappa 10.

La tabella 29 rappresenta le celle per l'abbinamento e quindi il numero di trattati e non trattati per sesso e livello di istruzione. Possiamo notare che il numero di laureati è più elevato di quello dell'esperimento precedente e arriva al 20% (più femmine che maschi) e sono molti i titoli intermedi (circa 60%), pur rimanendo assai significativi anche i titoli bassi, che oggi sono sotto l'obbligo scolastico (quasi il 20%, ma erano quasi il 40% nell'intera platea di trattati FSE). In sostanza si tratta di una platea mista più forte di quella precedente, cui sono abbinati sosia appropriati di forza simile.

Tabella 29 - Celle PSM solo chi ha svolto attività volte all'acquisizione di qualifiche

Tavola di contingenza fasce_per_psm Celle per PSM * vCHI

		vCHI		Totale	
		0 controllo	1 caso		
Celle per PSM (fasce_per_psm)	11 Maschi - basso titolo studio	N.	601	601	1202
		% in vCHI	12,8%	12,8%	12,8%
	12 Maschi - medio titolo studio	N.	1483	1483	2966
		% in vCHI	31,6%	31,6%	31,6%
	13 Maschi - alto titolo studio	N.	363	363	726
		% in vCHI	7,7%	7,7%	7,7%
	21 Femmine - basso titolo studio	N.	274	274	548
		% in vCHI	5,8%	5,8%	5,8%
	22 Femmine - medio titolo studio	N.	1336	1336	2672
		% in vCHI	28,4%	28,4%	28,4%
	23 Femmine - alto titolo studio	N.	642	642	1284
		% in vCHI	13,7%	13,7%	13,7%
Totale	N.	4699	4699	9398	
	% in vCHI	100,0%	100,0%	100,0%	

Il comportamento rispetto al lavoro successivo alla formazione viene di nuovo esplorato con la regressione logistica. Premettiamo che le percentuali di conseguimento di un lavoro appaiono subito diverse e la distanza con i sosia si accorcia molto, da 11,5 a 3,5 punti percentuali. Soprattutto però, riformulando il confronto a parità di condizioni, mentre restano significativi per favorire la ricerca rapida del lavoro essere maschi e giovani, con un leggero vantaggio di cittadinanze straniere, ma anche dell'essere nati in Emilia o vicino, piuttosto che lontano, il vantaggio dell'essere sosia rispetto all'essere trattato FSE si azzerava e non ha nessun rilievo statistico.

Possiamo quindi dire che il FSE attrae o si rivolge ad una platea debole, che resta tale sul mercato del lavoro, rispetto al reperimento di occasioni nei mesi e anni successivi, ma se l'intervento formativo è significativo, questo azzerava il gap di debolezza con i sosia ($\text{Exp}(B)=1,020$) non significativo e incluso nell'intervallo di confidenza). Ricordiamo che sia i trattati che i sosia hanno già lavorato precedentemente. Come abbiamo spiegato, per i limiti dell'archivio non possiamo fare lo stesso test sulle persone in cerca di prima occupazione, trattati e sosia.

Anche altre forme di intervento FSE potrebbero incidere sulla probabilità di lavorare, come i tirocini e gli accompagnamenti, ma su queste tipologie facciamo qualche approfondimento di seguito.

Tabella 30 - Regressione logistica, solo chi ha attività volte all'acquisizione di qualifiche

		Variabili nell'equazione					Exp(B)
		B	S.E.	Wald	gl	Sign.	
Fase 0	Costante	1,596	,028	3354,588	1	,000	4,933

Riepilogo del modello

Fase	Logaritmo della verosimiglianza - 2	R-quadrato di Cox e Snell	R-quadrato di Nagelkerke
1	8156,657 ^a	,038	,065

a. Stima terminata all'iterazione numero 5 perché le stime dei parametri sono state modificate in misura inferiore a ,001.

Tabella di classificazione^a

Osservato		Previsto			
		Lavorato dopo (cob_dopo_NEW)		Percentuale di correttezza	
		0 No	1 Sì		
Fase 1	Lavorato dopo	0 No	2	1582	,1
	(cob_dopo_NEW)	1 Sì	0	7814	100,0
Percentuale globale					83,2

a. Il valore di divisione è ,500

Variabili nell'equazione

		B	S.E.	Wald	gl	Sign.	Exp(B)	95% C.I. per EXP(B)	
								Inferiore	Superiore
Fase	vCHI	,020	,069	,084	1	,771	1,020	,891	1,169
1 ^a	Sesso per PSM	-,136	,058	5,381	1	,020	,873	,779	,979
	titolo_studio_PSM	-,011	,008	1,901	1	,168	,989	,974	1,005
	anno_nascita_PSM	,036	,003	192,530	1	<,001	1,036	1,031	1,042
	cittadinanza_PSM_sintesi	-,007	,045	,025	1	,873	,993	,909	1,085
	Fasce luogo di nascita	-,014	,003	17,816	1	<,001	,986	,980	,992
	mesidid	-,021	,002	72,595	1	<,001	,980	,975	,984
	Costante	-68,375	5,079	181,262	1	<,001	,000		

a. Variabili inserite nella fase 1: vCHI, Sesso per PSM, titolo_studio_PSM, anno_nascita_PSM, cittadinanza_PSM_sintesi, Fasce luogo di nascita, mesidid.

Un altro punto di osservazione sul lavoro svolto dopo la formazione riguarda il numero di occasioni di lavoro. Non potendo con le informazioni disponibili definire livello, tipo di contratto e durata (per l'assenza di specifiche informazioni e identificativo sicuro del rapporto di lavoro cui si riferiscono i diversi eventi, comunicazioni e cessazioni) possiamo tuttavia contare le comunicazioni di avviamento nella finestra temporale a disposizione.

Come mostra la figura successiva, abbiamo registrato 139mila comunicazioni di avvio rapporto sulle 48mila persone del nostro archivio trattati e sosia. Ci sono circa 12mila persone che hanno zero avviamenti e 1 che ne conta 167 (massimo), caso possibile per chi svolge lavori a giornata come ad esempio nel settore alberghiero-ristorazione o nell'interinale. La media è di 2,8 comunicazioni a persona.

Tabella 31

	Statistiche descrittive					
	N	Minimo	Massimo	Somma	Media	Deviazione std.
n_occasioni_dopo	48752	,00	167,00	139219,00	2,8557	5,41234
Numero di casi validi (listwise)	48752					

Da questo punto di osservazione possiamo quindi vedere se emergono differenze tra i soggetti trattati FSE e i sosia. Per questo esperimento utilizziamo la regressione lineare, dato che la variabile dipendente non è dicotomica ma metrica.

Utilizzando le stesse variabili usate nel modello logistico, dedicato alla probabilità di lavorare, notiamo che in questo caso le persone trattate nell'ambito FSE sono mediamente più dinamiche e hanno significativamente avuto in media più occasioni (quasi mezza occasione in più in media a persona se guardiamo il coefficiente B dell'equazione 0,40). Non è più significativo il genere, mentre essere giovani è ancora essenziale e significativo.

Il segno negativo del titolo di studio, anche se non significativo, ci segnala che queste occasioni si riferiscono a segmenti a bassa qualificazione. L'analisi del luogo di nascita evidenzia una maggiore frequenza di lavori per le persone provenienti dall'Africa, dall'Asia e dal sud Italia. Possiamo quindi dire che mentre si conferma la maggiore debolezza delle persone cui il FSE si rivolge, nei fatti, per una parte importante di loro il passaggio attraverso i servizi è seguito da una maggiore vivacità di presenza sul mercato rispetto ai loro sosia che, ricordiamo, sono del tutto analoghi per le caratteristiche di sesso, titolo di studio, età e provenienza.

Tabella 32 - Regressione lineare multipla (numero di occasioni di lavoro per tutti)

Riepilogo del modello^b

Modello	R	R-quadrato	R-quadrato adattato	Errore std. della stima	Durbin-Watson
1	,131 ^a	,017	,017	5,36588	1,997

a. Predittori: (costante), newcittadinanza2_PSM cittadinanza_PSM_sintesi, vCHI, sesso_PSM Sesso per PSM, anno_nascita_PSM, titolo_studio_PSM, fasce_prov_nascita_PSM Fasce luogo di nascita, ngiornidid_PSM

b. Variabile dipendente: n_occasioni_dopo

ANOVA^a

Modello		Somma dei quadrati	gl	Media quadratica	F	Sign.
1	Regressione	24612,681	7	3516,097	122,118	<,001 ^b
	Residuo	1403470,579	48744	28,793		
	Totale	1428083,260	48751			

a. Variabile dipendente: n_occasioni_dopo

b. Predittori: (costante), newcittadinanza2_PSM cittadinanza_PSM_sintesi, vCHI, sesso_PSM Sesso per PSM, anno_nascita_PSM, titolo_studio_PSM, fasce_prov_nascita_PSM Fasce luogo di nascita, ngiornidid_PSM

Coefficienti^a

Modello	Coefficienti non standardizzati		Coefficienti standardizzati		Sign.
	B	Errore standard	Beta	t	
1 (Costante)	-66,058	4,077		-16,203	<,001
vCHI	,402	,076	,037	5,257	<,001
sesso_PSM Sesso per PSM	-,032	,051	-,003	-,636	,525
anno_nascita_PSM	,035	,002	,078	17,012	<,001
titolo_studio_PSM	-,012	,006	-,011	-2,175	,030
ngiornidid_PSM	-,002	,000	-,123	-17,345	<,001
fasce_prov_nascita_PSM Fasce luogo di nascita	,016	,003	,033	5,325	<,001
newcittadinanza2_PSM cittadinanza_PSM_sintesi	,175	,038	,028	4,636	<,001

a. Variabile dipendente: n_occasioni_dopo

Ancora più interessante è la replica dello stesso esperimento sul database che abbiamo utilizzato prima e che contiene solo i partecipanti al FSE che hanno svolto corsi più strutturati per conseguire qualifiche o approfondimenti. In questo caso i partecipanti al FSE, rispetto ai sosia, mostrano un netto vantaggio sul mercato, con quasi una occasione in più in media (B 0,93), mentre l'aver partecipato al FSE diventa tra le variabili prese in esame quella più rilevante nel determinare il risultato (Beta 0,78, consente il confronto del peso dei fattori a parità di scala di misura). Non sono più presenti effetti per genere, provenienza e nazionalità, mentre essere giovani rimane centrale (ogni 7 anni in meno una occasione in più), a segnalare che si tratta di un'area dinamica, per alcuni presumibilmente di attraversamento rapido verso la stabilità e

per altri di maggiore precarietà temporale. Non abbiamo sufficienti informazioni nell'estrazione disponibile dalle COB sui rapporti di lavoro in un tempo più prolungato per fare analisi specifiche, che sarebbero facilmente realizzabili, su questi aspetti, approfondendo almeno nei 5 anni successivi (arco di tempo utilizzato da varie indagini tra cui Almalaurea) qualità dei rapporti di lavoro, stabilità del lavoro, continuità di lavoro nello stesso settore ed evoluzione della carriera di lavoro negli anni successivi.

Tabella 33 - Regressione lineare multipla (numero di occasioni di lavoro per chi ha svolto attività di formazione per l'acquisizione di qualifiche)

Riepilogo del modello^b

Modello	R	R-quadrato	R-quadrato adattato	Errore std. della stima	Durbin-Watson
1	,092 ^a	,008	,008	5,97309	1,975

a. Predittori: (costante), newcittadinanza2_PSM cittadinanza_PSM_sintesi, sesso_PSM Sesso per PSM, vCHI, anno_nascita_PSM, titolo_studio_PSM, ngiornidid_PSM, fasce_prov_nascita_PSM Fasce luogo di nascita

b. Variabile dipendente: n_occasioni_dopo

ANOVA^a

Modello		Somma dei quadrati	gl	Media quadratica	F	Sign.
1	Regressione	2865,397	7	409,342	11,473	<,001 ^b
	Residuo	335014,728	9390	35,678		
	Totale	337880,125	9397			

a. Variabile dipendente: n_occasioni_dopo

b. Predittori: (costante), newcittadinanza2_PSM cittadinanza_PSM_sintesi, sesso_PSM Sesso per PSM, vCHI, anno_nascita_PSM, titolo_studio_PSM, ngiornidid_PSM, fasce_prov_nascita_PSM Fasce luogo di nascita

Coefficienti^a

Modello		Coefficienti non standardizzati		Coefficienti standardizzati		Sign.
		B	Errore standard	Beta	t	
1	(Costante)	-26,789	11,378		-2,354	,019
	vCHI	,937	,148	,078	6,317	<,001
	sesso_PSM Sesso per PSM	-,044	,128	-,004	-,347	,728
	anno_nascita_PSM	,015	,006	,028	2,691	,007
	titolo_studio_PSM	-,035	,018	-,021	-1,954	,051
	ngiornidid_PSM	-,002	,000	-,089	-7,200	<,001
	fasce_prov_nascita_PSM Fasce luogo di nascita	,013	,008	,023	1,639	,101
	newcittadinanza2_PSM	,123	,104	,016	1,189	,234
	cittadinanza_PSM_sintesi					

a. Variabile dipendente: n_occasioni_dopo

Per approfondire l'effetto dei diversi percorsi proposti dal FSE abbiamo realizzato una analisi fattoriale per componenti principali, con l'obiettivo di analizzare la struttura dei dati considerando 15 variabili, sia quelle già utilizzate che definiscono il profilo delle persone, sia quelle che riepilogano le attività svolte con il FSE. Attraverso un modello con rotazione ortogonale dei fattori, per avere indicazioni più nette sul significato dei medesimi, abbiamo estratto 6 fattori che sintetizzano il 67,3% della varianza della matrice. Si tratta di una analisi con valenza del tutto esplorativa e gli indici del modello sono deboli (KMO pari a 51) ma la sua utilità consiste nel fornirci dei profili che considerano l'insieme delle informazioni da cui trarre poi ulteriori indicazioni utili.

Come si vede sotto dalla matrice dei pesi fattoriali sono stati denominati 6 fattori che caratterizzano in modo più marcato, dal punto di vista statistico, la storia di questa platea di persone, trattati e sosia insieme. Li abbiamo chiamati:

1. Seguiti con accoglienza e formazione leggera
2. Seguiti con qualifiche, formazione e accompagnamento
3. Emiliani scolarizzati non seguiti dal FSE
4. Formativi con strumenti FAD e PW
5. Giovani scolarizzati seguiti attraverso tirocini
6. Donne straniere seguite per le qualifiche regionali

Salvo il fattore n.3 emergono caratteristiche relative alla platea FSE, dato che la maggior parte delle informazioni riguardano queste esperienze.

Tabella 34 – Matura della matrice – analisi delle componenti principali

	Matrice dei componenti ruotati^a					
	Componente					
	Seguiti con accoglienza formazione leggera	Seguiti con qualifiche formazione accompagnamento	Emiliani scolarizzati non seguiti	Formati con FAD e PW	Giovani scolarizzati con tirocini	Donne straniere con qualifiche regionali
Sesso (sesso_PSM)	0,023	-0,101	0,102	0,051	-0,067	0,785
Anno nascita (anno_nascita_PSM)	-0,028	0,095	-0,040	0,021	0,738	-0,170
Titolo di studio (titolo_studio_PSM)	-0,079	0,100	0,518	0,047	0,375	0,289
Luogo nascita (fasce_prov_nascita_PSM)	-0,001	-0,012	-0,891	0,001	0,023	-0,070
Cittadinanza (newcittadinanza2_PSM)	0,017	0,026	0,869	-0,003	-0,134	-0,090
Numero attività di accoglienza, orientamento (N_Accoglienza_orientamento)	0,631	0,153	-0,018	-0,022	0,032	0,009
Numero tirocini (N_Tirocini)	0,042	-0,060	-0,019	-0,028	0,688	0,069
Numero corsi di formazione permanente (N_Form_permanente)	0,934	-0,103	0,001	-0,001	-0,019	0,029
Numero corsi per qualifiche (N_Form_qualifiche)	-0,091	0,824	0,085	0,189	0,081	-0,133
Numero corsi per qualifiche del sistema regionale (N_Form_qualifiche_reg)	0,163	0,511	-0,086	-0,136	-0,040	0,433
Numero accompagnamenti al lavoro (N_Accom_lavoro)	0,040	0,442	0,036	-0,033	-0,054	-0,329
Numero project work (N_Pwork)	-0,007	0,167	0,019	0,767	0,010	0,004
N. corsi formazione a distanza (N_Fad)	0,001	-0,012	-0,005	0,803	-0,014	0,025
N. ore di formazione permanente (ore_Form_permanente)	0,895	-0,074	0,000	0,016	-0,015	0,019
Numero ore di formazione per qualifiche (ore_Form_qualifiche)	-0,008	0,933	0,031	0,119	0,043	0,066

Metodo di estrazione: Analisi dei componenti principali.

Metodo di rotazione: Varimax con normalizzazione Kaiser.

a. Convergenza per la rotazione eseguita in 6 iterazioni.

Successivamente, il punteggio ricavato per ciascun fattore dall'analisi statistica è stato riclassificato, identificando tutti i soggetti nella cui esperienza si ritrova in modo molto marcato la presenza delle diverse attività indicate dai punteggi fattoriali visibili nella tabella. Non sono associazioni di attività scelte da noi, ma presenti nei dati relativi alle persone. Quindi nuove variabili sono state costruite a partire da questi punteggi, con valore 0 per chi ha un punteggio basso e 1 per coloro che hanno valore sopra 1 (dato che tutti i fattori hanno media zero e deviazione standard 1, si tratta delle persone che hanno un punteggio superiore ad una deviazione standard).

Questi nuovi gruppi, identificati da queste variabili, insieme ai vecchi gruppi, costruiti dall'incrocio tra sesso e fascia del titolo di studio, usati per il PSM, sono stati a loro volta analizzati con regressioni logistiche e lineari per stimare, a parità di finestra temporale, la loro specifica probabilità di trovare lavoro rispetto alla media generale e il numero medio di occasioni che può essere loro attribuito.

I risultati di questa analisi sono riportati nelle tabelle seguenti e nei grafici a dispersione, che commentiamo brevemente.

Tabella 35 – Sintesi degli indici di regressione logistica e regressione lineare

Gruppo	Probabilità di lavorare dopo (Exp(B))	Numero di occasioni dopo (Coefficiente B)
11 Maschi - basso titolo studio	1,124	0,127
12 Maschi - medio titolo studio	1,495	0,259
13 Maschi - alto titolo studio	1,164	-0,62
21 Femmine - basso titolo studio	0,642	-0,469
22 Femmine -medio titolo studio	1,062	-0,123
23 Femmine - alto titolo studio	1,175	-0,42
Sosia	1,165	-0,54
Trattati	0,858	0,54
Seguiti con accoglienza formazione leggera	1,171	0,106
Seguiti con qualifiche formazione accompagnamento	1,487	0,83
Emiliani scolarizzati non seguiti	1,21	-0,17
Formati con FAD e PW	1,381	0,117
Giovani scolarizzati con tirocini	1,499	0,158
Donne straniere con qualifiche regionali	0,959	-0,561

La media generale degli oltre 24mila trattati e degli altrettanti sosia è rappresentata dai simboli quadrato (trattati) e rotondo (sosia). L'asse orizzontale rappresenta la probabilità di aver lavorato dopo l'esperienza FSE e il valore 1 indica la probabilità di chi per ciascun gruppo non è nel gruppo e rappresenta il valore degli altri di riferimento. Come si può vedere i sosia stanno a destra perché hanno lavorato mediamente di più e i trattati a sinistra del valore 1. L'asse verticale rappresenta invece la probabilità di aver fatto più lavori, dove zero è l'indifferenza. Come si può vedere i sosia hanno avuto meno occasioni dei trattati, che hanno avuto mediamente mezza occasione in più (coefficiente 0,5), cosa che è positiva ma può anche indicare debolezza di queste occasioni.

In questo contesto è interessante notare dove si posizionano i gruppi più connotati, che abbiamo individuato con l'analisi per componenti principali.

In alto a destra il gruppo di maggior investimento da parte del fondo sociale, che hanno fatto corsi lunghi, acquisito qualifiche e che hanno anche avuto azioni di accompagnamento. Sono il gruppo che ha lavorato di più (quasi il 50% in più della media), e che ha anche avuto occasioni di lavoro più numerose.

All'estremo opposto del grafico le donne straniere che hanno partecipato al FSE anche per qualifiche regionali, ma che sono ancora sotto la media sia per probabilità di lavorare, sia per numero di occasioni.

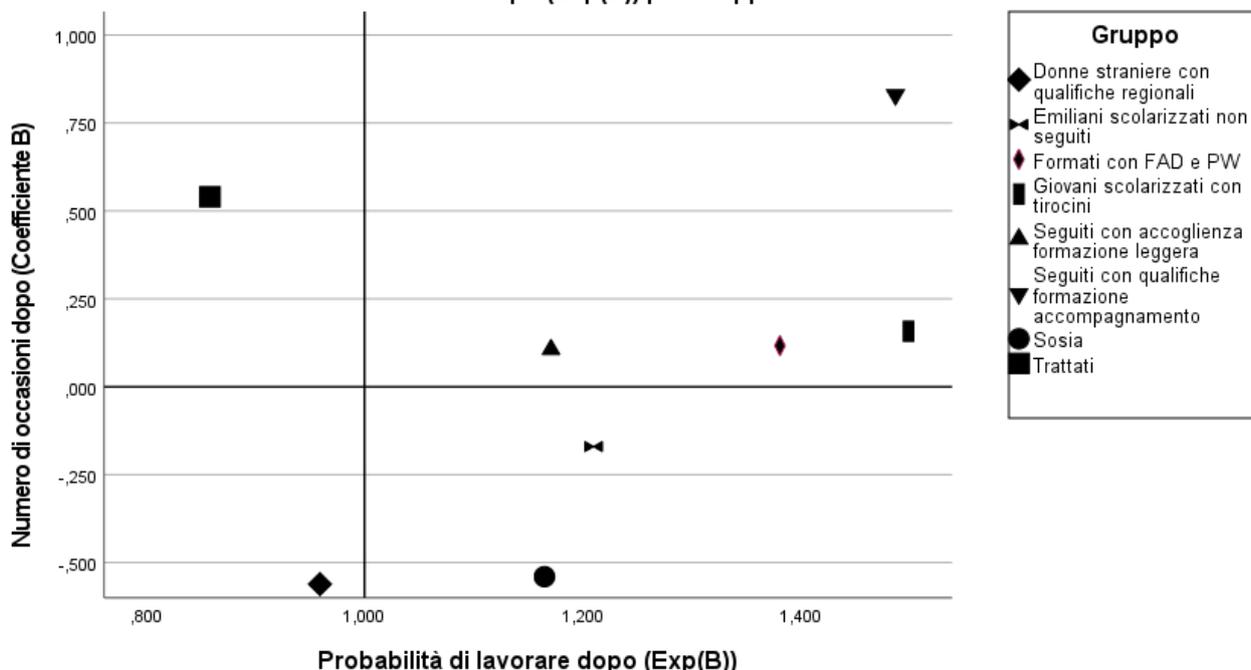
Interessante la posizione del gruppo di soggetti emiliani, a scolarità medio alta, che non è stato coinvolto nelle attività del fondo sociale. Si tratta di un gruppo di sosia indubbiamente più forte della media, che ha avuto una probabilità di lavoro superiore alla media e un numero di occasioni sotto la media. Questo gruppo costituisce un buon punto di confronto per i diversi gruppi di trattati.

Un altro gruppo interessante è quello dei trattati che sono stati coinvolti nel FSE e hanno svolto tirocini. La percentuale di chi ha lavorato è molto alta (oltre il 50% in più del totale), coerentemente con la funzione che ci attende proprio dai tirocini. Infine coloro che hanno svolto attività leggere hanno comunque avuto occasioni di lavoro sopra la media.

Pur mancando la possibilità di verificare la qualità delle occasioni, dato che stiamo lavorando sul totale delle persone che hanno partecipato al fondo sociale nel triennio, pare che la posizione sopra entrambe le medie possa ritenersi indizio solido di compensazione di fragilità.

Grafico 36

Grafico a dispersione di Numero di occasioni dopo la formazione (Coefficiente B) per Probabilità di lavorare dopo (Exp(B)) per Gruppo

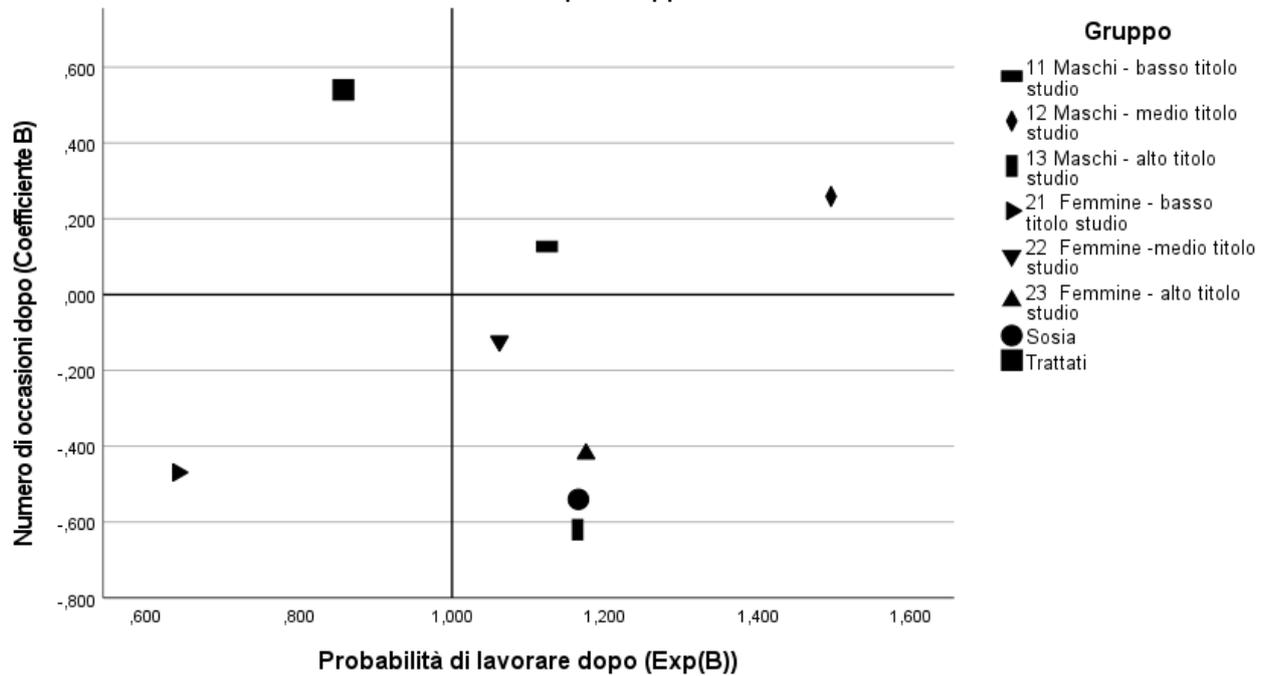


Il secondo grafico posiziona nel medesimo spazio cartesiano i gruppi che sono stati utilizzati per il PSM, divisi per sesso e fasce di istruzione. Ciascun gruppo contiene trattati e sosia e le probabilità sono sempre calcolate a parità di finestra temporale dedicata alla ricerca.

Come si può notare, il gruppo di gran lunga penalizzato è quello delle donne con basso livello di istruzione, in basso a sinistra nel grafico. I maschi a basso livello di istruzione sono di poco sopra entrambe le medie, mentre sono i maschi con istruzione intermedia che hanno la maggiore probabilità di lavoro e che come abbiamo visto sono anche un target privilegiato dei corsi più impegnativi del FSE. Maschi e femmine ad alta istruzione hanno più probabilità di lavorare della media ma un numero minore di occasioni, cosa che conferma il significato ambivalente degli avviamenti ripetuti.

Grafico 37

Grafico a dispersione di Numero di occasioni dopo (Coefficiente B) per Probabilità di lavorare dopo (Exp(B)) per Gruppo



Riepiloghiamo rapidamente i risultati di questo lavoro, che si possono ordinare su due livelli, di contenuto e di metodo.

Riguardo al contenuto possiamo dire che l'ampia platea di persone cui si rivolgono le iniziative finanziate dal FSE in Emilia Romagna sono più deboli della media sul mercato del lavoro e che per questo motivo hanno tendenzialmente meno probabilità di lavorare e meno occasioni di lavoro. Pare chiaro nei numeri, tuttavia, che esiste una capacità di compensazione, che cresce con il crescere della complessità e della dimensione delle proposte formative, che sembrano cambiare positivamente il posizionamento sul mercato delle persone coinvolte. Anche strumenti diretti al contatto con il lavoro, come i tirocini, paiono incidere con un effetto positivo.

Questo non toglie importanza al peso delle variabili strutturali come il genere, associato a una significativa debolezza, al livello di istruzione e alla provenienza, dove emergono le tradizionali segmentazioni del mercato del lavoro italiano che più di altri destina agli stranieri extracomunitari soprattutto i segmenti meno attrattivi del mercato.

Riguardo al metodo l'indagine ha dimostrato il notevole potenziale dei dati amministrativi che con pochi sforzi e bassi costi potrebbero consentire lavori rigorosi di valutazione, condotti su numeri non raggiungibili con le indagini campionarie (quasi 50mila soggetti), corrispondenti a tutti i soggetti coinvolti.

Le fonti amministrative, come è stato ricordato più volte nel corso del lavoro, hanno delle imperfezioni e spesso le estrazioni dagli archivi non sono sufficientemente allineate con le finalità delle indagini e i disegni di valutazione. Si tratta di problemi facilmente superabili se l'esercizio valutativo diventasse periodico e utilizzasse delle routine di estrazione dati e di analisi via via affinate e rese potenti per rispondere a tutte le principali domande che i soggetti che hanno responsabilità delle politiche attive si pongono. Le dimensioni longitudinale, qualitativa e settoriale, che non hanno potuto essere esplorate in questa indagine, sarebbero facilmente trattabile disponendo di tracciati più completi e di finestre temporali successive all'ingresso sul mercato del lavoro di almeno 5 anni, per seguire i processi di sviluppo e consolidamento delle posizioni e di avvio o meno di carriere professionali stabili.

Anche i problemi di riservatezza possono essere facilmente gestiti se si inquadrano all'interno di strategie e metodi studiati su misura e se si mettono sul piatto i grandi potenziali risparmi che derivano dall'usare per la valutazione dati che già esistono e dal trarre da indagini via via più affidabili indicazioni per spendere meglio quanto destinato all'investimento pubblico nelle politiche attive, a supporto del sistema produttivo regionale.