

## DELIBERAZIONE N. 6

### Il Consiglio di Amministrazione

*(Seduta del 3 novembre 2016)*

**VISTO** l'art. 18 del decreto legislativo 14/09/2015, n. 150 avente ad oggetto: "Disposizioni per il riordino della normativa in materia di servizi per il lavoro e di politiche attive, ai sensi dell'articolo 1, comma 3, della legge 10 dicembre 2014, n. 183", avente ad oggetto "Servizi e misure di politica attiva del lavoro" e che individua i servizi e le misure di politica attiva da erogare ai soggetti disoccupati allo scopo di costruire i percorsi più adeguati per l'inserimento e il reinserimento professionale nel mercato del lavoro;

**VISTA** l'articolo 9, comma 1, lettera e), del decreto legislativo 14 settembre 2015 n. 150, a norma del quale l'ANPAL provvede alla "definizione delle metodologie di profilazione degli utenti, allo scopo di determinarne il profilo personale di occupabilità, in linea con i migliori standard internazionali", nonché dei costi standard applicabili ai servizi e alle misure di cui all'articolo 18 del presente decreto;

**VISTO** l'articolo 18, comma 5, del decreto legislativo 14 settembre 2015 n. 150, secondo cui "sulla base delle informazioni fornite in sede di registrazione, gli utenti dei servizi per l'impiego vengono assegnati ad una classe di profilazione, allo scopo di valutarne il livello di occupabilità, secondo una procedura automatizzata di elaborazione dei dati in linea con i migliori standard internazionali";

**VISTO** l'art. 23, comma 1, del decreto legislativo 14 settembre 2015 n. 150, secondo cui "Ai disoccupati percettori della Nuova prestazione di Assicurazione Sociale per l'Impiego (NASpI) di cui al decreto legislativo 4 marzo 2015, n. 22, la cui durata di disoccupazione eccede i quattro mesi è riconosciuta, qualora ne facciano richiesta al centro per l'impiego presso il quale hanno stipulato il patto di servizio personalizzato di cui all'articolo 20, comma 1, ovvero mediante la procedura di cui all'articolo 20, comma 4, una somma denominata «assegno individuale di ricollocazione», graduata in funzione del profilo personale di occupabilità, spendibile presso i centri per l'impiego o presso i servizi accreditati ai sensi dell'articolo 12";

**VISTA** la nota tecnica “Il sistema di profilazione dei beneficiari dei servizi per il lavoro” inviata dall’Isfol il 28 luglio 2016;

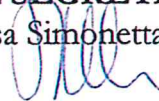
**CONSIDERATO** che la nota tecnica delinea una metodologia di profilazione automatizzata degli utenti dei servizi per l’impiego, che rispetta i requisiti previsti dalla normativa, nonché i caratteri di affidabilità e comparabilità con le metodologie adottate in altri paesi avanzati;

**DELIBERA**

di approvare la metodologia del sistema di profilazione dei beneficiari dei servizi per il lavoro, di cui alla nota tecnica “Il sistema di profilazione dei beneficiari dei servizi per il lavoro” inviata dall’Isfol il 28 luglio 2016 ed in allegato alla presente delibera.

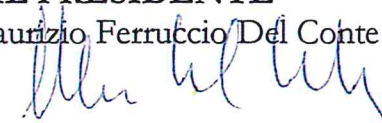
**IL SEGRETARIO**

Dott. ssa Simonetta D’Ottavio



**IL PRESIDENTE**

Prof. Maurizio Ferruccio Del Conte





Unione europea  
Fondo sociale europeo



*Ministero del Lavoro  
e delle Politiche Sociali*  
DIREZIONE GENERALE PER LE POLITICHE ATTIVE,  
I SERVIZI PER IL LAVORO E LA FORMAZIONE

# Il sistema di profilazione dei beneficiari dei servizi per il lavoro

*Nota tecnica*

Draft 28.07.2016

**ISFOL**  
ISTITUTO PER LO SVILUPPO DELLA  
FORMAZIONE PROFESSIONALE  
DEI LAVORATORI

## Premessa

**Nell'ambito delle politiche per il lavoro**, con il termine *profiling* (o profilazione) si intende lo sviluppo di una serie di tecniche e procedure per **acquisire da parte dell'offerta di servizi conoscenze approfondite sui beneficiari, effettivi o potenziali, allo scopo di migliorare l'attività di intermediazione**. La profilazione si configura pertanto come uno strumento di erogazione delle politiche di tipo preventivo, in grado di **delineare l'intensità e l'articolazione delle politiche** offerte a ciascun beneficiario. **In tale ottica l'approccio al *profiling* sottende due criteri differenti e complementari: il primo, di tipo verticale, prevede che ad un maggiore livello di svantaggio rilevato su un generico beneficiario di una politica pubblica corrisponda un maggior impiego di risorse in termini di spesa, mentre il secondo, orizzontale, considera il mix di servizi più appropriato a raggiungere l'obiettivo prefissato dalla politica erogata, determinato dal profilo specifico del beneficiario**. Nel primo caso la profilazione è finalizzata a stimare il livello di svantaggio o di disagio nel mercato del lavoro di individui appartenenti ad una specifica sottopopolazione. Il livello di svantaggio è generalmente identificato con il grado di occupabilità dei potenziali beneficiari di servizi per il lavoro. Il grado di occupabilità è assimilato alla probabilità che un individuo in cerca di occupazione possa trovare un lavoro in un determinato intervallo di tempo o alla probabilità che il periodo di disoccupazione possa protrarsi fino ad un termine prefissato, ad esempio quello che definisce la disoccupazione di lunga durata.

Le diverse finalità dell'**applicazione di tecniche di profilazione sono riconducibili nell'esperienza internazionale all'approccio anglosassone e a quello scandinavo**. Nel primo, in estrema sintesi, la finalità esplicita della politiche attive per il lavoro, e conseguentemente del sistema di profilazione, è legata alla riduzione della spesa per politiche passive (*work first*), mentre nel secondo le politiche attive sono volte ad offrire servizi mirati alla corretta collocazione occupazionale degli individui, considerando le loro caratteristiche specifiche (*human capital approach*). **In questo caso l'utilizzo della profilazione, determinando l'intensità e l'articolazione dei servizi di politica attiva più adeguati**, assicura una maggiore efficienza della spesa.

Lo sviluppo delle tecniche di *profiling* **nell'ambito dei servizi per il lavoro è nato in risposta all'esigenza di prevenire la permanenza nella disoccupazione associata alla fruizione di sussidi, allo scopo di ridurre la spesa pubblica per politiche passive**. In altra ottica il *profiling* rappresenta una chiave di passaggio dalle politiche passive alle politiche attive del lavoro (**come, ad esempio, l'assegno di ricollocazione, Dlgs 150/2015**) e una parallela modifica della destinazione della spesa verso le politiche attive, giudicate più efficienti, una volta esaurita la necessità di sostegno al reddito.

In senso più generale, una classificazione delle tecniche di *profiling* può essere declinata rispetto alla combinazione tra il livello di automatismo della procedura (*statistical, rules-based*) e l'introduzione di elementi discrezionali (*caseworker-based*), o rispetto al grado di complessità nell'utilizzo delle basi informative (*rules-based, data-only*). In molti casi l'utilizzo di un modello statistico-predittivo viene integrato con un processo di valutazione, più o meno codificato, da parte di operatori specializzati (*data-assisted*), dove la seconda componente è spesso utilizzata per catturare caratteristiche che incidono sul livello di occupabilità non altrimenti rilevabili (ad es. aspetti motivazionali, profilo specifico delle competenze, soft skills, ecc.).

**L'approccio alla profilazione può combinare in vari modi i diversi criteri e metodi disponibili: valutazione degli operatori, un insieme di regole oggettive, un sistema statistico-predittivo fondato su basi informative quantitative**. Diverse composizioni dei vari metodi suggeriscono una tassonomia sintetica e sufficientemente esaustiva dei sistemi di profilazione adottati dai vari paesi, combinando elementi discrezionali, forme di automatismo e **livello di complessità nell'utilizzo delle informazioni** (Artan Loxha, 2014).

## Le esperienze internazionali

A livello internazionale molti paesi hanno sperimentato, sviluppato e implementato in vario modo sistemi di profilazione.

Le tecniche utilizzate **presentano un'elevata variabilità tra gli approcci e nel corso degli anni si sono registrate** in alcuni casi modifiche anche rilevanti negli approcci e nei metodi utilizzati (Barnes et al., 2015); è il caso del Regno Unito e della Francia, che dopo una lunga sperimentazione di metodi strettamente quantitativi hanno optato, anche in seguito a modifiche strutturali del sistema di gestione delle politiche del lavoro, **per l'adozione di un approccio esclusivamente discrezionale.**

La gran parte dei paesi utilizza un sistema di tipo misto, affiancando al profiling statistico-predittivo la valutazione discrezionale da parte degli operatori (Kureková, 2014). Pochi paesi utilizzano esclusivamente una profilazione quantitativa, generalmente legata **all'approccio** anglosassone finalizzato alla sola riduzione della spesa per politiche passive (US). A volte **l'approccio misto prevede che** i risultati della profilazione quantitativa siano a disposizione degli operatori che possono deciderne **l'utilizzo** nella scelta delle misure da erogare (Finlandia). In altri casi, spesso dopo lunghe fasi di sperimentazione, si è optato per il solo approccio discrezionale (Germania, Francia, UK), accompagnato da un investimento rilevante sulle competenze degli operatori.

**Nell'utilizzo dell'approccio** statistico-predittivo, sia esclusivo che accompagnato da valutazione discrezionale, il ricorso a dati di origine amministrativa rappresenta un elemento costante, talvolta integrato da dati di fonte campionaria. Parallelamente, si registra in **maniera costante l'utilizzo di dati di natura longitudinale.**

La misura che identifica il livello di svantaggio, seppur circoscritta ad informazioni sul periodo di disoccupazione, è variamente declinata da paese a paese: **probabilità di rientrare al lavoro dopo un anno dall'inizio della disoccupazione (Olanda); rischio di permanenza prolungata nello stato di disoccupazione (Finlandia); rischio di disoccupazione fino all'esaurimento del diritto all'indennità (US); rischio di disoccupazione di lunga durata, oltre 6 mesi (Svezia), oltre 12 mesi (Australia, Irlanda); durata della disoccupazione (Danimarca).**

## La stima di un modello statistico-predittivo di profilazione

Va chiarito, in premessa, che la determinazione del livello di disagio di un individuo tramite la stima preventiva di un modello statistico-**predittivo comporta necessariamente un'approssimazione che sottende una serie di ipotesi.** In altre parole, il passaggio dal concetto di disagio alla sua misura opera, in primo luogo, una riduzione del concetto stesso verso i fattori che lo determinano e, successivamente, limita la misura dello svantaggio ai soli fattori osservabili. Va ricordato pertanto che le fonti di eterogeneità non osservabili rischiano di compromettere seriamente i risultati della profilazione. Tale elemento è legato alla qualità delle fonti informative disponibili e **alla scarsa propensione nel nostro Paese all'utilizzo di banche dati integrate.**

### *Approccio*

Sul piano operativo la definizione di un sistema di profilazione per stimare il livello di svantaggio nel mercato del lavoro può considerare due specificazioni: i) stima della probabilità di non trovare **un'occupazione in un determinato intervallo di tempo;** ii) stima della probabilità di trovarsi nella condizione di disoccupazione. Anche se, in senso più generale, tali misure sono associate al livello di svantaggio nel mercato del lavoro, le due definizioni sottendono due approcci differenti, mutuamente esclusivi e non complementari. Il primo prevede una definizione dinamica, dove il livello di svantaggio è correlato alla durata della permanenza nella disoccupazione; il secondo segue una definizione cross-sezionale, secondo la quale lo svantaggio è dovuto alle caratteristiche che rendono più probabile trovarsi nella condizione di disoccupazione piuttosto che in quella di occupato.

La letteratura internazionale non ha dubbi sul fatto che l'approccio longitudinale sia il più corretto. La definizione cross-sezionale considera implicitamente alla stessa stregua la disoccupazione di lunga durata e la disoccupazione di carattere frizionale: un individuo che si trova nella condizione di disoccupato ad una data specifica non presenta necessariamente un livello di disagio elevato, dal momento che brevi periodi di disoccupazione si osservano generalmente anche nelle transizioni job-to-job e nella mobilità volontaria. Il livello di disagio deve quindi essere correlato alla probabilità di permanere nella condizione di disoccupazione e non già alle caratteristiche che determinano la mera presenza nella condizione. **Nell'approccio cross-sezionale**, pur introducendo tra le specifiche dei modelli la durata della disoccupazione, le cose tendono a migliorare, ma tale accortezza non consente di modificare un approccio, in ultima analisi, concettualmente non corretto.

#### Fonti informative

Per seguire un approccio longitudinale è necessaria una struttura di dati di tipo panel: informazioni sulla popolazione (stimata o censuaria) dei disoccupati ad una data fissata (ingresso nella condizione di eleggibilità) sui quali sono rilevate le caratteristiche candidate a spiegare la permanenza nelle condizione; sono poi necessarie, sul medesimo panel di individui, informazioni sulla storia professionale, formativa e lavorativa o, quanto meno il dato sulla condizione occupazione rilevata ad una determinata distanza di tempo. La disponibilità di dati di carattere longitudinale è fortemente associata alle fonti amministrative; collezioni campionarie di dati ricavati da survey, pur se strutturati in senso longitudinale, presentano numerosi problemi (distorsioni dovute ad attrition, mancanza di correzione di incoerenze tra dati temporali, ecc.). **L'approccio dinamico impone la stima del modello su una sottopopolazione relativamente poco numerosa** (le persone in cerca di occupazione al tempo  $t_0$ ); la specifica della variabile obiettivo determina inoltre una partizione di tale popolazione tra individui usciti dalla disoccupazione e individui rimasti nella condizione **dopo uno specifico intervallo di tempo**. **L'utilizzo di una base di dati campionaria**, come nel caso italiano la rilevazione Istat sulle forze di lavoro, unita alla necessità di specificare il modello con una serie sufficiente di **covariate**, rischia di non produrre stime attendibili. **L'impiego di dati fonte amministrativa**, generalmente di carattere censuario, risolve, da un lato, i problemi di scarsa numerosità delle fonti campionarie ma, **dall'altro, presenta problemi dovuti alla** ridotta capacità informativa delle fonti derivate da registri. La soluzione, adottata largamente in sede internazionale e raccomandata da tempo anche nel nostro Paese, è **quella dell'integrazione tra banche dati amministrative**<sup>1</sup>.

#### Strumentazione

Le informazioni, acquisite da fonti campionarie o amministrative, alimentano una procedura statistica in grado di stimare la probabilità di permanere nella condizione di disoccupazione, che può essere declinata in vari modi, identificata con il livello di svantaggio. La strumentazione statistica si serve di modelli per la stima di probabilità (probit o logit) o di modelli di durata (della disoccupazione) e, **adottando l'ipotesi di una struttura gerarchica-territoriale dei dati**, di un approccio multilivello. La procedura restituisce per ciascuna caratteristica individuale (genere, età, titolo di studio, esperienza lavorativa pregressa, ecc.) un coefficiente che indica il peso che una specifica caratteristica ha nel determinare la probabilità di permanere nella condizione di disoccupato o nel determinare la durata della disoccupazione.

#### Applicazione

Una volta stimati tali coefficienti, nella parte operativa della procedura, vengono rilevate sui potenziali beneficiari dei servizi per il lavoro, ad esempio nel momento di presa in carico dei disoccupati presso i centri per **l'impiego**, **le medesime caratteristiche** utilizzate come regressori nel modello stimato e vengono applicati i rispettivi coefficienti, in modo da determinare per il singolo beneficiario il livello di svantaggio, compreso in un intervallo predeterminato (generalmente da 0 a 1) o ripartito in classi discrete.

<sup>1</sup> Una raccomandazione in tal senso è contenuta nella relazione finale della Commissione interistituzionale attivata presso il CNEL nel 2007: Ugo Trivellato, SaraLavoro: proposta per un Sistema di Archivi per Analisi sul lavoro. Il lavoro che cambia, Contributi tematici e Raccomandazioni. (<http://www.portalecnel.it/Portale/IndLavrapportiFinali.nsf/vwCapitoli?OpenView&Count=40>)

## Il sistema di profilazione dei beneficiari dei servizi per il lavoro in Italia: la proposta Isfol

### Scelte metodologiche

Nel caso italiano il metodo proposto è finalizzato ad attribuire una misura dello svantaggio sul mercato del lavoro in base al profilo dei beneficiari dei servizi per il lavoro in modo da determinare, in base a tale misura, il livello del servizio da erogare.

L'**approccio**, di tipo statistico-predittivo, prevede di osservare il profilo degli individui potenzialmente beneficiari dei servizi per il lavoro e identificare le caratteristiche che determinano differenti livelli di svantaggio, così da attribuire un peso a ciascuna caratteristica. Nella fase effettiva di presa in carico sono rilevate, su ciascun individuo, le medesime caratteristiche e, attribuendo il sistema di pesi identificato in precedenza, viene determinata la misura di svantaggio del singolo individuo.

Nella fase esplorativa **per l'identificazione del metodo da utilizzare**, sono state sperimentate due classi di modelli, cross-sezionale e longitudinale. **Il primo è stato utilizzato in continuità con l'esperienza del Programma Garanzia Giovani**, mentre il secondo rappresenta un accostamento agli standard internazionali. Alcuni elementi critici della classe di modelli cross-sezionale sono risultati attenuati nel modello stimato per il programma Garanzia Giovani, dal momento che la popolazione eleggibile era caratterizzata da un relativamente elevato grado di omogeneità (giovani NEET in età compresa tra 15 e 29 anni). La medesima classe di modelli applicata **sull'intera popolazione si è dimostrato meno adatto nel rispondere agli obiettivi della profilazione: il più elevato grado di eterogeneità dell'intera popolazione in età lavorativa non ha consentito di ottenere risultati adeguati**, data la specificazione voluta del modello. Per tale ragione e a seguito di varie sperimentazioni **si è optato per l'adozione di un modello longitudinale**.

La popolazione eleggibile è costituita dagli individui in cerca di occupazione all'anno  $t_0$  e con età compresa tra 15 e 64 anni.

**La base dati utilizzata per la stima del modello proviene dall'indagine Continua delle Forze di Lavoro dell'Istat**. Il particolare disegno della rilevazione consente di costruire una struttura longitudinale (Centra et al. 2001), in modo da produrre stime **sulle transizioni tra forme contrattuali e sull'ammontare e il profilo degli ingressi al lavoro nel corso di un anno**. In particolare è stato costruito un dataset pooled unendo le componenti panel RCFL 2011-2012 2012-2013 2013-2014 2014-2015. La base dati così costruita si compone di 39.573 osservazioni<sup>2</sup>.

La strumentazione econometrica utilizzata per determinare il peso che specifiche caratteristiche hanno sul livello di svantaggio degli individui potenzialmente beneficiari di servizi per il lavoro è stata determinata da due fattori: la variabile dipendente è di tipo dicotomico; le caratteristiche che incidono sulla propensione di trovarsi in condizioni di svantaggio sono sia individuali che territoriali.

Il primo fattore ha imposto **l'utilizzo di un modello logit** (Agresti, 1996; McCullagh, Nelder, 1989), adeguato a stimare la probabilità di sperimentare un evento in funzione di specifiche caratteristiche. Il secondo **fattore suggerisce l'utilizzo di un modello multilevel**; le unità su cui si rileva il fenomeno (unità di primo livello) risultano naturalmente aggregate in gruppi differenti (unità di secondo livello, nel caso specifico le regioni). In presenza di unità statistiche annidate in unità di livello gerarchicamente superiore, le unità appartenenti a uno stesso gruppo sono soggette a comportamenti simili a causa delle caratteristiche del **contesto che le raggruppa**. In altri termini, **l'ipotesi alla base è che le caratteristiche socio-economiche territoriali siano una sovrastruttura che condiziona sia la probabilità che un individuo disoccupato in un determinato istate temporale non sia più nella stessa condizione a distanza di un anno sia le caratteristiche individuali degli individui**. La misura del livello di svantaggio deve pertanto considerare la elevata **eterogeneità territoriale del fenomeno dell'occupabilità, evitando** il rischio di errori di specificazione, anche

---

<sup>2</sup> Nella stima del modello è stato utilizzato il coefficiente di riporto all'universo fornito dall'indagine, adattato alla struttura pooled e normalizzato.

elevati, con conseguenti errori di attribuzione del punteggio ai beneficiari, dal momento che una **determinata caratteristica (il genere, l'età, il titolo di studio, ecc.) può avere effetti molti diversi sull'occupabilità in contesti territoriali differenti.**

In ragione di ciò, la tecnica di profilazione si fondava sulla stima di un modello multilevel logistico a intercetta casuale, che rientra nella classe dei modelli di tipo gerarchico (Goldstein, 1995; Snijders e Bosker, 1999).

La variabile dipendente, ossia l'informazione scelta per definire il livello di disagio, è dunque una variabile dicotomica che si valorizza ad 1 nel caso in cui un **individuo in cerca di occupazione nell'anno t0 risulti non occupato anche nell'anno t1<sup>3</sup>**.

I predittori inseriti nel modello, ossia le caratteristiche che incidono sulla propensione di trovarsi in condizioni di svantaggio, di tipo sia individuale che territoriale, sono di seguito elencate:

*Schema 1 Le covariate inserite nel modello di profilazione*

Variabili individuali	Modalità	
Genere	Donna	
	Uomo	
Età in anni compiuti	Variabile continua	
Cittadinanza	Paesi EU	
	Italiana	
	Paesi Extra EU	
Titolo di studio più elevato conseguito	Fino alla licenza media	
	Qualifica professionale	
	Diploma	Istituto professionale
		Istituto tecnico
		Liceo
		Altro
	Laurea (diploma o triennale)	Scienze umanistiche
		Scienze sociali
		Scienze della salute
		Ingegneria, informatica e trasporti
		Altro
	Laurea (magistrale o vecchio ord.)	Scienze umanistiche
		Scienze sociali
		Scienze della salute
		Ingegneria, informatica e trasporti
		Scienze naturali
		Architettura
	Condizione professionale anno precedente	Altro inattivo
		Occupato
		In cerca di nuova occupazione
		In cerca di prima occupazione
	Durata della disoccupazione (in mesi)	Variabile continua
	Iscritto a corso di istruzione scolastico, universitario, IFP, IFTS, ITS	Si, No
Anno t0 (pooled)	2011	
	2012	
	2013	
	2014	
<b>Variabili territoriali</b>		
Tasso di occupazione 15-29 anni	Variabili continue	
Variazione del tasso di disoccupazione		
Incidenza delle famiglie a bassa intensità di lavoro		
Densità imprenditoriale		

<sup>3</sup> A differenza della gran parte dei modelli adottati a livello internazionali non è stato possibile utilizzare come variabile dipendente la durata della disoccupazione (o una variabile da questa derivata, es. probabilità di diventare disoccupato di lunga durata) a causa della particolare struttura del dataset derivato dai dati della rilevazione Istat FDL. Nello specifico nel dataset non sono disponibili informazioni riferite al periodo compreso tra i due momenti di rilevazione, elemento che impedisce di **ricostruire in maniera affidabile l'informazione sulla durata della disoccupazione.**



Nel modello è stata inserita la variabile relativa all'anno di riferimento del dataset pooled, con l'obiettivo di cogliere la componente congiunturale: i coefficienti stimati per i diversi anni sembrano riprodurre in larga misura il tasso di incremento del numero di persone in cerca di occupazione.

La selezione dei predittori da utilizzare a livello territoriale è stata condotta in seguito ad un'analisi esplorativa sui fattori che incidono sulla differenza del livello medio regionale di svantaggio e guidata dai criteri di parsimonia.

La procedura di stima adottata si compone di due passi:

1) La stima del modello nullo, vale a dire la misura della probabilità di restare non occupati a distanza di un anno rispetto alla sola regione. In questa fase è stato possibile verificare se la strumentazione utilizzata si adattasse al caso in esame. Il risultato ha confermato l'elevata variabilità del fenomeno tra le regioni, confermando la necessità di utilizzare un modello gerarchico al fine di considerare l'effetto che medesime caratteristiche hanno in regioni diverse.

2) La stima del modello completo. Alla specificazione del modello nullo sono state aggiunte le covariate che si sono dimostrate essere validi predittori della probabilità di restare non occupati a distanza di un anno.

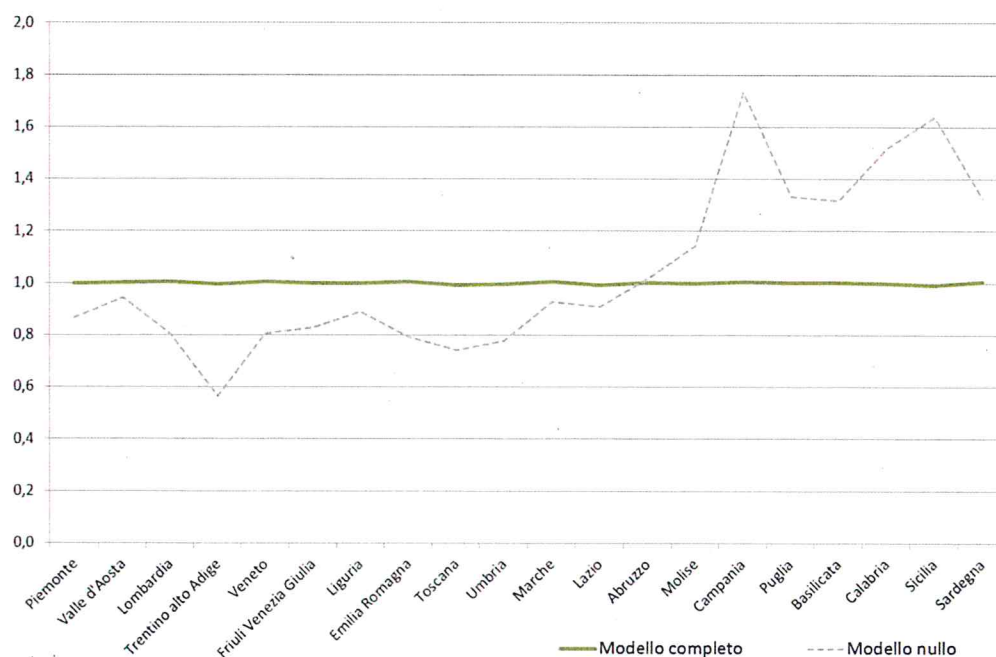
I parametri stimati del modello di regressione multilivello forniscono i pesi di ciascuna caratteristica individuale e territoriale che contribuisce a determinare la probabilità di restare non occupati a distanza di un anno vale a dire la misura del livello di svantaggio.

### I risultati

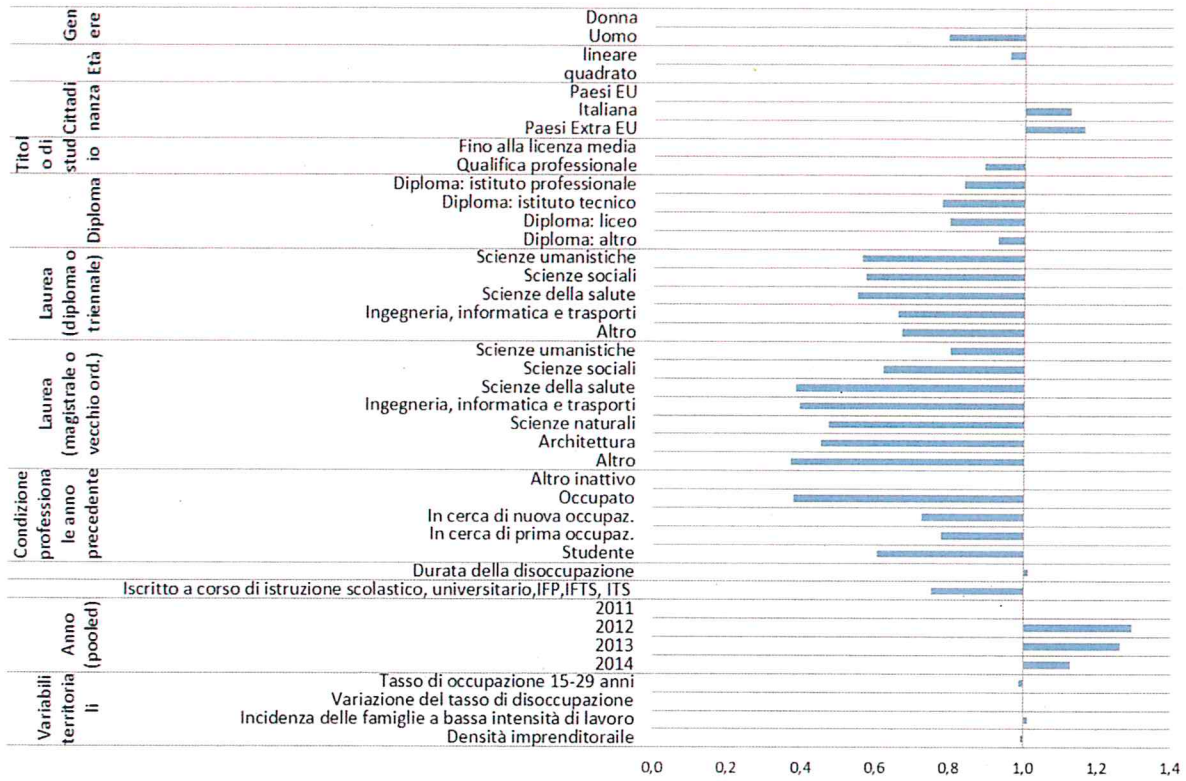
Di seguito sono riportati i risultati in ordine alle stime ottenute (effetti di primo e secondo livello) e alla distribuzione della probabilità stimata sulla popolazione eleggibile.

Il modello stimato mostra di cogliere in misura adeguata la componente di cluster (le regioni), fattore verosimilmente riconducibile ad un più elevato effetto di omogeneità dei mercati del lavoro regionali se riferiti all'intera popolazione.

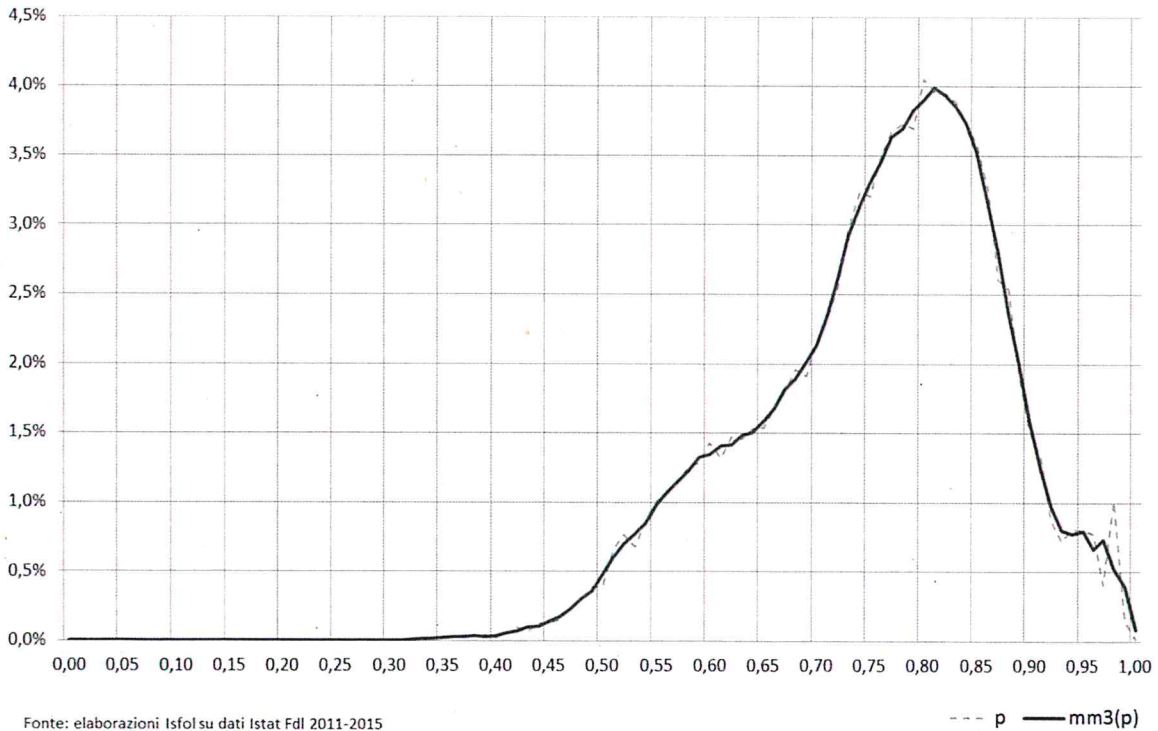
Stime degli effetti casuali di secondo livello: modello nullo e modello completo



Stime dei parametri del modello: odds ratio<sup>4</sup>



Distribuzione della probabilità stimata sulla popolazione eleggibile (media mobile di ordine 3)



<sup>4</sup> Il modello completo è riportato in appendice

## Validazione del modello

Al fine di validare il modello, assicurandone la robustezza e la capacità di rappresentare il fenomeno in maniera adeguata, si ritiene opportuno sviluppare alcuni test.

Sfruttando i dati di fonte amministrativa, ottenuti tramite integrazione di banche dati provenienti da Inps e dal MLPS (Archivio percettori di prestazioni di sostegno al reddito, Sistema statistico delle comunicazioni pobbliatorie, ecc.), si intende sottoporre il modello a test verificando se esso è effettivamente predittivo del livello di svantaggio.

La fase di test, in estrema sintesi, intende ricostruire tramite le fonti amministrative la popolazione eleggibile ad un determinato istante temporale e verificare se la frequenza effettivamente osservata dopo 12 mesi, date le caratteristiche del soggetto preso in esame, coincide statisticamente con la probabilità di rimanere in condizione di ricerca di occupazione a distanza di un anno stimata dal modello.

La fase di test è necessario al fine di un corretto utilizzo della strumentazione econometrica proposta e **della buona riuscita dell'azione di policy.**

## Riferimenti bibliografici

- Barnes S.A., Wright S., Irving P., Deganis I., (2015), Identification of latest trends and current developments in methods to profile jobseekers in European Public Employment Services: Final report, European Commission, Directorate-General for Employment, Social Affairs and Inclusion.
- Centra M., Discenza A.R., Rustichelli E., (2001), Strumenti per le analisi di flusso nel mercato del lavoro. Una procedura per la ricostruzione della struttura longitudinale della rilevazione trimestrale Istat sulle forze di lavoro, Isfol, Monografie sul Mercato del lavoro e le politiche per l'impiego, n. 2, 2001.
- Goldstein H. (1995), Multilevel Statistical Model, Edward Arnold, London.
- ISFOL, (2015), Garanzia Giovani: Prime analisi sui giovani presi in carico e sul sistema di profilazione degli utenti. <http://www.isfol.it/isfol-europa/garanzia-giovani/approfondimenti-tematici/prime-analisi-sui-giovani-presi-in-carico-e-sul-sistema-di-profilazione-degli-utenti>
- Kureková L.M., (2014), Review of Profiling Systems, Categorization of Jobseekers and Calculation of Unit Service Costs in Employment Services - Implications and Applications for Slovakia, Nezisková organizácia EPIC, Panenska 29, 811 03 Bratislava.
- Laboratorio Ida Rossi, Prova Progetto Valutazione (2014), Come capire (meglio) chi rischia di restare a lungo **senza lavoro? Risultati di un'analisi di profiling condotta nella provincia di Cuneo**, Nota di sintesi - Marzo 2014
- Loxha A., Advanced Profiling of Unemployed in Public Employment Services. A Critical Review of OECD Experiences and applications for Western Balkans. Vienna, March 4, 2014.
- McCullagh P., Nelder I. A. (1989), Generalized linear models, Chapman and Hail, London.
- Snijders T., Bosker R. (1999), An Introduction to Basic and Advanced Multilevel Modelling, Sage, London.
- Weinert P., Baukens M., Bollérot P., Pineschi-Gapègne M., Walwei U., (eds), 2001, Employability: from theory to practice, International social security series, vol. 7, Transaction Publishers, New Brunswick (USA).

## Appendice

**Modello: stima della probabilità delle persone in cerca di occupazione di permanere in condizione di non occupazione a distanza di un anno**

probabilità che un individuo disoccupato a t0 sia disoccupato a t1

Specifiche:

popolazione 15-64 anni che a t0 è disoccupata

sono inclusi a t0:

2 persone in cerca, con precedenti esperienze, ex-occupati

3 persone in cerca, con precedenti esperienze, ex-inattivi

4 persone in cerca, senza precedenti esperienze

DATI PANEL: RCFL 2011-2012 2012-2013 2013-2014 2014-2015

DATI PESATI

Numero osservazioni

39.573

**MODELLO NULLO**

Criterio di convergenza (PCONV=1.11022E-8) soddisfatto.

Statistiche di stima						
-2 Res Log Pseudo-Likelihood	195103.8					
Chi-quadrato generalizzato	39521					
Chi-quadrato gener. / DF	1					
Stime dei parametri di covarianza						
Param cov	Soggetto	Stima	Errore standard	Valore Z	Pr > Z	
Intercept	v_reg	0,09661	0,03331	2,9	0,0019	
Soluzioni per effetti fissi						
Effetto	Stima	Errore standard	Valore t	Pr >  t	Inferiore	Superiore
Intercept	1,2256	0,07238	16,93	<.0001	1,0741	1,3771
Soluzione per gli effetti casuali						
Soggetto	Stima	Err std prev	Valore t	Pr >  t	Inferiore	Superiore
Piemonte	-0,1443	0,08278	-1,74	0,0813	-0,3066	0,01796
Valle d'Aosta	-0,0586	0,2153	-0,27	0,7856	-0,4806	0,3635
Lombardia	-0,2145	0,07866	-2,73	0,0064	-0,3687	-0,0603
Trentino alto Adige	-0,5699	0,1224	-4,65	<.0001	-0,8099	-0,3299
Veneto	-0,2169	0,08576	-2,53	0,0114	-0,385	-0,0488
Friuli Venezia Giulia	-0,1863	0,1143	-1,63	0,1031	-0,4103	0,03771
Liguria	-0,1174	0,1036	-1,13	0,2572	-0,3204	0,08566
Emilia Romagna	-0,2332	0,08526	-2,74	0,0062	-0,4003	-0,0661
Toscana	-0,2988	0,08639	-3,46	0,0005	-0,4681	-0,1294
Umbria	-0,2514	0,1126	-2,23	0,0256	-0,472	-0,0307
Marche	-0,0739	0,1005	-0,74	0,4623	-0,2708	0,1231
Lazio	-0,0973	0,08016	-1,21	0,225	-0,2544	0,05986
Abruzzo	0,0189	0,1037	0,18	0,8553	-0,1843	0,2221
Molise	0,1353	0,1584	0,85	0,3929	-0,1751	0,4457
Campania	0,5507	0,08116	6,79	<.0001	0,3916	0,7098
Puglia	0,2874	0,08273	3,47	0,0005	0,1253	0,4496
Basilicata	0,2747	0,1342	2,05	0,0408	0,01156	0,5378
Calabria	0,4189	0,09253	4,53	<.0001	0,2376	0,6003
Sicilia	0,493	0,08219	6	<.0001	0,3319	0,6541
Sardegna	0,2833	0,09408	3,01	0,0026	0,09892	0,4677

**MODELLO COMPLETO**

Criterio di convergenza (PCONV=1.11022E-8) soddisfatto.

Statistiche di stima						
-2 Res Log Pseudo-Likelihood	202793.3					
Chi-quadrato generalizzato	41672.5					
Chi-quadrato gener. / DF	1,05					
Stime dei parametri di covarianza						
Param cov	Soggetto	Stima	Errore standard	Valore Z	Pr > Z	
Intercept	v_reg	0,00031	0,00117	0,26	0,3964	
Soluzioni per effetti fissi						
Caratteristiche a t0	Stima	Errore standard	Valore t	Pr >  t	Inferiore	Superiore
Intercetta	3,15750	0,38350	8,2	<.0001	2,349	3,967
Uomo	-0,23150	0,02707	-8,6	<.0001	-0,285	-0,179
Età in anni compiuti	-0,04204	0,00884	-4,8	<.0001	-0,059	-0,025
Età in anni compiuti al quadrato	0,00069	0,00011	6,1	<.0001	0,000	0,001
<b>Cittadinanza (base= Paesi EU)</b>						
Italiana	0,11850	0,06299	1,9	0,0600	-0,005	0,242
Paesi Extra EU	0,15090	0,07415	2,0	0,0418	0,006	0,296
<b>Titolo di studio (base=fino alla licenza media)</b>						
Qualifica professionale	-0,11530	0,05099	-2,3	0,0238	-0,215	-0,015
Diploma: istituto professionale	-0,17700	0,05193	-3,4	0,0007	-0,279	-0,075
Diploma: istituto tecnico	-0,25390	0,03562	-7,1	<.0001	-0,324	-0,184
Diploma: liceo	-0,22550	0,05018	-4,5	<.0001	-0,324	-0,127
Diploma: altro	-0,07485	0,07739	-1,0	0,334	-0,227	0,077
Laurea (diploma o triennale): Scienze umanistiche	-0,57680	0,12070	-4,8	<.0001	-0,814	-0,340
Laurea (diploma o triennale): Scienze sociali	-0,55980	0,10880	-5,2	<.0001	-0,773	-0,347
Laurea (diploma o triennale): Scienze della salute	-0,60020	0,14610	-4,1	<.0001	-0,887	-0,314
Laurea (diploma o triennale): Ingegneria, informatica e trasporti	-0,41610	0,20850	-2,0	0,0460	-0,825	-0,007
Laurea (diploma o triennale): Altro	-0,40290	0,16310	-2,5	0,0135	-0,723	-0,083
Laurea (magistrale, specialistica, vecchio ordinamento): Scienze umanistiche	-0,22410	0,08974	-2,5	0,0125	-0,400	-0,048
Laurea (magistrale, specialistica, vecchio ordinamento): Scienze sociali	-0,48030	0,06782	-7,1	<.0001	-0,613	-0,347
Laurea (magistrale, specialistica, vecchio ordinamento): Scienze della salute	-0,95890	0,15540	-6,2	<.0001	-1,263	-0,654
Laurea (magistrale, specialistica, vecchio ordinamento): Ingegneria, informatica e trasporti	-0,92810	0,12660	-7,3	<.0001	-1,176	-0,680
Laurea (magistrale, specialistica, vecchio ordinamento): Scienze naturali	-0,75210	0,15060	-5,0	<.0001	-1,047	-0,457
Laurea (magistrale, specialistica, vecchio ordinamento): Architettura	-0,79680	0,15270	-5,2	<.0001	-1,096	-0,498
Laurea (magistrale, specialistica, vecchio ordinamento): Altro	-0,99190	0,27730	-3,6	0,0003	-1,535	-0,448
<b>Condizione professionale anno precedente (base= Altro inattivo)</b>						
Cond. anno prec: Occupato	-0,97300	0,06461	-15,1	<.0001	-1,100	-0,846
Cond. anno prec: In cerca di nuova occupaz.	-0,32150	0,06219	-5,2	<.0001	-0,443	-0,200
Cond. anno prec: In cerca di prima occupaz.	-0,25110	0,07156	-3,5	0,0004	-0,391	-0,111
Cond. anno prec: Studente	-0,50640	0,07994	-6,3	<.0001	-0,663	-0,350
<b>Durata della disoccupazione</b>	0,01320	0,00075	17,7	<.0001	0,012	0,015
<b>Iscritto a corso di istruzione scolastico, universitario,IFP,IFTTS, ITS</b>	-0,28870	0,07515	-3,8	0,0001	-0,436	-0,141
<b>Anno t0 (Anno=2011)</b>						
2012	0,25970	0,04058	6,4	<.0001	0,180	0,339
2013	0,23400	0,04533	5,2	<.0001	0,145	0,323
2014	0,12080	0,05087	2,4	0,0176	0,021	0,221
<b>Tasso di occupazione 15-29 anni*</b>	-0,01064	0,00322	-3,3	0,0010	-0,017	-0,004
<b>Variazione del tasso di disoccupazione**</b>	0,00297	0,00097	3,1	0,0021	0,001	0,005
<b>Incidenza delle famiglie a bassa intensità di lavoro***</b>	0,01121	0,00555	2,0	0,0433	0,000	0,022
<b>Densità imprenditoriale****</b>	-0,00650	0,00373	-1,7	0,0816	-0,014	0,001

\* A livello provinciale a t0

\*\*Variazione relativa annua a livello provinciale

\*\*\*Indicatore a livello regionale Eurostat : "People living in households with very low work intensity " --> The indicator is defined as the number of persons living in a household having a work intensity below a threshold set at 0.20. The work intensity of a household is the ratio of the total number of months that all working-age household members have worked during the income reference year and the total number of months the same household members theoretically could have worked in the same period. A working-age person is a person aged 18-59 years, with the exclusion of students in the age group between 18 and 24 years.

\*\*\*\*Istat: Imprese per 1000 abitanti (anno 2012)

Soluzione per gli effetti casuali						
Soggetto	Stima	Err std prev	Valore t	Pr >  t	Inferiore	Superiore
Piemonte	-0.0025	0.01663	-0.15	0.8808	-0.0351	0.03011
Valle d'Aosta	0.00105	0.01752	0.06	0.9521	-0.0333	0.0354
Lombardia	0.00379	0.01618	0.23	0.8148	-0.0279	0.0355
Trentino alto Adige	-0.0053	0.01737	-0.31	0.7586	-0.0394	0.02871
Veneto	0.00547	0.01672	0.33	0.7433	-0.0273	0.03824
Friuli Venezia Giulia	-0.0008	0.01735	-0.04	0.9656	-0.0348	0.03326
Liguria	-0.0016	0.0172	-0.09	0.9269	-0.0353	0.03214
Emilia Romagna	0.00615	0.01671	0.37	0.7128	-0.0266	0.03891
Toscana	-0.0092	0.01712	-0.54	0.5895	-0.0428	0.02432
Umbria	-0.0034	0.0173	-0.2	0.8432	-0.0373	0.03048
Marche	0.00457	0.01715	0.27	0.7901	-0.0291	0.03819
Lazio	-0.0069	0.01647	-0.42	0.6746	-0.0392	0.02537
Abruzzo	0.00256	0.01719	0.15	0.8816	-0.0311	0.03624
Molise	0.00023	0.01747	0.01	0.9894	-0.034	0.03447
Campania	0.00586	0.01656	0.35	0.7234	-0.0266	0.03831
Puglia	0.00213	0.01656	0.13	0.8975	-0.0303	0.03459
Basilicata	0.00094	0.01741	0.05	0.9572	-0.0332	0.03505
Calabria	-9E-05	0.01702	-0.01	0.9956	-0.0335	0.03326
Sicilia	-0.0066	0.01684	-0.39	0.6974	-0.0396	0.02646
Sardegna	0.00362	0.01701	0.21	0.8315	-0.0297	0.03695