

# **Forze di Lavoro**

## **Dati individuali con abbinamenti a 12 mesi di distanza**

### **1. Nota introduttiva**

La rilevazione trimestrale sulle forze di lavoro (RTFL) rappresenta una fonte informativa preziosa per le analisi sulla mobilità del mercato del lavoro e per la costruzione di modelli microeconomici volti a stimare le probabilità di transizione tra i diversi stati (occupato, disoccupato, inattivo). La RTFL incorpora nel suo disegno campionario una struttura longitudinale derivante dal sistema di rotazione delle famiglie nei campioni trimestrali che prevede, infatti, la reintervista degli individui a distanza di 3, 12 e 15 mesi dalla prima intervista. Teoricamente, in base a questo schema, la metà degli individui intervistati sono gli stessi a un trimestre e a un anno di distanza dalla prima occasione di rilevazione.

Nella RTFL, tuttavia, la ricostruzione della struttura longitudinale è ostacolata dall'assenza di un codice che permetta di identificare con assoluta certezza il soggetto intervistato in occasioni di indagine diverse. La funzione di riconoscimento del medesimo individuo deve pertanto essere demandata ad alcune variabili chiave che sono soggette a errori di misura, in sede di rilevazione o di registrazione dei dati. Ai fini dell'abbinamento è stata utilizzata una procedura messa a punto in collaborazione con l'Università di Padova che consente l'abbinamento di record con chiavi parzialmente errate. Tale procedura massimizza il numero di unità abbinate a condizione che la probabilità di agganciare individui che in realtà sono diversi sia al di sotto di una certa soglia prefissata<sup>1</sup>.

Nell'utilizzare la componente longitudinale è inoltre fondamentale tenere conto del fatto che non si tratta di un vero e proprio panel perché un individuo, intervistato la prima volta in uno dei comuni campione, non viene reintervistato se nell'arco di tempo tra la prima e la successiva intervista ha cambiato residenza o si è trasferito all'estero. La componente longitudinale non rappresenta, quindi, tutta la popolazione, ma solo quella residente in uno stesso comune sia all'inizio sia alla fine del periodo considerato. Nei file longitudinali qui diffusi sono presenti i record della sola popolazione longitudinale di almeno 15 anni di età a inizio periodo. E' da sottolineare infine che i pesi di riporto all'universo "longitudinali" sono ottenuti in modo da assicurare la coerenza con le più importanti stime trasversali correntemente diffuse e relative alla popolazione complessiva della RTFL<sup>2</sup>.

---

<sup>1</sup> Per approfondimenti si veda in allegato:

“Le matrici di transizione della Rilevazione trimestrale sulle forze di lavoro. Nota metodologica”

a cura di C. Ceccarelli, A.R. Discenza, A. Paggiaro, S. Rosati e N. Torelli – ISTAT – Roma 12 dicembre 2002.

<sup>2</sup> Per approfondimenti si veda in allegato:

“La mobilità nel mercato del lavoro: principali risultati (Aprile 1998 – aprile 2002)”

a cura di M. Albisinni e A.R. Discenza – ISTAT Approfondimenti – Roma 12 dicembre 2002.

Tutti i file di dati elementari abbinati (file longitudinali) sono forniti in formato ascii, hanno lo stesso tracciato record (cfr. in allegato: “Forze di Lavoro: Tracciato record dei dati abbinati a 12 mesi di distanza. Aprile 1993-aprile 2002 – File STANDARD.rtf”) e la seguente struttura :

### Struttura file STANDARD

TRIMESTRE 1												TRIMESTRE 2	CONCORD	COEPANEL	
VARIABILI DI BASE INDIVIDUALI										VARIABILI DI ANALISI	VARIABILI DI ANALISI				
ANNO	TRIMESTRE	REGIONE	N. COMPONENTI LA FAMIGLIA P/90	RELAZIONE CON L'INTESTATARIO DEL FOGLIO DI FAMIGLIA	SESSO	ETA' IN ANNI COMPIUTI	STATO CIVILE	MOTIVO DELL' ASSENZA DAL COMUNE DI RESIDENZA	TITOLO DI STUDIO PIU' ELEVATO CONSEGUITO	ANALISI DEI TITOLI DI STUDIO SUPERIORI PIU' ELEVATI DELLA LICENZA MEDIA INFERIORI	NUMERO D'ORDINE DELLA PERSONA NEL FOGLIO INDIVIDUALE	Variabili del trimestre 1 in formato standard	Variabili del trimestre 2 in formato standard (escluso titolo di studio, relazione con intestatario, regione, ripartizione, sesso)	Identificatore per il tipo di abbinamento effettuato	coefficiente per il riporto alla popolazione longitudinale

All'interno di ciascun file di dati longitudinali sono incluse le informazioni relative ai due trimestri considerati riguardanti gli individui abbinati. Per ciascun individuo si dispone di variabili:

- identificative;
- di analisi relative al trimestre iniziale;
- di analisi relative al trimestre finale;

sempre per ciascun individuo è presente l'indicatore del tipo di abbinamento effettuato e il coefficiente di riporto.

Tutte le variabili vengono fornite con modalità analoghe a quelle del file standard trimestrale dei dati trasversali, vale a dire con oscuramento di alcuni campi per motivi di riservatezza e con riagggregazione di alcune modalità di risposta per le variabili a codifica complessa (settore di attività economica, professione, ecc.).

In considerazione dei cambiamenti intervenuti nei questionari dal 1993 al 2002 alcune variabili non sono sempre disponibili per tutti i periodi considerati e altre, invece, potrebbero essere codificate in modo diverso tra un anno e l'altro. Per l'individuazione delle variabili presenti in ciascun file (sia per il trimestre iniziale che per il trimestre finale) occorre fare riferimento all'allegato “Forze di Lavoro: Tracciato record dei dati abbinati a 12 mesi di distanza. Aprile 1993-aprile 2002 – File STANDARD.rtf”.

Nei file è presente il peso di riporto all'universo longitudinale vincolato alla popolazione longitudinale che consente di calcolare le matrici di transizione.

E' importante precisare che la componente longitudinale basata sui dati della RTFL descrive solo i flussi tra le diverse condizioni intervenuti per la popolazione longitudinale, mentre i campioni trasversali della RTFL forniscono una stima della distribuzione per condizione della popolazione iniziale e finale. Esiste una quota di popolazione (gli usciti dalla popolazione iniziale e gli entrati nella popolazione finale) per i quali non è possibile stimare i flussi tra condizione. Di questa parte, che non viene rappresentata dalla componente longitudinale, si possono comunque ottenere informazioni parziali per differenza tra i dati di stock trasversali e i dati di stock relativi alla sola popolazione longitudinale. In particolare per quella parte di popolazione che muore, cambia residenza o emigra all'estero si può stimare la condizione a inizio periodo. Specularmente, di quella parte della popolazione che si è iscritta in anagrafe o ha compiuto 15 anni nel periodo sotto osservazione si può stimare la condizione a fine periodo.

Nel calcolare queste stime occorre fare molta attenzione. Le stime relative alla popolazione longitudinale sono ottenute dal campione longitudinale che è un sottoinsieme (più che dimezzato) dei campioni trasversali dei due trimestri messi a confronto. In teoria, le stime di tutti gli aggregati relativi alla popolazione longitudinale devono essere al più uguali a quelle relative alla popolazione complessiva essendo la prima un sottoinsieme della seconda. Potrebbe accadere, però, soprattutto per gli aggregati di modesta entità, che tale condizione teorica non sia rispettata. Ciò è essenzialmente dovuto al fatto che le stime relative a gruppi di contenuta consistenza hanno un elevato errore campionario. Tale errore è ancora più alto per le stesse stime ottenute dal campione longitudinale. Può accadere quindi che le stime puntuali riferite alla popolazione longitudinale possono risultare leggermente superiori alle corrispondenti stime per la popolazione complessiva pur avendo intervalli di confidenza sovrapposti. Essendo impossibile tenere sotto controllo questa eventualità, la procedura di calcolo dei pesi di riporto è studiata per minimizzare il rischio di incoerenze. In linea di massima la coerenza viene garantita per i principali domini territoriali/demografici e sottogruppi di popolazione.