

Nota metodologica

1. Introduzione

La Rilevazione campionaria sulle forze di lavoro costituisce la principale fonte statistica sul mercato del lavoro italiano. Da essa derivano le stime ufficiali degli occupati e delle persone in cerca di lavoro. Il suo utilizzo per analisi di tipo sia congiunturale sia strutturale è quanto mai ampio: l'evoluzione dei principali indicatori del mercato del lavoro può essere studiata in modo disaggregato a livello territoriale, settoriale e per le principali caratteristiche sociodemografiche della popolazione.

La Rilevazione campionaria sulle forze di lavoro è stata profondamente ristrutturata a partire dal 2004 negli aspetti contenutistici, definatori, tecnici e organizzativi. La ristrutturazione, dettata in primo luogo dalla necessità di adeguare l'indagine ai nuovi standard comunitari, vuole anche rispondere alle accresciute esigenze conoscitive nonché all'obiettivo di miglioramento della qualità della rilevazione.¹ In confronto al passato, i principali cambiamenti introdotti nel 2004 riguardano la periodicità dell'indagine, non più svolta in una specifica settimana per ciascun trimestre ma distribuita su tutte le settimane dell'anno; i criteri di classificazione degli individui secondo la condizione professionale; l'ampliamento dei contenuti informativi; il ricorso a una rete di rilevazione professionale, composta da oltre 300 rilevatori; l'utilizzo di tecniche *computer-assisted* in sostituzione dei tradizionali modelli di rilevazione cartacei, in particolare le tecniche Capi (*Computer assisted personal interview*) e Cati (*Computer assisted telephone interview*); l'attivazione di un complesso sistema informativo-informatico che supporta lo svolgersi dell'indagine, dal monitoraggio della qualità del lavoro sul campo all'acquisizione per via telematica dei nominativi delle famiglie dai comuni del campione.²

In occasione dell'ultima ristrutturazione l'Istat ha ricostruito le serie storiche dei principali indicatori del mercato del lavoro per raccordare i risultati ottenuti attraverso la precedente metodologia con quelli della nuova rilevazione avviata nel 2004. Il raccordo, effettuato con una tecnica statistico-econometrica, ha riguardato inizialmente il periodo IV trimestre 1992 – IV trimestre 2003. Un successivo lavoro ha consentito di disporre di serie storiche ricostruite fin dal 1977 per i principali aggregati³. Tali ricostruzioni, nonostante abbiano riguardato un elevato numero di variabili, non arrivano al livello di dettaglio informativo delle tavole di media annua e quindi per alcune delle informazioni ivi contenute non esistono termini di paragone omogenei per gli anni precedenti al 2004. Il 2 marzo 2015, in occasione del rilascio dei dati di media del 2014, sono stati diffusi i dati ricostruiti mensili, trimestrali e annuali dal 2002 al 2014, tenendo conto della ricostruzione statistica delle serie di popolazione a seguito delle risultanze del Censimento della popolazione e delle abitazioni del 2011, già diffusa dall'Istituto il 14 gennaio 2015.

Nella lettura dei dati, infine, è bene prestare attenzione al livello di errore statistico dovuto alla natura campionaria della rilevazione, che può essere sensibile per gli aggregati meno consistenti e per i livelli di

¹ Con riferimento agli standard comunitari, si veda in particolare quanto riportato nel regolamento (Ce) n. 577/98 del Consiglio del 9 marzo 1998 relativo all'organizzazione dell'indagine campionaria sulle forze di lavoro nella Comunità, e nel regolamento (Eu) n. 545/2014 del Parlamento europeo e del Consiglio del 15 maggio 2014.

² Per ulteriori dettagli si veda Istat: *La rilevazione sulle forze di lavoro – Contenuti, metodologie, organizzazione*. Istat: Roma, 2006. (Metodi e norme, n. 32) disponibile all'indirizzo Internet http://www.istat.it/dati/catalogo/20060830_00/.

³ Tali dati sono consultabili sul *datawarehouse* I.Stat.

dettaglio più elevati. Per la valutazione del livello di precisione delle stime fino al livello provinciale si consiglia di consultare il seguente link (<http://www.istat.it/it/archivio/210086>) e accedere alle tavole provinciali.

2. Universo di riferimento

L'universo di riferimento dell'indagine è costituito da tutti i componenti delle famiglie residenti in Italia, anche se temporaneamente all'estero. Sono escluse le famiglie residenti in Italia che vivono abitualmente all'estero e i membri permanenti delle convivenze (ospizi, brefotrofi, istituti religiosi, caserme, eccetera). La popolazione residente comprende le persone, di cittadinanza italiana e straniera, che risultano iscritte alle anagrafi comunali.

3. Unità di rilevazione e di analisi

L'unità di rilevazione è la famiglia di fatto. Questa va intesa come un insieme di persone legate da vincoli di matrimonio, parentela, affinità, adozione, tutela o da vincoli affettivi, coabitanti e aventi dimora abituale nello stesso comune, anche se non residenti nello stesso domicilio. Una famiglia può essere costituita anche da una sola persona. Un questionario individuale viene posto a ciascun componente della famiglia di fatto. Unità di analisi sono sia gli individui sia le famiglie.

4. Periodicità e riferimenti temporali

A partire dal 2004, la rilevazione è divenuta continua in quanto le informazioni sono rilevate con riferimento a tutte le settimane dell'anno tenuto conto di un'opportuna distribuzione del campione complessivo nelle 13 settimane di ciascun trimestre.

I risultati dell'indagine, elaborati e diffusi con cadenza trimestrale, restituiscono in tal modo un'immagine media del trimestre piuttosto che relativa a un preciso istante temporale.

A partire da ottobre 2009 è iniziata anche la diffusione delle stime mensili dei principali indicatori del mercato del lavoro, rivolte a offrire una più tempestiva informazione. Tali stime, prodotte con un'opportuna metodologia statistica, sono diffuse a distanza di circa 30 giorni dalla fine di ciascun mese di riferimento.

In generale, le informazioni raccolte si riferiscono a una settimana, detta *di riferimento*, che va dal lunedì alla domenica successiva. L'età degli intervistati viene considerata in relazione alla domenica della settimana di riferimento. Alcune informazioni (ad esempio, stato civile, livello di istruzione, cittadinanza) sono invece rilevate con riferimento al momento in cui viene svolta l'intervista. Quest'ultima viene generalmente condotta nel corso della settimana *di rilevazione*, quella seguente alla settimana di riferimento. D'altro canto, qualora non sia possibile effettuare l'intervista nella settimana *di rilevazione*, si può condurla in una delle cosiddette settimane *di coda*, ovvero in una delle quattro settimane seguenti quella di rilevazione.

5. Disegno campionario

Il campione utilizzato è stato progettato per assicurare prefissati livelli di precisione delle stime trimestrali a livello regionale e provinciali in media d'anno. Il campione è a due stadi, rispettivamente comuni e famiglie, con stratificazione delle unità di primo stadio. Le unità di primo stadio sono stratificate all'interno di ciascuna provincia sulla base della dimensione demografica dei comuni. In seguito a tale stratificazione i comuni vengono suddivisi in autorappresentativi e non autorappresentativi. I primi, costituiti da quelli con popolazione superiore a una soglia prestabilita che varia da provincia a provincia, sono inclusi nel campione con certezza; i secondi, la cui popolazione è al di sotto delle soglie prefissate, sono invece raggruppati in strati omogenei per provincia in modo da ottenere livelli costanti di popolazione complessiva. Essi entrano nel campione attraverso

un meccanismo di selezione casuale che prevede l'estrazione di un comune da ciascuno strato con probabilità proporzionale al peso demografico di ciascun comune nello strato. I comuni campione non autorappresentativi vengono sostituiti quando non sono più in grado di fornire nuove famiglie campione. Per ciascun comune viene estratto dalla lista anagrafica un campione casuale semplice di famiglie. Nell'anno 2017 in ciascun trimestre sono state intervistate circa 65 mila famiglie (circa 150 mila individui) residenti in 1.410 comuni coinvolti nella rilevazione, distribuiti in tutte le province del territorio nazionale.

A ciascuna famiglia del campione, detta famiglia base, sono associate tre famiglie sostitute a cui si ricorre nel caso in cui la famiglia base non voglia o non possa partecipare alla rilevazione. Ciascun gruppo di quattro famiglie, composto dalla famiglia base e da tre famiglie sostitute, è chiamato *quartina*.

L'estrazione delle famiglie campione dalle liste anagrafiche dei comuni viene effettuata con cadenza annuale, nel mese di marzo, in modo da assicurare la formazione dei campioni per quattro rilevazioni consecutive, dal terzo trimestre di un anno al secondo trimestre dell'anno successivo. Il campione delle unità di secondo stadio è caratterizzato da una struttura longitudinale che segue uno schema di rotazione del tipo 2-2-2 per cui ogni famiglia viene intervistata per due trimestri successivi, esce temporaneamente dal campione per due rilevazioni, vi rientra per due ulteriori tornate per poi uscire definitivamente. Tale sistema di rotazione consente di mantenere invariata metà della composizione del campione in due trimestri consecutivi e in trimestri a distanza di un anno l'uno dall'altro. La scelta del sistema di rotazione è tale da conciliare in maniera ottimale le esigenze di costruzione di stime di "livello" e di "variazione": quanto maggiore è il numero di famiglie che si rinnovano di periodo in periodo, tanto maggiore è l'affidabilità delle stime di livello; parimenti, quanto più consistente è la quota di famiglie in comune da una rilevazione all'altra, tanto più risultano stabili le stime relative alle differenze tra periodi successivi (stime di variazione). Tale schema permette altresì di effettuare analisi longitudinali e di flusso, importanti per uno studio della dinamica del mercato del lavoro.

Le stime si ottengono assegnando a ciascuna osservazione campionaria un peso, chiamato *coefficiente di riporto all'universo*, che esprime il numero delle unità della popolazione rappresentate dall'unità campionaria in questione. Per il calcolo dei coefficienti di riporto all'universo viene utilizzato uno stimatore di ponderazione vincolata. I pesi sono ottenuti risolvendo un problema di minimo vincolato in cui le stime campionarie di alcune variabili ausiliarie sono poste uguali ai corrispondenti totali noti da fonti esterne all'indagine. Le informazioni esterne utilizzate fanno riferimento essenzialmente alla distribuzione della popolazione per sesso e classi di età a diversi livelli di dettaglio territoriale (regioni, province, grandi comuni).

La popolazione di riferimento di ciascun trimestre si ottiene dalla popolazione residente (aggiornata alla data più prossima al trimestre di riferimento dell'indagine), al netto dei membri permanenti delle convivenze.

Avvertenze

Simboli convenzionali

Nelle tavole statistiche sono state adoperate le seguenti convenzioni:

Linea (-) a) quando il fenomeno non esiste; b) quando il fenomeno esiste e viene rilevato, ma i casi non si sono verificati.

Due puntini (..) per i numeri che non raggiungono la metà della cifra dell'ordine minimo considerato.

Arrotondamenti

Per effetto degli arrotondamenti in migliaia operati direttamente dall'elaboratore, i dati delle tavole possono non coincidere tra loro per qualche unità in più o in meno. Per lo stesso motivo, non sempre è stato possibile realizzare la quadratura verticale e orizzontale nell'ambito della stessa tavola.

Ripartizioni geografiche

NORD

Piemonte, Valle d'Aosta/Vallée d'Aoste, Lombardia, Liguria (*Nord-ovest*),

Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna (*Nord-est*);

CENTRO

Toscana, Umbria, Marche, Lazio;

MEZZOGIORNO

Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna.