

Nota metodologica

Caratteristiche e contenuti informativi della rilevazione

Nel quadro del Sistema europeo delle statistiche su retribuzioni e costo del lavoro, la rilevazione sulla struttura delle retribuzioni (usualmente definita con la sigla SES, *Structural Earning Statistics*) è una indagine quadriennale che ha l'obiettivo di monitorare la struttura e la distribuzione delle retribuzioni negli Stati membri dell'Unione Europea.

La rilevazione SES è disciplinata dal Regolamento del consiglio 530/99, dal Regolamento attuativo della commissione 1738/2005 e dal Regolamento della commissione 698/06 sulla valutazione della qualità. La SES è stata anche inserita nel Regolamento della commissione 831/02 sul rilascio dei microdati per scopi scientifici.

La struttura dell'indagine sulle retribuzioni si riferisce all'anno 2006, per le variabili annuali, e a Ottobre 2006 per i dati mensili.¹

La SES relativa all'anno 2006 ha come universo di riferimento le imprese/enti e istituzioni con almeno 10 dipendenti attive nell'industria, nei servizi orientati al mercato e nei servizi sociali e personali (escluso il settore dell'Agricoltura, silvicoltura e pesca e della Amministrazione Pubblica). Per il ciclo 2006 si è proceduto ad un trattamento metodologico differenziato per il settore privato e per quello a parziale controllo pubblico rispetto al settore pubblico.

Nel settore privato vengono rilevate tutte le imprese con almeno 250 dipendenti, nel settore pubblico la metodologia, di seguito presentata, non effettua selezioni a priori sulla dimensione dell'ente/istituzione.

Nel settore privato sono state rilevate 6.015 imprese (a cui corrispondono a 14.200 unità locali) e 137.210 dipendenti. Nel settore pubblico le stime finali riguardano 5.699 unità locali e 17.816 dipendenti.

Le informazioni fornite dalla rilevazione, basata su criteri armonizzati, permettono di misurare l'influenza della struttura occupazionale e dell'orario di lavoro sui differenziali retributivi nazionali (territoriali, settoriali, di genere, etc.) in funzione di alcune caratteristiche personali e professionali dei singoli dipendenti. Le stime derivate dalla SES consentono di analizzare sia gli aspetti collegati al "capitale umano" in termini di abilità e conoscenze (tramite caratteristiche quali l'età, il livello di istruzione e formazione acquisito e l'anzianità aziendale), sia quelli strettamente correlati all'inserimento nel posto di lavoro (come la professione, il tipo e la durata del contratto di lavoro).²

In particolare, le informazioni obbligatorie a livello di unità locali riguardano:

- l'attività economica, a livello di due cifre della classificazione Nace Rev1.1 (corrispondente alla Ateco 2002);
- la localizzazione geografica, secondo la classificazione NUTS1 (ripartizioni territoriali);
- la dimensione dell'unità locale e dell'impresa (ente o istituzione);
- il tipo di controllo economico e finanziario;
- l'esistenza di contratti collettivi di lavoro.

Quelle relative ai dipendenti sono :

- genere ed età;
- professione, secondo la classificazione ISCO88 a una cifra;
- livello di istruzione e formazione personale più elevato raggiunto, secondo la classificazione ISCED;
- anzianità di servizio nella struttura dove il lavoratore è inserito;
- durata del tempo di lavoro (a tempo pieno o a tempo parziale) e tipo di contratto (a tempo indeterminato, determinato o apprendistato);
- numero di ore retribuite e di straordinario nel mese di ottobre e nell'anno;
- numero di settimane a cui si riferisce la retribuzione annuale, numero di giorni di assenza;
- retribuzione lorda, per straordinari e per indennità di turno nel mese di ottobre;
- retribuzioni lorde annuali.

¹ Si è scelto il mese di ottobre in quanto ritenuto, nei diversi Paesi, quello in cui si rilevano i minori tassi di assenza per malattia e non sono presenti feste nazionali che riducano il numero dei giorni lavorati. L'Eurostat lascia comunque facoltà ai Paesi di scegliere un altro mese per la rilevazione qualora il mese di ottobre non abbia tali caratteristiche.

² L'Eurostat acquisisce direttamente i microdati, che vengono diffusi su New Cronos http://epp.eurostat.ec.europa.eu/portal/page/portal/labour_market/earnings/database, in maniera aggregata secondo modalità concordate con i Paesi prima dell'avvio della rilevazione. I microdati sono anche diffusi, previa l'anonimizzazione dei record per i quali non è possibile garantire la confidenzialità del dato individuale., sia dal Safe Center dell'Eurostat sia dall'Istat come "File per la ricerca".

Gli indicatori prodotti, in particolare quelli retributivi e sull'orario del lavoro, incorporano per ciascun paese gli effetti delle condizioni del mercato del lavoro nazionale (grado di flessibilità, assetti istituzionali e del *welfare*), come pure la composizione dell'occupazione (per settore di attività economica e posizione di lavoro) o la specifica distribuzione per qualifica o per titolo di studio.

In particolare, la SES fornisce l'indicatore del *gender pay gap*. Sulla base delle stime del 2006, a livello europeo si è stabilito di fornire annualmente tale informazione, anche negli anni non coperti dalla rilevazione, secondo una metodologia stabilita all'interno del gruppo di lavoro su *Labour Market Statistics* (LAMAS) dell'Eurostat. A ogni Paese è stata lasciata facoltà di selezionare le fonti nazionali da utilizzare per aggiornare l'indicatore negli anni non coperti dalla rilevazione. L'Istat utilizza i dati forniti dalla rilevazione annuale sulle famiglie Eu-Silc.

Settore privato

L'impostazione metodologica della rilevazione garantisce i confronti, a parità di attività economica con i precedenti cicli di rilevazione, pur presentando elementi di innovazione sia rispetto all'insieme di riferimento, sia nel processo di produzione statistica.

Rispetto al ciclo del 2002, nel settore privato, oltre all'ampliamento del campo di rilevazione al settore dei servizi sociali e personali, vi è stato un miglioramento della rappresentatività statistica dei dipendenti per i quali si richiedono le informazioni individuali, mantenendo l'obiettivo di fornire indicatori a livello di impresa (come nel 2002) e di unità locale (come nel 1996).

L'universo di riferimento della rilevazione è rappresentato dall'archivio ASIA 2005 delle imprese con almeno 10 dipendenti.

Il piano di campionamento è a due stadi: il primo è costituito dalle imprese stratificate per settori di attività economica (a due cifre della classificazione Nace Rev.1.1), per sette classi di dipendenti (10-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1000 e oltre) e per le cinque ripartizioni territoriali (Nord-ovest, Nord-est, Centro, Sud e Isole). Il secondo stadio è rappresentato da un campione di dipendenti la cui numerosità dipende dalla dimensione dell'impresa in cui è occupato.

Le imprese con almeno 250 addetti sono tutte comprese nel campione.

La numerosità campionaria delle imprese da 10 a 249 dipendenti, risultante dalle tecniche di allocazione multivariata adottate, è stata sovradimensionata sulla base dei tassi di caduta di risposta osservati nella rilevazione SES del 2002, per le attività economiche da C a K, e nella rilevazione sulla struttura del costo del lavoro (RCL) del 2004, per le attività economiche M, N e O.

Le imprese della parte campionaria della rilevazione sono state selezionate secondo il campionamento di Pareto, un caso particolare di selezione a probabilità variabile.³

Prospetto 1 - Numerosità campionaria di secondo stadio per classi di dipendenti dell'impresa

| classe di dipendenti | numero di dipendenti da campionare |
|----------------------|---------------------------------------|
| da 10 a 19 | tutti |
| da 20 a 49 | 20 |
| da 50 a 99 | 25 |
| da 100 a 249 | 35 |
| da 250 a 499 | 40 |
| da 500 a 999 | 50 |
| da 1000 a 1999 | 60 |
| da 2000 a 3999 | 65 |
| da 4000 a 7499 | 75 |
| da 7500 a 9999 | 100 |
| 10000 e oltre | 200 |

³ Falorsi P.D., Ballin M., De Vitiis C., Scepi G., «Principi e metodo del software generalizzato per la definizione del disegno di campionamento nelle indagini sulle imprese condotte dall'Istat». *Statistica Applicata*, Vol 10, n.2 (1998).

Le numerosità campionarie di secondo stadio, cioè il numero di dipendenti da campionare all'interno di ciascuna impresa, sono state prefissate secondo la tipologia d'impresa definita da classi di ampiezza, adattando le indicazioni suggerite da Eurostat⁴ all'esperienza maturata nel ciclo 2002. Il Prospetto 1 riporta la numerosità campionaria di secondo stadio secondo la dimensione dell'impresa.

Per le imprese con meno di 250 dipendenti la selezione dei dipendenti è stata effettuata dalle imprese in modo casuale, secondo un passo di campionamento fornito dall'Istat, affinché la selezione avvenisse in modo sistematico, con probabilità uguali e senza re-immissione.

Per le imprese di maggiore dimensione, sono stati utilizzati, in via sperimentale⁵, i dati ricavati dall'archivio delle dichiarazioni fiscali dei modelli 770 dell'anno 2005, riferiti all'anno d'imposta 2004 che hanno consentito sia la stratificazione del campione di secondo stadio, sia la selezione guidata dei dipendenti da campionare. In particolare, il patrimonio informativo delle dichiarazioni 770 è stato utilizzato per il 93 per cento delle 3.383 grandi imprese considerate: si tratta di quelle per le quali la partita Iva (o codice fiscale) ha trovato l'aggancio con il corrispettivo dei sostituti d'imposta presente nell'archivio dei modelli fiscali 770.

Attraverso i codici fiscali (o partita Iva) dei percipienti, i dipendenti dell'impresa sono stati stratificati per sei classi di età (<20, 20-30, 30-40, 40-50, 50-60, >60 anni) e per genere, variabili che presentano un elevato livello di correlazione con i livelli retributivi percepiti dai dipendenti.

Alle grandi imprese sono stati forniti i codici fiscali dei dipendenti cui riferire le informazioni, lasciando la facoltà di non utilizzarli e procedere ad una selezione casuale dei dipendenti-campione come per le imprese di minore dimensione. Inoltre, su esplicita richiesta, l'impresa ha avuto a disposizione le numerosità definite per strato, al fine di poter effettuare le sostituzioni del dipendente-campione nei seguenti casi: cessazione del rapporto di lavoro, assenza di ore effettivamente lavorate dal dipendente nel mese di ottobre, assunzione del dipendente successiva al 31 dicembre 2005, errore nel codice fiscale trasmesso dall'Istat.

Complessivamente il campione finale di imprese è risultato pari a 16.553 imprese corrispondenti a circa 370.000 dipendenti attesi.

La rilevazione è condotta mediante autocompilazione di questionari inviati per posta alle imprese selezionate, con solleciti telefonici e postali volti a garantire una buona copertura del campione. Inoltre, alle imprese è stata data, per la prima volta, l'opportunità di compilare un questionario elettronico, scaricabile via web con password fornita individualmente.

Il questionario contiene una sezione dedicata ai dati riferiti all'impresa nel suo complesso e una per rilevare le caratteristiche socio-economiche di ciascun dipendente campionato. Al fine di garantire la tutela del segreto statistico e della privacy dell'impresa, i codici fiscali dei dipendenti da campionare sono trasmessi alle grandi imprese esclusivamente nella copia elettronica, o inviati alle imprese che hanno preferito compilare il modello cartaceo.

Il tasso di risposta è stato complessivamente pari al 36,3 per cento in termini di imprese, rappresentative del 45,7 per cento dei dipendenti totali attesi.

Il campione finale rispondente comprende 6.015 imprese e 137.210 dipendenti ed è risultato sufficiente ad assicurare la significatività statistica dei risultati, dato l'iniziale sovracampionamento del 30 per cento. In questo modo si è corrisposto ai severi requisiti di qualità statistica richiesti dal Regolamento comunitario.

Per le variabili soggette a regolamento è stata adottata una procedura di controllo e correzione con integrazione delle mancate risposte parziali. In particolare si sono adottati metodi basati sulle tecniche del "miglior donatore" e sull'integrazione con valori mediani calcolati sui rispondenti rispetto a sub-domini di quelli di stima finale.

Tale procedura si basa sul confronto con i dati elementari desunti direttamente dalle rilevazioni Istat sul Sistema dei conti delle imprese (per quelle con almeno 100 addetti) e sulle Piccole e medie imprese (per quelle con meno di 100 addetti) e dall'archivio Oros dell'Istat, che contiene i dati a livello di singola impresa, derivanti dal modello DM10 dell'INPS. L'archivio Oros ha permesso il confronto puntuale dei livelli di retribuzione lorda e di oneri sociali a livello di impresa relativamente a: numero di dipendenti, dipendenti per qualifica professionale e tipo di orario di lavoro (tempo pieno o tempo parziale).

⁴ Doc. Eurostat/E1/Sal/5.2/02(Room document) PHARE 1999 Multi-Country Programme: Pilot Project on Labour Cost and Earnings Statistics- Model SES Questionnaires: One for Local Units and One for Enterprises -prepared by Christophe Demunter- January 2002.

⁵ La sperimentazione è stata condotta nell'ambito del Gruppo di Lavoro Istat MEF/Dipartimento delle politiche fiscali per l'utilizzo a fini statistici dell'archivio dei Modelli 770.

Per integrare le mancate risposte relative alle variabili titolo di studio e professione (secondo la classificazione ISCO88) è stata utilizzata una tecnica probabilistica che individua il “miglior donatore”, utilizzando il software RIDA (Ricostruzione informazioni con donatore automatica) sviluppato presso l’Istat. La tecnica si basa sul calcolo della distanza minima ponderata tra due unità rispetto ad un numero di variabili stabilite. La distanza minima subisce successive correzioni con un fattore di penalità che evita che lo stesso donatore possa essere utilizzato più volte, in modo tale da assicurare al massimo la varianza originaria delle variabili.

Il riporto dei dati all’universo è stato effettuato con l’ausilio della procedura solitamente utilizzata per le indagini campionarie sulle imprese, definendo i pesi a livello di dipendente, di impresa e, a differenza del 2002, anche per unità locale.

La metodologia adottata si basa sullo stimatore di ponderazione vincolata e si avvale delle informazioni strutturali ausiliarie fornite dall’archivio delle imprese di riferimento (ASIA 2006) per la stima dei pesi finali da assegnare alle unità rispondenti del primo e del secondo stadio campionario (imprese e dipendenti). Si assume che le variabili oggetto di indagine siano strettamente correlate con quelle strutturali ausiliarie. Inoltre, è garantito il rispetto di uguaglianza tra i “totali noti” delle variabili ausiliarie e le stime delle stesse, introducendo fattori correttivi dei pesi così calcolati.⁶

Le variabili ausiliarie utilizzate sono costituite dal numero di imprese e dal numero di dipendenti.

I domini di stima sono stati definiti a livello di sottosezione di attività economica (25 modalità), classe di dipendenti (10-19; 20-49; 50-99; 100-249; 250-499; 500 e oltre) e le 5 macroregioni della classificazione NUTS1.

Per quanto concerne il peso dell’unità locale, il modello di rilevazione ha registrato, a livello di singolo dipendente, le informazioni relative all’unità locale dove effettivamente lavora: numero di dipendenti medi, attività economica, ripartizione geografica e provincia di appartenenza. Sulla base di queste informazioni è stato effettuato l’aggancio con l’universo delle Unità Locali del 2006 (Istat), individuando 14.200 unità locali di riferimento. La probabilità di estrarre la singola unità locale è stata legata alla probabilità di estrarre almeno un suo dipendente.

Infine, per ciascun peso (impresa, unità locale e dipendente) sono stati definiti i rispettivi fattori correttivi di calibrazione che, come per i fattori correttivi sulle mancate risposte, si basano su un’ipotesi di stretta correlazione tra variabile oggetto di indagine e le variabili strutturali ausiliarie.

Settore Pubblico

L’estensione della rilevazione al settore pubblico, si è basata per i settori dell’Istruzione e della Sanità, che rappresentano la quota più consistente del settore, sull’utilizzo congiunto di fonti amministrative, fiscali e statistiche, senza ricorrere a indagine diretta sugli enti/istituzioni pubbliche. Questa integrazione di informazioni provenienti da fonti diverse ha avuto l’obiettivo di assicurare i livelli di qualità delle stime, richieste da Eurostat, per la prima volta a livello di singolo dipendente, rispetto alle quali non esistono riferimenti statistici di controllo desumibili da altre fonti ufficiali.

Attraverso l’analisi dei contenuti informativi degli archivi amministrativi e statistici disponibili a livello nazionale, è stato predisposto un complesso lavoro di armonizzazione delle classificazioni e definizioni presenti nelle diverse fonti. Un risultato di rilievo è stato quello di verificare il livello di copertura delle informazioni rese disponibili dalle diverse fonti rispetto alle richieste del Regolamento europeo.

Per l’integrazione delle variabili definite obbligatorie da Eurostat e non presenti nelle fonti amministrative si sono utilizzate le informazioni provenienti da due rilevazioni dirette condotte dall’Istat presso le famiglie: la Rilevazione Eu-Silc (*Statistics on Income and Living Conditions*) riferita al 2006 e la Rilevazione sulle forze di lavoro riferita al quarto trimestre 2006.

Per l’Istruzione le fonti amministrative e fiscali utilizzate sono rappresentate da archivi amministrativi provenienti da tre diversi Enti: archivio dei modelli fiscali 770 anno di riferimento 2006 (fonte: Ministero Economia e Finanze, MEF); cedolini mensili stipendiali di ottobre 2006 (fonte: MEF, Sistema SPT); liste dei dirigenti, del personale educativo di ruolo e non di ruolo, docenti dell’infanzia, scuola primaria, scuole medie e

⁶ Tali fattori correttivi tengono conto della mancata risposta totale dovuta all’impossibilità di intervistare alcune imprese selezionate per non disponibilità delle stesse o per irrisolvibilità.

superiori per l'anno scolastico 2005-2006 (fonte: Ministero della Istruzione dell'Università e della Ricerca); liste del personale docente e non docente universitario del 2006 (fonte: Consorzio interuniversitario Cineca).

Per la Sanità le fonti sono costituite dalla lista delle unità istituzionali appartenenti al settore delle Amministrazioni Pubbliche S13 (compilata dall'Istat) e l'archivio dei modelli fiscali 770 per l'anno di riferimento 2005 (fonte: Agenzia delle Entrate).

Secondo quanto richiesto dal Regolamento europeo, per l'istruzione è stato possibile individuare l'universo delle unità locali delle scuole (effettuando successivamente la selezione di quelle con almeno 10 dipendenti) e dell'organico effettivamente presente nel mese di ottobre 2006. Si noti che questo mese si caratterizza come quello di avvio delle attività scolastiche e quindi registra, rispetto alla media annuale del settore, una minore presenza dei "docenti supplenti", come pure dei contratti "a tempo determinato", che sono invece più diffusi nei mesi successivi di assestamento dell'organico scolastico⁷.

Per la Sanità, in assenza di informazioni statisticamente controllate relative alle unità locali, si è fatto riferimento all'organico medio nell'anno e all'insieme delle unità locali afferenti alla stessa struttura sanitaria.

Come si evince dal precedente elenco di fonti considerate, per la Sanità si rileva una minore disponibilità di informazioni rispetto all'Istruzione: il primo comparto è caratterizzato da una struttura di governo regionale, il secondo ha una gestione centralizzata dei dati amministrativi e fiscali (MIUR e MEF). La differenza nell'anno di riferimento dei dati dell'archivio dei modelli fiscali 770 (2005 per il primo, 2006 per il secondo) è strettamente dipendente dalle diversa tempestività dell'insieme delle fonti disponibili, che deve essere considerato come un vincolo strutturale di cui si tiene conto nella procedura di stima utilizzata⁸.

Pur partendo da input differenti, i due universi di riferimento per l'Istruzione e la Sanità pubblica sono stati definiti tramite un approccio statistico-metodologico simile, atto a garantire i livelli di qualità delle stime e un uniforme trattamento dei dati.

Il passaggio da archivi amministrativi alla definizione degli universi di riferimento è sostanzialmente il risultato di un processo ETL (*Extract, Transform, Load*), basato su una stessa chiave di aggancio, rappresentata dalla partita Iva per le persone giuridiche e dal codice fiscale per le persone fisiche, con conseguente individuazione univoca e diretta delle unità di analisi rappresentate dai lavoratori dipendenti.

L'integrazione delle informazioni provenienti da archivi amministrativi con quelle provenienti dalle rilevazioni presso le famiglie ha implicato ulteriori passaggi con una accurata omogeneizzazione delle classificazioni, definizioni e nomenclature utilizzate. Inoltre, poiché la Rilevazione sulle forze di lavoro ed Eu-Silc non hanno in comune una stessa chiave di aggancio, è stata necessaria la costruzione di una procedura di record linkage basata sulla ricostruzione dei codici fiscali per i dipendenti derivati dalla Rilevazione sulle forze di lavoro⁹ e alla validazione dei codici fiscali dei dipendenti Eu-Silc. Per quest'ultima rilevazione, i codici fiscali sono stati ricostruiti in fase di check e si è proceduto unicamente alla verifica della congruenza tra attività economica rilevata da indagine e quella risultante dall'aggancio con l'archivio dei modelli fiscali 770.

Individuati gli universi di riferimento totali, sono stati esclusi, come previsto dal Regolamento dell'Ue, i dipendenti occupati in strutture con meno di 10 dipendenti. Per la definizione della dimensione occupazionale, nell'Istruzione è stata introdotta una distinzione tra Scuola e Università. Per la Scuola "l'unità locale" è rappresentata dall'insieme delle sedi afferenti alla stessa struttura scolastica e la stime della dimensione occupazionale è stata resa disponibile dalle liste del MIUR (in particolare la Sezione "Anagrafe del personale"), mentre per l'Università la dimensione è rappresentata dai dipendenti della struttura universitaria complessiva. Un approccio analogo a quello relativo alle università è stato adottato per il settore sanitario: in questo settore "l'unità locale" è rappresentata dalle singole sedi in cui è articolata la struttura sanitaria rilevata nella lista S13 ma le informazioni disponibili non permettono la distinzione del personale dipendente a livello di "unità locale".

⁷ La minore presenza di dipendenti con contratti di lavoro "flessibili" ha un effetto anche sui livelli delle retribuzioni che incorporano un'anzianità maggiore e quindi livelli complessivamente più elevati. Inoltre per il settore della Scuola si fa riferimento alle ore contrattuali assegnate per livello di insegnamento (infanzia, primario, secondario inferiore e superiore), inferiori alle ore effettivamente lavorate stimate nel resto dell'economia. Ciò comporta indicatori della retribuzione oraria relativamente più elevati.

⁸ Per la Sanità i dati retributivi relativi al 2005 sono stati aggiornati in fase di stime finali considerando sia i dati di Contabilità Nazionale dell'Istat sia i dati relativi al Conto Annuale della Ragioneria Generale di Stato del MEF.

⁹ La ricostruzione del codice fiscale è il risultato di un processo di normalizzazione applicato alla sezione Dati Anagrafici del modello di Rilevazione sulle forze di lavoro.

Il processo di selezione sulla dimensione occupazionale ha determinato la definizione dei due universi di riferimento di dipendenti peranti in strutture pubbliche con almeno 10 dipendenti, pari a 1.207.720 nell'Istruzione e 717.561 nella Sanità.

La numerosità campionaria è stata definita della considerando una proporzione fissa dell'1 per cento negli strati; questi sono definiti dall'incrocio di sei classi di retribuzioni¹⁰, sei classi di età¹¹, genere e professione; quest'ultima modalità riguarda solo l'istruzione per la quale sono disponibili cinque classi della ISCO88 a una cifra¹².

In ciascuno strato sono stati selezionati tutti i dipendenti i cui codici fiscali erano presenti contemporaneamente nell'universo di riferimento e nelle Rilevazioni sulle forze di lavoro o Eu-Silc (quindi con un approccio di campionamento ragionato); la restante numerosità campionaria, in ogni strato, è stata raggiunta estraendo casualmente altri codici fiscali.

Per il settore dell'istruzione il campione è risultato composto da 10.791 dipendenti, dei quali 7.565 presenti nella Rilevazione sulle forze di lavoro o in Eu-Silc; nel settore Sanità invece il campione include 7.053 dipendenti, di cui 4.394 rilevati anche dalla Rilevazione sulle forze di lavoro o da Eu-Silc.

La composizione dei campioni estratti massimizza quindi l'aggancio tra l'universo di riferimento dei due settori pubblici e i campioni rilevati nelle due rilevazioni Istat.

La valutazione e la correzione delle informazioni campionarie ha utilizzato le stesse tecniche e software applicati nel controllo di quelle relative al settore privato, con alcuni adattamenti. In particolare, la correzione è stata subordinata gerarchicamente alle tre variabili cardine della SES: retribuzione del mese di ottobre (esclusivamente per l'Istruzione), retribuzione annuale e ore effettivamente lavorate.

Per le variabili non individuate nelle fonti utilizzate (la percentuale di part-time il numero di giorni di ferie usufruite e le retribuzioni mensili per turni lavoro festivo e/o notturno) sono state approntate alcune procedure di stima basata sulle relazioni intercorse tra esse e le variabili di cui si aveva informazione. Le variabili derivabili dalle due rilevazioni Istat e non presenti in tutti i record appartenenti al campione sono state considerate mancate risposte parziali.

Nel settore della Sanità, il numero di variabili da stimare è risultato maggiore che nell'istruzione, data la minore disponibilità di fonti utilizzabili.

Le altre variabili richieste dal Regolamento, tutte di tipo qualitativo e tutte relative alle caratteristiche personali del dipendente, sono state soggette a un processo di imputazione che utilizza le tecniche del "miglior donatore".

A tale scopo è stato usato, come per il settore privato, il software RIDA, che ha permesso di imputare tutti i valori mancanti relativi alle seguenti variabili: livello di istruzione, professione, modalità di contratto, tipo di contratto, data di assunzione e data di nascita nei casi di errori nella registrazione del codice fiscale.

Al fine di uniformare le metodologie di stima con il comparto privato, il riporto dei dati all'Universo utilizza la stessa metodologia.

Sono state ottenute ponderazioni distinte per gli enti/istituzioni (unità di primo stadio) e per i dipendenti (unità di secondo stadio) e per le unità locali. Per il settore della Sanità le fonti disponibili non permettono di fornire stime a livello di unità locali ma solo a livello delle unità di primo stadio, rappresentate dall'Ente/Istituzione nel suo complesso.

A conferma dell'efficacia del processo di integrazione sviluppato, si è verificata che la quota di situazioni di ridondanza e di incongruenza tra le fonti è inferiore al 2%.

¹⁰ Minore di 6000 euro; da 6.000 a 19.999 euro; da 20.000 a 31.999 euro; da 32.000 a 39.999 euro; da 40.000 a 49.999 euro; 50.000 euro ed oltre.

¹¹ Da 16 a 19 anni, da 20 a 29 anni; da 30 a 39 anni; da 40 a 49 anni; da 50 a 59 anni; 60 anni ed oltre.

¹² Della classificazione ISCO88 a un digit sono state considerate le professioni direttive e impiegatizie, con l'esclusione di quelle operaie in quanto non applicabili per contratto al personale del settore Istruzione.