

File Standard

Inserimento professionale dei
laureati
Indagine 2001

Manuale utente

Il manuale, curato da Paola Ungaro, è frutto della collaborazione fra l'Unità operativa Formazione universitaria del Servizio DEM, l'Unità operativa Metodologie e tecniche per la tutela della riservatezza dell'informazione statistica del Servizio MPS e l'Unità operativa Metodologie per le indagini sociali del Servizio CIN.

Gli Allegati A e B sono di Fausta Sottili, l'Allegato D di Claudia De Vitiis (paragrafi 1 e 2.1) e Adriano Pareto (paragrafo 2.2), l'Allegato E di Alessandra Capobianchi.

I programmi per l'elaborazione del file dati sono stati predisposti da Fausta Sottili.

La progettazione dell'indagine e la rilevazione dei dati sono state curate da Simona Pace.

INDICE

| | |
|--|----|
| PREMESSA | 4 |
| 1. L'INDAGINE SULL'INSERIMENTO PROFESSIONALE DEI LAUREATI | 6 |
| 1.1 Le principali caratteristiche e il disegno dell'indagine | 6 |
| 1.2 Il processo di produzione dei dati | 7 |
| 1.3 I contenuti informativi | 8 |
| 2. LE VARIABILI CONTENUTE NEL FILE | 11 |
| ALLEGATO A: IL TRACCIATO RECORD | |
| ALLEGATO B: I CARTELLINI | |
| ALLEGATO C: DEFINIZIONI | |
| ALLEGATO D: STRATEGIA DI CAMPIONAMENTO E LIVELLO DI PRECISIONE DELLE STIME | |
| 1. Strategia di campionamento | |
| 1.1. Descrizione generale del disegno di campionamento | |
| 1. 2. Procedimento per il calcolo delle stime | |
| 2. Valutazione del livello di precisione delle stime | |
| 2.1. Calcolo della varianza campionaria | |
| 2.2. Presentazione sintetica degli errori campionari | |
| ALLEGATO E: METODOLOGIA PER LA STIMA DEL RISCHIO DI VIOLAZIONE DELLA RISERVATEZZA | |

Premessa

Il Decreto Legislativo n° 322 del 6/9/1989 regola la diffusione delle informazioni statistiche prodotte nell'ambito del Sistema Statistico Nazionale al fine di garantire la riservatezza dei rispondenti. In particolare, per la diffusione di dati elementari, l'articolo 10, comma 2, dispone quanto segue: *"Sono distribuite altresì, ove disponibili, su richiesta motivata e previa autorizzazione del Presidente dell'ISTAT, collezioni campionarie di dati elementari, resi anonimi e privi di ogni riferimento che ne permetta il collegamento con singole persone fisiche e giuridiche"*.

Nell'osservanza di tale legge l'ISTAT ha adottato misure e tecniche che rendono impossibile, o altamente improbabile, il collegamento dei dati rilasciati con l'unità statistica a cui si riferiscono. Per tale motivo sui file originali delle indagini vengono apportate alcune modifiche, nell'intento di garantire la massima protezione ai dati contenendo al minimo l'eventuale perdita di informazione.

Le metodologie di protezione dei dati - che in passato si concretizzavano nell'accorpamento delle modalità di singole variabili e/o nell'oscuramento integrale di altre - sono state di recente modificate nella direzione di ampliare le informazioni a disposizione dell'utente, attraverso un più ampio ricorso alla tecnica della soppressione locale. Tale tecnica prevede che i valori assunti dalle variabili che comportano un elevato rischio di identificazione vengano oscurati a livello di singoli record e non sull'intero file dati (V. Allegato E).

Nelle prossime pagine, dopo una breve descrizione delle fasi principali dell'indagine sull'inserimento professionale dei laureati, viene riportato il tracciato record che descrive le variabili contenute nel file standard. Per agevolare l'elaborazione dei dati e l'interpretazione dei risultati, negli allegati a seguire sono riportati il questionario, le classificazioni e le definizioni adottate, e vengono illustrate le caratteristiche del disegno di campionamento e la metodologia adottata per la protezione dei dati.

Con l'occasione si ricorda al richiedente che si impegna a:

- utilizzare i dati soltanto per gli scopi dichiarati;
- non fornire a terzi i dati elementari, consentendone l'accesso, sotto la propria responsabilità, soltanto alle persone direttamente coinvolte nel lavoro per il quale essi sono stati richiesti;
- citare la fonte Istat nell'eventuale divulgazione di elaborazioni dei dati;
- inviare alla Biblioteca dell'Istat due copie delle pubblicazioni eventualmente prodotte con l'utilizzo dei dati ottenuti.

1. L'indagine sull'inserimento professionale dei laureati

1.1 Le principali caratteristiche e il disegno dell'indagine

Nel 2001 l'Istituto Nazionale di Statistica ha realizzato la quinta edizione dell'indagine sull'inserimento professionale dei laureati, una delle tre rilevazioni - la prima ad essere varata - del Sistema di indagini sulla transizione scuola-lavoro.

Si tratta di un sistema integrato di rilevazioni finalizzate ad analizzare la condizione occupazionale dei giovani in possesso di titolo di studio post-obbligo, che ad oggi comprende, oltre alla rilevazione sui laureati, analoghe indagini sui diplomati di scuola secondaria superiore¹ e sui diplomati universitari².

Le indagini sulla transizione scuola-lavoro sono fortemente coerenti dal punto di vista della metodologia adottata ed utilizzano questionari analoghi nella struttura e, fin dove possibile, nei contenuti informativi. In tutti e tre i casi le rilevazioni sono triennali, condotte su singole leve di studenti a circa tre anni dal conseguimento del titolo³ e realizzate con tecnica **C.A.T.I.** (Computer Assisted Telephone Interview).

La popolazione di riferimento dell'indagine 2001 sull'inserimento professionale dei laureati è costituita dai **laureati dell'anno solare 1998 in tutte le sedi universitarie italiane** per un totale di **129.307 individui** (57.674 uomini e 71.633 donne).

L'indagine è campionaria e prevede la formazione di un **campione a uno stadio stratificato per sesso, sede universitaria di appartenenza e corso di laurea**.

Il campione teorico è risultato pari a 30.052 unità; per far fronte alle mancate risposte totali, nel corso della rilevazione è stato selezionato un ulteriore campione di 6.321 unità. Il tasso di campionamento complessivo è risultato pari al **28,1%**.

I rispondenti all'indagine sono stati 20.844 con un tasso di risposta del **57,3%**.

¹ Cfr. Istat (2000): Percorsi di studio e di lavoro dei diplomati – Indagine 1998, Collana Informazioni.

² Cfr. Istat (2001): Inserimento professionale dei diplomati universitari – Indagine 1999, Collana Informazioni.

³ La scelta di condurre le interviste a distanza di tre anni dal conseguimento del titolo di studio - oltre a tenere conto dell'eventuale espletamento del servizio militare - è coerente con i lunghi tempi di ricerca della prima occupazione che caratterizzano il mercato del lavoro italiano. Solo dopo un certo lasso di tempo dall'uscita dal sistema d'istruzione, infatti, è possibile indagare sugli esiti occupazionali, specie se l'attenzione è volta anche alla qualità del lavoro svolto.

1.2 Il processo di produzione dei dati

L'indagine si è svolta in due fasi: la prima, che ha coinvolto gli atenei, ha consentito di costruire l'elenco da cui estrarre i nominativi da contattare; la seconda, rivolta ai laureati selezionati nel campione, ha permesso la raccolta dei dati relativi ai singoli percorsi universitari e post-laurea.

Più in particolare, nell'ambito della prima fase è stato realizzato un censimento dei laureati dell'anno 1998 presso tutti gli atenei italiani⁴. Tale elenco ha costituito la base necessaria alla formazione del campione di laureati da intervistare. Oltre alle notizie anagrafiche, le università hanno fornito una prima sintetica descrizione dei percorsi accademici dei laureati (data di conseguimento del titolo, voto riportato, durata degli studi). Queste informazioni sono state in seguito utilizzate per effettuare controlli di coerenza con quanto successivamente dichiarato dagli intervistati.

A differenza delle precedenti indagini sui laureati – che, lo si ricorda, utilizzavano per la raccolta dei dati questionari autocompilati ed inviati per posta dai rispondenti – la rilevazione del 2001 è stata realizzata mediante tecnica C.A.T.I.. Nella seconda fase di rilevazione, quindi, il campione di laureati è stato contattato telefonicamente ed intervistato da una ditta specializzata. Le interviste sono state effettuate nel corso di circa 80 giorni lavorativi; il numero medio di interviste effettuate quotidianamente da ciascun intervistatore è stato pari a 7,7.

Particolare cura, durante la rilevazione, è stata posta alla fase di codifica manuale dei quesiti aperti sulla professione: gli operatori, infatti, dopo aver registrato la descrizione delle professioni svolte dall'intervistato e dal padre hanno proceduto alla loro codifica secondo la Classificazione ISTAT delle professioni 1991⁵.

Benché la tecnica C.A.T.I. consenta di controllare la coerenza delle risposte rilasciate dall'intervistato già durante la fase di rilevazione, il file dati è stato sottoposto ad un ulteriore processo di controllo e successiva correzione. Le incoerenze e le incompatibilità sono state eliminate mediante specifici software basati su metodologie deterministiche e probabilistiche. In particolare i metodi probabilistici utilizzati si basano sulla ricerca di un laureato-donatore,

⁴ La raccolta delle liste individuali, realizzata perlopiù per via telematica, si è protratta per circa sette mesi, al termine dei quali il file dati è stato sottoposto a controlli di completezza e coerenza rispetto ai dati di fonte amministrativa prodotti dal Ministero dell'Istruzione, dell'Università e della Ricerca. Data la natura telefonica dell'indagine, nella fase di controllo, particolare attenzione è stata rivolta com'è ovvio alla revisione e correzione dei numeri telefonici forniti e alla ricerca dei numeri mancanti.

⁵ Cfr. Istat (1991): Classificazione delle professioni, Metodi e norme, serie C, n.12.

con caratteristiche simili a quelle del laureato che presenta il dato mancante o errato, che possa donare l'informazione.

Si è proceduto, quindi, alla costruzione delle stime utilizzando degli stimatori di ponderazione vincolati per sesso, corso e sede universitaria, secondo la metodologia di stima standard utilizzata dall'Istat (V. Appendice E).

I dati sono quindi stati elaborati e sottoposti a validazione mediante l'utilizzo di serie storiche e il confronto con altre indagini (Rilevazione trimestrale delle forze di lavoro e altre indagini sulla transizione scuola-lavoro).

Il processo di produzione dei dati si conclude con la loro diffusione attraverso più canali: il tradizionale volume, il sito internet dell'Istat, il file standard, il Laboratorio di Analisi dei Dati Elementari ADELE.

1.3 I contenuti informativi

Il questionario utilizzato per l'indagine sui laureati – strutturato in maniera da assicurare la confrontabilità con le informazioni raccolte dalle indagini sulla transizione dei diplomati di scuola secondaria superiore e dei diplomati universitari e l'integrazione con altre fonti quali la Rilevazione trimestrale delle forze di lavoro - si articola in cinque sezioni. La prima sezione è dedicata al curriculum degli studi e alle attività di qualificazione, la seconda al lavoro, la terza alla ricerca del lavoro, mentre nella quarta e nella quinta si raccolgono notizie relative alla famiglia d'origine e informazioni anagrafiche.

Le sezioni del questionario e i principali quesiti di snodo sono presentati nello schema 1; di seguito vengono invece illustrati gli obiettivi e i principali contenuti delle diverse sezioni.

SEZIONE 1 – CURRICULUM

Il percorso formativo intrapreso rappresenta certamente una determinante degli esiti occupazionali. Per tale motivo l'indagine propone un'ampia sezione nella quale viene analizzato il curriculum degli studi attraverso la raccolta di informazioni sulla carriera sia scolastica sia universitaria: indirizzo scolastico ed accademico concluso, rendimento, eventuale possesso di altri titoli di studio, percorso accademico (interruzioni/trasferimenti di corso), modalità di frequenza dell'università, soddisfazione per scelte di studio effettuate, ulteriore impegno in attività formative.

Le informazioni rilevate in tale sezione, integrandosi con quelle provenienti dalle fonti amministrative (indagini MIUR sull'istruzione universitaria), costituiscono un utile patrimonio per lo studio delle carriere accademiche degli studenti dei corsi di laurea.

SEZIONE 2 – LAVORO

Considerata l'elevata mobilità occupazionale che caratterizza i primi anni della vita lavorativa giovanile, la sezione sul lavoro si apre con quesiti mirati alla ricostruzione della storia lavorativa dell'intervistato (esperienze di lavoro rifiutate o interrotte dopo la laurea), proseguendo poi con l'analisi delle caratteristiche del lavoro attualmente svolto: data di inizio, tipo di lavoro, tipo di contratto, orario, guadagno mensile, professione, posizione nella professione, settore di attività economica, attività formativa sul luogo di lavoro, utilizzo del computer e delle lingue straniere, grado di soddisfazione rispetto al lavoro svolto, ecc..

In considerazione del crescente interesse nei confronti delle forme di lavoro non continuative, molti dei quesiti che nella precedente edizione erano rivolti solo a quanti dichiaravano di essere impegnati in lavori continuativi sono oggi rivolti anche a coloro che svolgono lavori occasionali o stagionali.

SEZIONE 3 – RICERCA DI LAVORO

I quesiti di questa sezione, rivolti esclusivamente a coloro che si dichiarano alla ricerca di un'occupazione (che siano o meno attualmente occupati), sono volti a rilevare notizie sia sulle modalità di ricerca del lavoro sia sulle caratteristiche del lavoro desiderato (tipo e orario di lavoro preferiti, propensione a cambiare città, attese retributive).

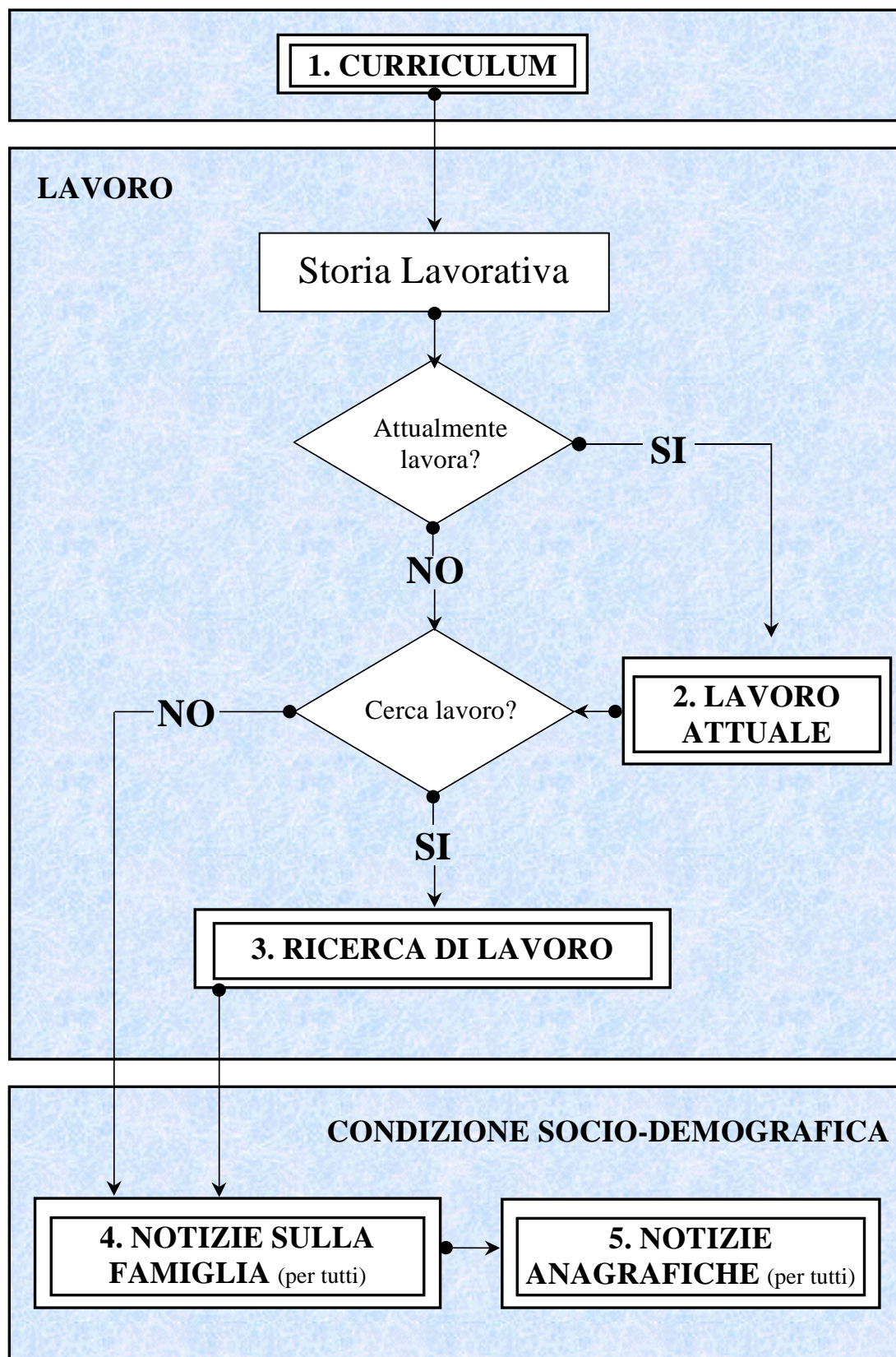
SEZIONE 4 – NOTIZIE SULLA FAMIGLIA D'ORIGINE

La sezione mira ad ottenere una descrizione dell'ambiente socio-culturale di provenienza degli intervistati attraverso la raccolta di informazioni sulla struttura familiare, sul livello di istruzione dei genitori e sulla loro condizione professionale, tutti elementi utili sia ad analizzare l'effetto delle condizioni sociali di origine sul successo negli studi, sia a confrontare la posizione sociale dei giovani rispetto a quelle della famiglie d'origine.

SEZIONE 5 – NOTIZIE ANAGRAFICHE

Questa sezione, uniformata alle altre indagini dell'Istat in base a definizioni internazionali, richiede indicazioni sulle caratteristiche di base dell'intervistato, relative alla cittadinanza, alla provincia di abitazione e di lavoro, al sesso, all'anno di nascita, allo stato civile.

Schema 1: Sezioni e snodi del questionario sull'inserimento professionale dei laureati del 1998



2. Le variabili contenute nel file

L'adozione di una metodologia di protezione dei dati che si basa sulla tecnica di soppressione locale (vedi Allegato E) ha determinato la presenza di alcuni blank nelle variabili corso di laurea, sede universitaria, sesso, residenza ed età.

Inoltre per queste e per alcune altre variabili (mese di laurea, guadagno mensile, provincia di lavoro, stato civile) sono state costruite delle aggregazioni.

Nel tracciato record (Allegato A) vengono riportati nell'ordine:

1. il nome delle variabili (colonna "Variabile"), definito dal numero del corrispondente quesito nel questionario di rilevazione;
2. la posizione (colonne "da posizione" "a posizione");
3. l'ampiezza del campo in byte (colonna "n. di posizioni");
4. la descrizione delle variabili (colonna "Descrizione variabile"), che definisce il contenuto informativo di ciascuna variabile e le codifiche adottate, rimandando ai cartellini (Allegato B) solo nei casi in cui si sia proceduto a codifiche diverse da quelle presenti nel questionario.

La variabile COEFF individua il coefficiente di riporto all'universo, cioè il peso che ogni unità campionaria assume nella popolazione di riferimento. Ad esempio, se in un record il valore di COEFF è pari a 5, significa che le informazioni contenute in quel record sono riferite a 5 laureati dell'anno 1998.

Allegato A

IL TRACCIATO RECORD

| Varia- bile | Da posi- zione | A posi- zione | N. di posi- zioni | Descrizione variabile | |
|--|----------------------|---------------------|-------------------------|--|--|
| progr_u | 1 | 6 | 6 | Identificativo del laureato | |
| INFORMAZIONI RILASCIATE DALLE UNIVERSITA' | | | | | |
| Sede | 7 | 7 | 1 | Ripartizione geografica della sede universitaria | (Alleg. C Cartellino A) |
| Corso | 8 | 9 | 2 | Corso di laurea | (Alleg. C Cartellino B) |
| Gruppo | 10 | 11 | 2 | Gruppo di corsi di laurea | (Alleg. C Cartellino C) |
| Sezione 1: CURRICULUM | | | | | |
| q1_1 | 12 | 13 | 2 | Tipo di maturità | |
| q1_2a | 14 | 15 | 2 | Voto maturità | |
| q1_3a | 16 | 16 | 1 | Possesso di altra laurea | 1=Si |
| q1_3b | 17 | 17 | 1 | Possesso di diploma univ. o scuola diretta a fini speciali | 2=Si |
| q1_4 | 18 | 18 | 1 | Altro titolo univers. conseguito prima o dopo il 1998 | |
| q1_5 | 19 | 19 | 1 | Motivo iscrizione ad un nuovo corso | |
| q1_6 | 20 | 20 | 1 | Laurea in corso o fuori corso | |
| q1_7 | 21 | 21 | 1 | Anno fuori corso | |
| q1_8 | 22 | 24 | 3 | Voto massimo laurea | |
| q1_9 | 25 | 27 | 3 | Voto laurea | |
| q1_10 | 28 | 28 | 1 | Lode | |
| q1_11 | 29 | 30 | 2 | Trimestre di laurea | 1=Gen. Feb. Mar. 2=Apr. Mag. Giu. 3=Lug. Ago. Set. 4=Ott. Nov. Dic. |
| q1_12 | 31 | 31 | 1 | Interruzione precedente corso universitario | |
| q1_13 | 32 | 33 | 2 | Gruppo disciplinare del corso interrotto | (Alleg. C Cartellino C) |
| q1_14 | 34 | 34 | 1 | Sede universitaria nella stessa città in cui viveva | |
| q1_15 | 35 | 35 | 1 | Trasferimento in altra città | |
| q1_16 | 36 | 36 | 1 | Motivo del mancato trasferimento | |
| q1_17 | 37 | 37 | 1 | Per quanto tempo si è trasferito | |
| q1_18 | 38 | 38 | 1 | Abitazione dopo il trasferimento | |
| q1_19 | 39 | 39 | 1 | Frequenza a corsi privati preparazione esami | |
| q1_20 | 40 | 40 | 1 | Lavoro durante gli studi | |
| q1_21 | 41 | 41 | 1 | Si iscriverebbe di nuovo allo stesso corso | |
| q1_22 | 42 | 42 | 1 | Tipo di corso a cui si iscriverebbe | |
| q1_23 | 43 | 44 | 2 | Gruppo disciplinare del corso a cui si iscriverebbe | (Alleg. C Cartellino C) |
| q1_24 | 45 | 45 | 1 | Motivo per cui non si reiscriverebbe allo stesso corso | |
| q1_25 | 46 | 46 | 1 | Conseguimento dell'abilitazione professionale | |
| q1_26a | 47 | 47 | 1 | Formazione ulteriore: Dottorato di ricerca | |
| q1_26b | 48 | 48 | 1 | Formazione ulteriore: Specializzazione post-laurea | |
| q1_26c | 49 | 49 | 1 | Formazione ulteriore: Master | |
| q1_26d | 50 | 50 | 1 | Formazione ulteriore: Corso di laurea | |
| q1_26e | 51 | 51 | 1 | Formazione ulteriore: Corso di Diploma univ. (o S.D.F.S.) | |
| q1_26f | 52 | 52 | 1 | Formazione ulteriore: Borsa di studio universitaria | |
| q1_26g | 53 | 53 | 1 | Formazione ulteriore: Borsa di studio extrauniversitaria | |
| q1_26h | 54 | 54 | 1 | Formazione ulteriore: Stage o tirocinio | |
| q1_26i | 55 | 55 | 1 | Formazione o aggiornamento (oltre 3 mesi/ 300 ore) | |
| q1_26l | 56 | 56 | 1 | Altra attività di studio e formazione | |

| Sezione 2: LAVORO | | | | | |
|-------------------|-----|-----|---|---|---|
| q2_1 | 57 | 57 | 1 | Dopo la laurea opportunità di lavoro rifiutate | |
| q2_2 | 58 | 59 | 2 | Motivo del rifiuto | |
| q2_3 | 60 | 60 | 1 | Ha iniziato qualche lavoro poi interrotto | |
| q2_4 | 61 | 61 | 1 | Tipo di lavoro | |
| q2_5 | 62 | 62 | 1 | Lavoro autonomo | |
| q2_6a | 63 | 63 | 1 | Anno di inizio del lavoro | |
| q2_6b | 64 | 65 | 2 | Mese di inizio del lavoro | |
| q2_7 | 66 | 67 | 2 | Per quale motivo lo ha interrotto | |
| q2_8 | 68 | 69 | 2 | Attività formativa retribuita | |
| q2_9 | 70 | 70 | 1 | Lavora | |
| q2_10 | 71 | 71 | 1 | Ore di lavoro la settimana precedente | |
| q2_11 | 72 | 73 | 2 | N° di ore retribuite | |
| q2_12 | 74 | 74 | 1 | Cerca lavoro | |
| q2_13 | 75 | 76 | 2 | Motivo per cui NON cerca lavoro | |
| q2_14 | 77 | 77 | 1 | Tipo di lavoro | |
| q2_15 | 78 | 78 | 1 | Lavoro iniziato prima o dopo la laurea | |
| q2_16a | 79 | 79 | 1 | Anno di inizio del lavoro | |
| q2_16b | 80 | 81 | 2 | Mese di inizio del lavoro | |
| q2_17 | 82 | 82 | 1 | Laurea requisito necessario | |
| q2_18 | 83 | 83 | 1 | Tipo laurea necessaria | |
| q2_19 | 84 | 84 | 1 | Era richiesta una votazione minima | |
| q2_20 | 85 | 85 | 1 | Ha dovuto trasferirsi | |
| q2_21 | 86 | 86 | 1 | Tempo pieno o ridotto in part-time | |
| q2_22 | 87 | 87 | 1 | Part-time per scelta o per mancanza di altro lavoro | |
| q2_23 | 88 | 89 | 2 | Quante ore sono retribuite settimanalmente | |
| q2_24a | 90 | 90 | 1 | Blank | |
| q2_24b | 91 | 95 | 5 | Guadagno mensile netto in classi | 0=non indicato 1= fino a €800 2= da €800,01 a €1100 3= da €1100,01 a €1500 4= oltre €1500 |
| q2_25a | 96 | 96 | 1 | Blank | Riservato Istat |
| q2_25b | 97 | 100 | 4 | Professione del laureato | (Alleg. C Cartellino D) |
| q2_26 | 101 | 101 | 1 | Settore pubblico o privato | |
| q2_27 | 102 | 102 | 1 | Settore di attività | |
| q2_28 | 103 | 104 | 2 | Settore di "Altre attività" | |
| q2_29 | 105 | 106 | 2 | Settore di "Industria" | |
| q2_30 | 107 | 107 | 1 | N° dei dipendenti impresa, ente o studio | |
| q2_31 | 108 | 108 | 1 | Posizione nella professione | |
| q2_32 | 109 | 110 | 2 | Posizione professione lavoratori indipendenti | |
| q2_33 | 111 | 111 | 1 | Partita IVA | |
| q2_34 | 112 | 112 | 1 | Ritenuta d'acconto | |
| q2_35 | 113 | 114 | 2 | Tipo di contratto | |
| q2_36 | 115 | 116 | 2 | Posizione professione lavoratori dipendenti | |
| q2_37 | 117 | 117 | 1 | Contributi per la pensione | |
| q2_38 | 118 | 118 | 1 | Partecipazione a corsi di formazione | |
| q2_39 | 119 | 119 | 1 | Durata dei corsi | |

| | | | | | |
|---|-----|-----|---|---|---|
| q2_40 | 120 | 120 | 1 | Utilizzo pc | |
| q2_41a | 121 | 122 | 2 | Utilizzo pc: per scrivere | |
| q2_41b | 123 | 124 | 2 | Utilizzo pc: per disegnare | |
| q2_41c | 125 | 126 | 2 | Utilizzo pc: per navigare su internet | |
| q2_41d | 127 | 128 | 2 | Utilizzo pc: per consultare archivi | |
| q2_41e | 129 | 130 | 2 | Utilizzo pc: per il trattamento del suono | |
| q2_41f | 131 | 132 | 2 | Utilizzo pc: per inserire dati | |
| q2_41g | 133 | 134 | 2 | Utilizzo pc: per spedire la posta elettronica | |
| q2_41h | 135 | 136 | 2 | Utilizzo pc: per creare prodotti multimediali | |
| q2_41i | 137 | 138 | 2 | Utilizzo pc: per programmare | |
| q2_41l | 139 | 140 | 2 | Utilizzo pc: per fare calcoli | |
| q2_41m | 141 | 142 | 2 | Utilizzo pc: altro | |
| q2_42a | 143 | 143 | 1 | Quanto spesso legge in una lingua straniera? | |
| q2_42b | 144 | 144 | 1 | Quanto spesso scrive/parla in lingua straniera? | |
| q2_43a | 145 | 145 | 1 | Soddisfazione: trattamento economico | |
| q2_43b | 146 | 146 | 1 | Soddisfazione: possibilità di carriera | |
| q2_43c | 147 | 147 | 1 | Soddisfazione: mansioni svolte | |
| q2_43d | 148 | 148 | 1 | Soddisfazione: stabilità/sicurezza | |
| q2_43e | 149 | 149 | 1 | Soddisfazione: grado di autonomia del lavoro | |
| q2_43f | 150 | 150 | 1 | Soddisfazione: utilizzo conoscenze universitarie | |
| q2_44 | 151 | 151 | 1 | Valutazione della laurea rispetto al lavoro | |
| q2_45 | 152 | 153 | 2 | In che modo ha trovato il suo attuale lavoro? | |
| q2_46 | 154 | 154 | 1 | C'è stata una persona utile per iniziare il lavoro? | |
| q2_47 | 155 | 155 | 1 | Come questa persona è stata utile? | |
| q2_48 | 156 | 156 | 1 | Di chi si tratta? | |
| q2_49 | 157 | 157 | 1 | Cerca un nuovo lavoro? | |
| Sezione 3: RICERCA DI LAVORO | | | | | |
| q3_1 | 158 | 158 | 1 | Periodo dell'ultima iniziativa di ricerca | |
| q3_2 | 159 | 159 | 1 | Preferenza tempo pieno o part-time | |
| q3_3 | 160 | 160 | 1 | Preferenza lavoro dipendente o autonomo | |
| q3_4 | 161 | 161 | 1 | Disponibilità a lavorare all'estero | |
| q3_5 | 162 | 162 | 1 | Disponibilità a cambiare città | |
| q3_6a | 163 | 163 | 1 | Blank | |
| q3_6b | 164 | 168 | 5 | Minimo guadagno mensile desiderato in classi | 0=non indicato 1= fino a €800 2= da €800,01 a €1100 3= da €1100,01 a €1500 4= oltre €1500 |
| q3_7 | 169 | 169 | 1 | Disponibilità a cominciare entro 15 giorni | |
| Sezione 4: NOTIZIE SULLA FAMIGLIA DI ORIGINE | | | | | |
| q4_1 | 170 | 170 | 1 | Ha fratelli/sorelle? | |
| q4_2 | 171 | 171 | 1 | Quanti? | |
| q4_3 | 172 | 172 | 1 | Titolo di studio del padre | |
| q4_4 | 173 | 173 | 1 | Titolo di studio della madre | |
| q4_5 | 174 | 174 | 1 | Condizione occupazionale del padre | |
| q4_6a | 175 | 175 | 1 | Blank | Riservato Istat |
| q4_6b | 176 | 179 | 4 | Professione del padre del laureato | (Alleg. C Cartellino D) |
| q4_7 | 180 | 180 | 1 | Padre: posizione nella professione | |
| q4_8 | 181 | 182 | 2 | Padre: posizione profess. lavoratori indipendenti | |
| q4_9 | 183 | 184 | 2 | Padre: posizione profess. lavoratori dipendenti | |
| q4_10 | 185 | 185 | 1 | Padre: settore di attività | |

| | | | | | |
|---------------------------------------|------|-----|----|---|--|
| q4_11 | 186 | 186 | 1 | Condizione occupazionale della madre | |
| q4_12 | 187 | 187 | 1 | Madre: posizione nella professione | |
| q4_13 | 188 | 189 | 2 | Madre: posizione profess. lavoratori indipendenti | |
| q4_14 | 1901 | 191 | 2 | Madre: posizione profess. lavoratori dipendenti | |
| q4_15 | 192 | 192 | 1 | Madre: settore di attività | |
| Sezione 5: NOTIZIE ANAGRAFICHE | | | | | |
| q5_1 | 193 | 193 | 1 | Cittadinanza | |
| q5_2 | 194 | 194 | 1 | Blank | Riservato Istat |
| q5_3 | 195 | 197 | 3 | Blank | Riservato Istat |
| q5_4 | 198 | 198 | 1 | Prov. residenza all'iscrizione universitaria stessa dell'attuale? | |
| q5_5 | 199 | 201 | 3 | blank | Riservato Istat |
| regione | 202 | 203 | 2 | Regione di residenza attuale | (Alleg. C Cartellino E) |
| rip_res | 204 | 204 | 1 | Ripartizione geografica di residenza attuale | (Alleg. C Cartellino A) |
| q5_6a | 205 | 206 | 2 | Con chi abita? Solo | |
| q5_6b | 207 | 208 | 2 | Con chi abita? Padre | |
| q5_6c | 209 | 210 | 2 | Con chi abita? Madre | |
| q5_6d | 211 | 212 | 2 | Con chi abita? Fratelli/sorelle | |
| q5_6e | 213 | 214 | 2 | Con chi abita? Amici | |
| q5_6f | 215 | 216 | 2 | Con chi abita? Coniuge/convivente | |
| q5_6g | 217 | 218 | 2 | Con chi abita? Figli | |
| q5_6h | 219 | 220 | 2 | Con chi abita? Altri parenti | |
| q5_6i | 221 | 222 | 2 | Con chi abita? Altro | |
| q5_7 | 223 | 223 | 1 | Provincia in cui lavora | 1=stessa provincia di residenza 2=altra provincia |
| Reg_lav | 224 | 224 | 1 | Regione in cui lavora | 1=stessa regione di residenza 2=altra regione |
| q5_8 | 225 | 225 | 1 | Sesso | |
| q5_9 | 226 | 226 | 1 | Obblighi di leva | |
| q5_10 | 227 | 227 | 1 | Quando li ha assolti? | |
| q5_11a | 228 | 229 | 2 | Età in classi | 24= fino a 24 anni 25= 25 anni 26= 26 anni 27= 27 anni 28= 28 e 29 anni 30= 30 anni e più |
| q5_12 | 230 | 230 | 1 | Stato civile | 1=coniugato/a, convivente 2=altro |
| coeff | 231 | 243 | 13 | coefficiente di riporto | (10 decimali) |

Allegato B

I CARTELLINI

CARTELLINO A - Ripartizioni geografiche

- 1 Nord-ovest *(Comprende: Piemonte, Valle d'Aosta, Lombardia, Liguria)*
- 2 Nord-est *(Comprende: Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna)*
- 3 Centro *(Comprende: Toscana, Umbria, Marche, Lazio)*
- 4 Sud *(Comprende: Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria)*
- 5 Isole *(Comprende: Sicilia, Sardegna)*
- 9 Estero

CARTELLINO B - Corsi di laurea

- 1 Matematica
- 2 Fisica
- 3 Altri corsi del gruppo scientifico
(Comprende: Astronomia, Informatica, Scienze dell'informazione, Scienze nautiche)
- 5 Chimica
- 6 Farmacia
- 7 Altri corsi del gruppo chimico-farmaceutico
(Comprende: Chimica e tecnologia farmaceutiche, Chimica industriale)
- 8 Scienze geologiche
- 9 Scienze biologiche
- 10 Altri corsi del gruppo geo-biologico
(Comprende: Scienze ambientali, Scienze naturali)
- 11 Medicina e chirurgia
- 12 Odontoiatria e protesi dentarie
- 13 Ingegneria meccanica
- 14 Ingegneria elettronica
- 15 Ingegneria chimica
- 16 Ingegneria civile ed edile
- 17 Ingegneria informatica
- 18 Ingegneria elettrica
- 19 Ingegneria delle telecomunicazioni
- 20 Ingegneria gestionale
- 21 Ingegneria ambiente e territorio
- 22 Altri corsi del gruppo ingegneria
(Comprende: Ingegneria: aeronautica, aerospaziale, civile per la difesa del suolo e la pianificazione territoriale, dei materiali, delle tecnologie industriali, edile, elettrotecnica, forestale, mineraria, navale, navale e meccanica, nucleare)
- 23 Architettura
- 24 Altri corsi del gruppo architettura
(Comprende: Disegno industriale, Pianificazione territoriale e urbanistica, Pianificazione territoriale, urbanistica ed ambientale, Storia e conservazione dei beni architettonici e ambientali, Urbanistica)
- 25 Scienze agrarie *(Comprende: Scienze e tecnologie agrarie)*
- 26 Scienze forestali *(Comprende: Scienze forestali e ambientali)*
- 27 Medicina veterinaria
- 28 Scienze e tecnologie alimentari *(Comprende: Scienze delle preparazioni alimentari)*

- 29** Altri corsi del gruppo agrario
(Comprende: Agricoltura tropicale e sub-tropicale, Biotecnologie agro-industriali, Scienze agrarie tropicali e sub-tropicali, Scienze della produzione animale)
- 30** Economia e commercio
- 31** Scienze economiche e bancarie
- 32** Scienze statistiche ed economiche
- 33** Scienze statistiche e demografiche
- 34** Altre Scienze statistiche
(Comprende: Scienze statistiche ed attuariali, Scienze statistiche, Scienze statistiche demografiche e sociali, Statistica e informatica per l'azienda)
- 35** Economia aziendale
- 36** Economia politica
- 37** Economia bancaria, finanziaria e assicurativa
- 38** Economia del commercio internazionale e mercati valutari
- 39** Altri corsi del gruppo economico-statistico
(Comprende: Commercio internazionale e mercati valutari, Discipline economiche e sociali, Economia amministrazioni pubbliche e istituzioni internazionali, Economia bancaria, Economia del turismo, Economia delle istituzioni e mercati finanziari, Economia e legislazione per l'impresa, Economia marittima e dei trasporti, Scienze bancarie e assicurative, Scienze economiche, Scienze economiche e sociali)
- 40** Scienze politiche
- 41** Sociologia
- 42** Scienze della comunicazione
- 43** Altri corsi del gruppo politico-sociale
(Comprende: Relazioni pubbliche, Scienze internazionali e diplomatiche)
- 44** Giurisprudenza *(Comprende: Scienze dell'amministrazione)*
- 45** Lettere
- 46** Materie letterarie
- 47** Filosofia
- 48** Storia
- 49** Conservazione dei beni culturali
- 50** Altri corsi del gruppo letterario
(Comprende: Discipline arti, musica e spettacolo, Filologia e storia dell'Europa orientale, Geografia, Musicologia)
- 51** Lingue e letterature straniere moderne
- 52** Lingue e letterature straniere
- 53** Lingue e civiltà orientali *(Comprende: Lingue e letterature orientali)*
- 54** Altri corsi del gruppo linguistico
(Comprende: Interprete, Traduttore, Traduzione e interpretazione)
- 55** Pedagogia
- 56** Scienze dell'educazione *(Comprende: Scienze della formazione primaria)*
- 57** Psicologia

CARTELLINO C - Gruppi di corsi di laurea

- 1** Gruppo scientifico
- 2** Gruppo chimico-farmaceutico
- 3** Gruppo geo-biologico
- 4** Gruppo medico
- 5** Gruppo ingegneria
- 6** Gruppo architettura
- 7** Gruppo agrario
- 8** Gruppo economico-statistico
- 9** Gruppo politico-sociale
- 10** Gruppo giuridico
- 11** Gruppo letterario
- 12** Gruppo linguistico
- 13** Gruppo insegnamento
- 14** Gruppo psicologico
- 15** Gruppo educazione fisica

CARTELLINO D - Professioni

- 0** Non indicato
- 1** Legislatori, dirigenti e imprenditori
- 2** Professioni intellettuali, scientifiche e di elevata specializzazione
- 3** Professioni intermedie (tecnici)
- 4** Professioni esecutive relative all'amministrazione e gestione
- 5** Professioni relative alle vendite ed ai servizi per le famiglie
- 6** Artigiani, operai specializzati e agricoltori
- 7** Conduttori di impianti, operatori di macchinari
fissi e mobili e operai di montaggio industriale
- 8** Personale non qualificato
- 9** Forze armate

CARTELLINO E – Regioni di residenza

- 1** Piemonte
- 2** Valle d'Aosta
- 3** Lombardia
- 4** Trentino Alto Adige
- 5** Veneto
- 6** Friuli Venezia Giulia
- 7** Liguria
- 8** Emilia Romagna
- 9** Toscana
- 10** Umbria
- 11** Marche
- 12** Lazio
- 13** Abruzzo
- 14** Molise
- 15** Campania
- 16** Puglia
- 17** Basilicata
- 18** Calabria
- 19** Sicilia
- 20** Sardegna
- 99** Estero

Allegato C

DEFINIZIONI

- 1) *VOTO DI LAUREA*: è espresso in centodecimi, anche quando la scala prevista è diversa.
- 2) *OCCUPATO*: chi svolge un'attività lavorativa retribuita (ivi incluso apprendistato e contratti di formazione e lavoro), anche se non regolarizzata da contratto; sono invece escluse tutte le attività di qualificazione (dottorato di ricerca, borsa di studio, tirocinio, stage eccetera) anche se retribuite.
- 3) *PERSONA IN CERCA DI OCCUPAZIONE*: i laureati che dichiarano di essere in cerca di lavoro.
- 4) *TIPO DI LAVORO*:
 - Continuativo*: sono tutti i lavori che vengono svolti con continuità, cioè con una cadenza regolare, con o senza regolare contratto, con o senza un termine previsto.
 - Stagionale*: sono lavori che vengono svolti in modo continuativo, ma solo in alcuni periodi dell'anno.
 - Occasionale* (o saltuario): lavori vengono svolti di tanto in tanto (per brevi periodi), senza una cadenza regolare o prevedibile.
- 5) *LAVORO PART-TIME*: il lavoro si definisce part-time quando il lavoratore, in base ad un accordo con il datore di lavoro, è tenuto a prestare un orario inferiore a quello normalmente in vigore per gli altri occupati dell'ambiente di lavoro in cui opera. Anche un lavoratore indipendente può lavorare part-time, ad esempio un negoziante che svolge la propria attività nel suo negozio solo la mattina o il pomeriggio.
- 6) *GUADAGNO MENSILE NETTO*: guadagno percepito mensilmente da cui sono stati detratti i contributi, le imposte eccetera.

7) POSIZIONE NELLA PROFESSIONE:

Lavoratore dipendente: chi presta il suo lavoro alle dipendenze di un datore di lavoro pubblico o privato, è pagato con un salario o stipendio e non potrebbe svolgere il proprio lavoro se altri non gli mettessero a disposizione i mezzi per farlo.

Lavoratore indipendente: chi svolge un lavoro autonomo da solo o con altri soci, essendo proprietario di tutti gli strumenti tecnici e organizzativi necessari.

Consulente/collaboratore: chi, pur risultando formalmente lavoratore autonomo, presta la propria opera di consulenza o collaborazione per conto di un committente, operando prevalentemente (ma non necessariamente) presso di esso e con i mezzi tecnici e organizzativi da questo predisposti.

8) TIPO DI CONTRATTO:

A tempo indeterminato: non prevede nessuna scadenza del rapporto di lavoro e può essere a tempo pieno o part-time.

Formazione e lavoro: rapporto di lavoro subordinato a termine - della durata massima di 24 mesi - avente duplice contenuto, di lavoro e formazione. Può essere instaurato con lavoratori di età compresa tra i 16 ed i 32 anni (le Regioni possono elevare il limite di età).

Apprendistato: rapporto con contenuto misto di formazione e lavoro (di durata variabile tra i 18 mesi e 4 anni) che prevede che l'imprenditore o azienda si impegni a fornire al giovane apprendista la formazione necessaria per diventare un lavoratore qualificato.

Lavoro interinale: contratto a tempo determinato mediante il quale un'impresa di fornitura di lavoro temporaneo (impresa fornitrice) pone uno o più lavoratori, da essa assunti, a disposizione di un'impresa che ne utilizzi la prestazione lavorativa (impresa utilizzatrice) per esigenze di carattere temporaneo. E' quindi un rapporto trilaterale tra lavoratore, agenzia di lavoro temporaneo e impresa utilizzatrice.

Collaborazione coordinata e continuativa: rapporto di lavoro che prevede che il lavoratore presti - in forma coordinata e continuativa nel tempo - la propria opera di collaborazione per conto di un committente senza vincolo di subordinazione, operando prevalentemente (ma non necessariamente) presso di esso e con i mezzi tecnici e organizzativi da questo predisposti.

ALLEGATO D

STRATEGIA DI CAMPIONAMENTO E LIVELLO DI PRECISIONE DELLE STIME

1. Strategia di campionamento

La *popolazione di interesse* dell'indagine – ossia l'insieme delle unità statistiche relativamente alle quali si intende investigare – è costituita dai laureati che hanno conseguito il diploma di laurea nel corso dell'anno 1998 – ossia tre anni prima della data di osservazione dell'indagine - in tutte le sedi universitarie italiane.

L'indagine è volta ad acquisire una vasta gamma di informazioni sulle condizioni dei laureati a circa tre anni dal conseguimento della laurea e si propone lo studio della transizione dall'università al mondo del lavoro. L'intervallo di tre anni dovrebbe consentire di analizzare l'inserimento nel mondo del lavoro in una fase di relativa stabilità nella condizione occupazionale.

Gli obiettivi conoscitivi più significativi riguardano lo stato occupazionale, la stabilità dell'occupazione, la posizione nella professione e l'area di attività economica distintamente per sesso, corso di laurea e ripartizione geografica di residenza.

I *domini di interesse* sono l'intero territorio nazionale, le ripartizioni geografiche e i gruppi di corsi di laurea incrociati con il sesso e le regioni per un numero limitato di stime.

1.1. Descrizione generale del disegno di campionamento

La progettazione del disegno campionario ha tenuto conto di due elementi: a) per ciascuna sede universitaria si disponeva della lista dei laureati dell'anno 1998; b) la rilevazione sarebbe stata effettuata mediante intervista telefonica. Di conseguenza, è stato definito un disegno di campionamento ad uno stadio stratificato e la stratificazione delle unità della popolazione è stata definita sulla base dell'incrocio delle tre variabili: *corso di laurea*, *sede universitaria* e *sesso* (gli strati così definiti sono risultati 2253).

La numerosità campionaria complessiva è stata determinata in modo da garantire che gli errori di campionamento attesi delle principali stime riferite ai diversi domini di interesse non superassero prefissati livelli, tenendo conto anche dei tassi di risposta attesi. La dimensione campionaria teorica complessiva è risultata pari a 30.052 unità.

Una volta definite le numerosità campionarie teoriche negli strati, la selezione delle unità campionarie è stata effettuata senza reimmissione e con probabilità uguali.

Allo scopo di illustrare l'allocazione della dimensione campionaria tra i domini di interesse, viene riportata nel Prospetto 1 la distribuzione dei laureati dell'anno 1998 distinta per sesso e per gruppo di corsi di laurea, nell'universo dei laureati e nel campione teorico. Il campione realizzato è stato di 20.844 unità.

Prospetto 1 – Distribuzione dei laureati, nell'universo e nel campione teorico, per sesso e gruppo di corsi di laurea - Anno 1998

| GRUPPO DI CORSI | MASCHI | | FEMMINE | | TOTALE | |
|-----------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | Laureati Universo | Laureati Campione | Laureati Universo | Laureati Campione | Laureati Universo | Laureati Campione |
| Gruppo scientifico | 2.681 | 792 | 2.312 | 711 | 4.993 | 1.503 |
| Gruppo chimico-farmaceutico | 1.659 | 628 | 2.763 | 809 | 4.422 | 1.437 |
| Gruppo geo-biologico | 1.933 | 674 | 3.496 | 932 | 5.429 | 1.606 |
| Gruppo medico | 3.369 | 890 | 3.631 | 937 | 7.000 | 1.827 |
| Gruppo ingegneria | 12.563 | 2.466 | 2.091 | 811 | 14.654 | 3.277 |
| Gruppo architettura | 3.928 | 852 | 3.720 | 827 | 7.648 | 1.679 |
| Gruppo agrario | 1.491 | 578 | 1.111 | 516 | 2.602 | 1.094 |
| Gruppo economico-statistico | 13.086 | 2.426 | 11.580 | 2.194 | 24.666 | 4.620 |
| Gruppo politico-sociale | 4.404 | 1.026 | 5.652 | 1.257 | 10.056 | 2.283 |
| Gruppo giuridico | 8.094 | 1.531 | 11.396 | 2.044 | 19.490 | 3.575 |
| Gruppo letterario | 2.993 | 854 | 9.940 | 1.890 | 12.933 | 2.744 |
| Gruppo linguistico | 649 | 418 | 7.220 | 1.485 | 7.869 | 1.903 |
| Gruppo insegnamento | 399 | 388 | 3.727 | 889 | 4.126 | 1.277 |
| Gruppo psicologico | 606 | 430 | 2.813 | 797 | 3.419 | 1.227 |
| Totale | 57.855 | 13.953 | 71.452 | 16.099 | 129.307 | 30.052 |

1.2. Procedimento per il calcolo delle stime

Le stime prodotte dall'indagine sono principalmente stime di frequenze assolute e relative riferite ai laureati dell'anno 1998; una stima di interesse è data, ad esempio, dal numero totale di laureati che lavorano a un anno dal conseguimento del diploma.

Il principio su cui è basato ogni metodo di stima campionaria è che le unità appartenenti al campione rappresentino anche le unità della popolazione che non sono incluse nel campione. Questo principio viene realizzato attribuendo a ogni unità campionaria un peso che denota il numero di unità della popolazione rappresentate dalla unità medesima. Se, ad esempio, a una unità campionaria viene attribuito un peso pari a 30, ciò significa che questa unità rappresenta se stessa ed altre 29 unità della popolazione che non sono state incluse nel campione.

Al fine di rendere più chiara la successiva esposizione, introduciamo la seguente simbologia:

- d , indice del livello di riferimento delle stime (domini di interesse);
- i , indice di laureato;
- h , indice dello strato (sede \times corso \times sesso);
- N_h , totale dei laureati dello strato h ;
- n_h^* , numerosità campionaria teorica (ossia estratta) dello strato h ;
- n_h , numerosità campionaria realizzata nello strato h ;
- H , numero totale di strati;
- y , generica variabile oggetto di indagine;
- Y_{hi} , valore osservato della variabile y sull' i -mo laureato dello strato h ;
- δ_{hid} , variabile indicatrice pari a 1 se l'unità i dello strato h appartiene al dominio di interesse.

Se, ad esempio, y rappresenta la condizione lavorativa (espressa dalle due modalità lavora, non lavora), si avrà $Y_{hi} = 1$ se il laureato i -mo lavora e $Y_{hi} = 0$ altrimenti.

Si supponga di voler stimare con riferimento a un generico dominio d , il totale della variabile in esame, espresso dalla relazione:

$$Y_d = \sum_{h=1}^{H_d} \sum_{i=1}^{n_h} Y_{hi} \delta_{hid} \quad (1)$$

La stima del totale (1), viene determinata mediante uno stimatore corretto stratificato espresso mediante la seguente formula:

$$\hat{Y}_d = \sum_{h=1}^{H_d} \hat{Y}_h \quad (2)$$

dove

$$\hat{Y}_h = \sum_{i=1}^{n_h} Y_{hi} \delta_{hid} W_h \quad (3)$$

Nella (3) si è indicato con W_h il peso finale da attribuire ai rispondenti dello strato h . Tale peso finale, uguale per tutte le unità appartenenti a un medesimo strato della popolazione, viene ottenuto come prodotto del peso diretto D_h , definito come inverso della probabilità di inclusione di tali unità nel campione, e dell'inverso del tasso di risposta osservato nello strato h , r_h . In simboli si ha:

$$W_h = D_h \cdot \frac{1}{r_h} = \frac{N_h}{n_h^*} \cdot \frac{n_h^*}{n_h} \quad (4)$$

Pertanto, la stima di Y_d è semplicemente

$$\hat{Y}_d = \sum_{h=1}^H \bar{y}_{hd} N_h, \quad (5)$$

dove \bar{y}_{hd} è data da

$$\bar{y}_{hd} = \frac{1}{n_h} \sum_{i=1}^{n_h} Y_{hi} \delta_{hid}.$$

2. Valutazione del livello di precisione delle stime

2.1. Calcolo della varianza campionaria

Le principali statistiche per valutare la variabilità campionaria delle stime sono l'errore di campionamento assoluto e l'errore di campionamento relativo.

La stima dell'errore di campionamento assoluto e dell'errore di campionamento relativo di una generica stima \hat{Y}_d sono definite rispettivamente dalle seguenti espressioni:

$$\hat{\sigma}(\hat{Y}_d) = \sqrt{\hat{Var}(\hat{Y}_d)} \quad (6)$$

$$\hat{\varepsilon}(\hat{Y}_d) = \frac{\hat{\sigma}(\hat{Y}_d)}{\hat{Y}_d} \quad (7)$$

La stima della varianza di \hat{Y}_d , indicata nella (6) come $\hat{Var}(\hat{Y}_d)$, è stata calcolata ottenuta sulla base delle espressioni per la stima della varianza dello stimatore corretto per il campionamento stratificato:

$$\hat{Var}(\hat{Y}_d) = \sum_{h=1}^H \frac{N_h^2}{n_h} \left(1 - \frac{n_h}{N_h}\right) \sum_{i=1}^{n_h} \frac{(Y_{hid} - \bar{y}_{hd})^2}{n_h - 1} \quad (8)$$

Gli errori campionari espressi dalla (6) e dalla (7) consentono di valutare il grado di precisione delle stime; inoltre, la conoscenza dell'errore assoluto permette anche di costruire un intervallo di confidenza al livello α , sulla base della relazione

$$\Pr\{Y_d - k\hat{\sigma}(\hat{Y}_d) \leq \hat{Y}_d \leq Y_d + k\hat{\sigma}(\hat{Y}_d)\} = \alpha \quad (9)$$

Avendo, ad esempio, prefissato $\alpha=0.05$, l'intervallo di confidenza è dato da:

$$\{\hat{Y}_d - 1,96 \cdot \hat{\sigma}(\hat{Y}_d); \hat{Y}_d + 1,96 \cdot \hat{\sigma}(\hat{Y}_d)\}$$

2.2. Presentazione sintetica degli errori campionari

Ad ogni stima \hat{Y}_d è associato un errore campionario relativo $\hat{\varepsilon}(\hat{Y}_d)$ e quindi, per consentire un uso corretto delle stime fornite dall'indagine, sarebbe necessario presentare, per ogni stima pubblicata, anche il corrispondente errore di campionamento relativo. Tuttavia, non è possibile soddisfare questa esigenza di informazione, sia per motivi di tempi e di costi di elaborazione, sia perché le tavole della pubblicazione risulterebbero eccessivamente appesantite e di non agevole consultazione per l'utente finale. Inoltre, non sarebbero in ogni caso disponibili gli errori delle stime non pubblicate, che l'utente può ricavare in modo autonomo.

Per questi motivi, generalmente, si ricorre ad una presentazione sintetica degli errori relativi, basata sul *metodo dei modelli regressivi*. Tale metodo si basa sulla determinazione di una funzione matematica che mette in relazione ciascuna stima con il proprio errore relativo.

Il modello utilizzato per le stime di frequenze assolute è il seguente:

$$\log \hat{\varepsilon}^2(\hat{Y}_d) = a + b \log(\hat{Y}_d) \quad (10)$$

dove i parametri a e b vengono stimati mediante il metodo dei minimi quadrati.

Nella presente indagine è stato stimato un modello di tipo (10) per ciascuno dei seguenti domini di interesse:

- D1. totale Italia e totale Italia distintamente per sesso;
- D2. ripartizioni geografiche di residenza in totale e distintamente per sesso;
- D3. gruppi di corsi di laurea in totale e distintamente per sesso.

Per calcolare gli errori di campionamento è stato utilizzato un software generalizzato, messo a punto dall'Istat, che consente di calcolare per le variabili prescelte tra quelle rilevate, oltre agli errori di campionamento assoluti e relativi, anche gli intervalli di confidenza; inoltre, permette di costruire modelli regressivi del tipo (10) per la presentazione sintetica degli errori di campionamento.

Il prospetto 2 riporta i valori dei coefficienti a e b e dell'indice di determinazione R^2 del modello utilizzato per l'interpolazione degli errori campionari delle stime riferite ai domini D1 e D2. Informazioni analoghe sono riportate nel prospetto 4 per il dominio D3.

Sulla base delle informazioni contenute nei suddetti prospetti è possibile calcolare l'errore relativo di una determinata stima di frequenza assoluta \hat{Y}_d^* , riferita ai domini D1 e D2, e al dominio D3, nel modo di seguito descritto.

Dalla (10) mediante semplici passaggi si ricava:

$$\hat{\varepsilon}(\hat{Y}_d^*) = \sqrt{\exp(a + b \log(\hat{Y}_d^*))}. \quad (11)$$

Prospetto 2 - Valori dei coefficienti a e b e dell'indice di determinazione R^2 (%) del modello per l'interpolazione degli errori campionari delle stime riferite ai laureati per sesso e ripartizione geografica di residenza

| DOMINIO DI STIMA Ripartizione per sesso | | a | b | R^2 |
|--|----------------------|-----------------------|------------------------|---------------------|
| NORD-OVEST | Maschi | 1,58003 | -0,98894 | 98,11 |
| | Femmine | 1,87635 | -1,03539 | 99,34 |
| | <i>Totale</i> | <i>1,88611</i> | <i>-1,03281</i> | <i>99,41</i> |
| NORD-EST | Maschi | 1,39548 | -0,94932 | 96,51 |
| | Femmine | 1,57496 | -0,97869 | 97,92 |
| | <i>Totale</i> | <i>1,60933</i> | <i>-0,98438</i> | <i>98,14</i> |
| CENTRO | Maschi | 1,33204 | -0,94031 | 97,88 |
| | Femmine | 1,77546 | -1,00252 | 98,99 |
| | <i>Totale</i> | <i>1,73249</i> | <i>-0,99919</i> | <i>99,26</i> |
| SUD | Maschi | 1,43849 | -0,94019 | 96,90 |
| | Femmine | 1,84330 | -1,00207 | 97,19 |
| | <i>Totale</i> | <i>1,82618</i> | <i>-1,00036</i> | <i>99,08</i> |
| ISOLE | Maschi | 1,29948 | -0,93602 | 95,09 |
| | Femmine | 1,57443 | -0,98344 | 97,96 |
| | <i>Totale</i> | <i>1,70527</i> | <i>-1,00645</i> | <i>98,57</i> |
| TOTALE | Maschi | 1,67208 | -1,00019 | 98,51 |
| | Femmine | 1,88449 | -1,02344 | 99,06 |
| | <i>Totale</i> | <i>1,84709</i> | <i>-1,01849</i> | <i>99,02</i> |

Se, per esempio, la generica stima \hat{Y}_d^* si riferisce agli individui di sesso maschile dell'Italia Centrale, è possibile introdurre nella (11) i valori dei parametri a e b ($a=1,33204$, $b=-0,94031$) riportati nella settima riga del prospetto 2 e ricavare il corrispondente errore relativo.

Una volta calcolato l'errore relativo è possibile costruire l'intervallo di confidenza al 95% come:

$$\left\{ \hat{Y}_d^* - 1,96 \cdot \hat{\varepsilon}(\hat{Y}_d^*) \cdot \hat{Y}_d^* ; \hat{Y}_d^* + 1,96 \cdot \hat{\varepsilon}(\hat{Y}_d^*) \cdot \hat{Y}_d^* \right\}.$$

Allo scopo di facilitare il calcolo degli errori campionari per le stime che si riferiscono alle 5 ripartizioni geografiche e al totale Italia, nel prospetto 3 sono riportati i valori interpolati degli errori di campionamento relativi di alcune stime di frequenze assolute riferite al totale dei laureati.

Il prospetto 5, invece, contiene, per ciascun gruppo di corsi di laurea e per valori crescenti di stime di frequenze percentuali, i corrispondenti valori assoluti delle stime e i valori degli errori relativi percentuali, calcolati introducendo nella (11) i valori di a e b riportati nel prospetto 4.

Per ciascun dominio il suddetto prospetto ha la seguente struttura:

| Stime di frequenze percentuali | Stime di frequenze assolute | Errori relativi percentuali interpolati |
|---------------------------------------|------------------------------------|--|
| 1% | $\hat{Y}_d(1)$ | $\hat{\varepsilon}[\hat{Y}_d(1)]$ |
| 2% | $\hat{Y}_d(2)$ | $\hat{\varepsilon}[\hat{Y}_d(2)]$ |
| | | |
| 50% | $\hat{Y}_d(50)$ | $\hat{\varepsilon}[\hat{Y}_d(50)]$ |

Le informazioni contenute nei prospetti 3 e 5 consentono di calcolare l'errore relativo di una generica stima di frequenza assoluta o relativa mediante due procedimenti che risultano di facile applicazione, anche se conducono a risultati meno precisi di quelli ottenibili applicando direttamente la formula (11).

Il primo metodo consiste nell'approssimare l'errore relativo della stima di interesse con quello, riportato nei prospetti, corrispondente al livello di stima che più vi si avvicina.

Il secondo metodo, più preciso del primo, si basa sull'uso di una formula di interpolazione lineare per il calcolo degli errori di stime non comprese tra i valori forniti nei prospetti. In tal caso, l'errore campionario della stima \hat{Y}_d^* , si ricava mediante l'espressione:

$$\hat{\varepsilon}(\hat{Y}_d^*) = \hat{\varepsilon}(\hat{Y}_d^{k-1}) + \frac{\hat{\varepsilon}(\hat{Y}_d^{k-1}) - \hat{\varepsilon}(\hat{Y}_d^k)}{\hat{Y}_d^k - \hat{Y}_d^{k-1}} (\hat{Y}_d^* - \hat{Y}_d^{k-1}) \quad (12)$$

dove \hat{Y}_d^{k-1} e \hat{Y}_d^k sono i valori delle stime entro i quali è compresa la stima \hat{Y}_d^* , mentre $\hat{\varepsilon}(\hat{Y}_d^{k-1})$ e $\hat{\varepsilon}(\hat{Y}_d^k)$ sono i corrispondenti errori relativi.

Prospetto 3 - Valori interpolati degli errori campionari relativi percentuali delle stime riferite al totale dei laureati per ripartizione geografica di residenza

| STIMA | RIPARTIZIONE GEOGRAFICA | | | | | Italia |
|--------|-------------------------|----------|--------|------|-------|--------|
| | Nord-Ovest | Nord-Est | Centro | Sud | Isole | |
| 100 | 23,8 | 23,2 | 23,8 | 24,9 | 23,1 | 24,1 |
| 200 | 16,6 | 16,5 | 16,9 | 17,6 | 16,3 | 17,0 |
| 300 | 13,5 | 13,5 | 13,8 | 14,4 | 13,3 | 13,8 |
| 400 | 11,6 | 11,7 | 11,9 | 12,4 | 11,5 | 11,9 |
| 500 | 10,4 | 10,5 | 10,7 | 11,1 | 10,3 | 10,6 |
| 600 | 9,4 | 9,6 | 9,7 | 10,2 | 9,4 | 9,7 |
| 700 | 8,7 | 8,9 | 9,0 | 9,4 | 8,7 | 9,0 |
| 800 | 8,1 | 8,3 | 8,4 | 8,8 | 8,1 | 8,4 |
| 900 | 7,7 | 7,9 | 7,9 | 8,3 | 7,6 | 7,9 |
| 1.000 | 7,3 | 7,5 | 7,5 | 7,9 | 7,3 | 7,5 |
| 1.500 | 5,9 | 6,1 | 6,2 | 6,4 | 5,9 | 6,1 |
| 2.000 | 5,1 | 5,3 | 5,3 | 5,6 | 5,1 | 5,2 |
| 2.500 | 4,5 | 4,8 | 4,8 | 5,0 | 4,6 | 4,7 |
| 5.000 | 3,2 | 3,4 | 3,4 | 3,5 | 3,2 | 3,3 |
| 7.500 | 2,6 | 2,8 | 2,8 | 2,9 | 2,6 | 2,7 |
| 10.000 | 2,2 | 2,4 | 2,4 | 2,5 | - | 2,3 |
| 15.000 | 1,8 | 2,0 | 1,9 | 2,0 | - | 1,9 |
| 20.000 | 1,5 | - | - | - | - | 1,6 |
| 25.000 | - | - | - | - | - | 1,5 |
| 50.000 | - | - | - | - | - | 1,0 |
| 75.000 | - | - | - | - | - | 0,8 |

Prospetto 4 - Valori dei coefficienti a e b e dell'indice di determinazione R² (%) del modello per l'interpolazione degli errori campionari delle stime riferite ai laureati per sesso e gruppo di corsi di laurea

| DOMINIO DI STIMA Gruppo di corsi per sesso | | a | b | R ² |
|---|---------------|----------------|-----------------|----------------|
| Gruppo scientifico | Maschi | 0,90391 | -0,91734 | 90,64 |
| | Femmine | 0,98760 | -0,96495 | 95,21 |
| | <i>Totale</i> | <i>0,93732</i> | <i>-0,94257</i> | <i>94,04</i> |
| Gruppo chimico-farmaceutico | Maschi | 1,08887 | -1,04554 | 96,15 |
| | Femmine | 1,26476 | -0,99758 | 96,26 |
| | <i>Totale</i> | <i>1,20633</i> | <i>-1,00904</i> | <i>97,31</i> |
| Gruppo geo-biologico | Maschi | 1,09098 | -0,98729 | 97,39 |
| | Femmine | 1,46771 | -1,00977 | 98,45 |
| | <i>Totale</i> | <i>1,28869</i> | <i>-0,99165</i> | <i>98,30</i> |
| Gruppo medico | Maschi | 1,71767 | -1,08378 | 96,84 |
| | Femmine | 1,72952 | -1,08489 | 96,09 |
| | <i>Totale</i> | <i>1,74470</i> | <i>-1,07716</i> | <i>97,40</i> |
| Gruppo ingegneria | Maschi | 1,81429 | -1,01149 | 96,58 |
| | Femmine | 1,16577 | -1,04480 | 95,62 |
| | <i>Totale</i> | <i>1,84790</i> | <i>-1,03116</i> | <i>98,00</i> |
| Gruppo architettura | Maschi | 1,54061 | -0,97718 | 93,97 |
| | Femmine | 2,04085 | -1,07428 | 99,06 |
| | <i>Totale</i> | <i>1,60225</i> | <i>-0,98934</i> | <i>95,88</i> |
| Gruppo agrario | Maschi | 0,98103 | -1,00295 | 96,68 |
| | Femmine | 0,63484 | -0,99603 | 95,88 |
| | <i>Totale</i> | <i>0,87138</i> | <i>-1,00446</i> | <i>97,39</i> |
| Gruppo economico-statistico | Maschi | 1,43251 | -0,91495 | 92,72 |
| | Femmine | 1,79869 | -0,98498 | 96,83 |
| | <i>Totale</i> | <i>1,79258</i> | <i>-0,97967</i> | <i>97,03</i> |
| Gruppo politico-sociale | Maschi | 1,90643 | -1,04468 | 98,64 |
| | Femmine | 1,90989 | -1,03728 | 99,10 |
| | <i>Totale</i> | <i>1,93435</i> | <i>-1,03995</i> | <i>99,19</i> |
| Gruppo giuridico | Maschi | 2,01921 | -1,01504 | 97,18 |
| | Femmine | 2,17363 | -1,03975 | 99,54 |
| | <i>Totale</i> | <i>2,22539</i> | <i>-1,04428</i> | <i>99,55</i> |
| Gruppo letterario | Maschi | 1,43140 | -0,99331 | 96,54 |
| | Femmine | 1,99423 | -1,02883 | 99,15 |
| | <i>Totale</i> | <i>1,81190</i> | <i>-1,00991</i> | <i>98,98</i> |
| Gruppo linguistico | Maschi | 0,53789 | -1,06020 | 96,57 |
| | Femmine | 1,96934 | -1,02754 | 98,26 |
| | <i>Totale</i> | <i>1,43319</i> | <i>-0,94765</i> | <i>95,79</i> |
| Gruppo insegnamento | Maschi | -0,64290 | -0,93132 | 82,03 |
| | Femmine | 1,62007 | -1,02059 | 97,56 |
| | <i>Totale</i> | <i>1,06219</i> | <i>-0,93125</i> | <i>92,94</i> |
| Gruppo psicologico | Maschi | -0,11248 | -1,02808 | 98,06 |
| | Femmine | 1,45928 | -1,04945 | 99,53 |
| | <i>Totale</i> | <i>0,96650</i> | <i>-0,97990</i> | <i>96,91</i> |

Prospetto 5 - Valori interpolati degli errori campionari delle stime riferite al totale dei laureati per gruppo di corsi di laurea

| PERC. (%) | Gruppo scientifico | | Gruppo chimico-farmaceutico | | Gruppo geo-biologico | | Gruppo medico | | Gruppo ingegneria | |
|--------------|---------------------------|---------------------------|-----------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| | Stima valore assol. | Errore relat. perc. | Stima valore assol. | Errore relat. perc. | Stima valore assol. | Errore relat. perc. | Stima valore assol. | Errore relat. perc. | Stima valore assol. | Errore relat. perc. |
| 1 | 50 | 25,3 | 44 | 27,0 | 54 | 26,3 | 70 | 24,3 | 147 | 19,3 |
| 2 | 100 | 18,2 | 88 | 19,0 | 109 | 18,6 | 140 | 16,7 | 293 | 13,5 |
| 3 | 150 | 15,1 | 133 | 15,5 | 163 | 15,2 | 210 | 13,4 | 440 | 10,9 |
| 4 | 200 | 13,2 | 177 | 13,4 | 217 | 13,2 | 280 | 11,5 | 586 | 9,4 |
| 5 | 250 | 11,8 | 221 | 12,0 | 271 | 11,8 | 350 | 10,2 | 733 | 8,4 |
| 6 | 300 | 10,9 | 265 | 10,9 | 326 | 10,8 | 420 | 9,2 | 879 | 7,6 |
| 7 | 350 | 10,1 | 310 | 10,1 | 380 | 10,0 | 490 | 8,5 | 1.026 | 7,1 |
| 8 | 399 | 9,5 | 354 | 9,5 | 434 | 9,4 | 560 | 7,9 | 1.172 | 6,6 |
| 9 | 449 | 9,0 | 398 | 8,9 | 489 | 8,8 | 630 | 7,4 | 1.319 | 6,2 |
| 10 | 499 | 8,5 | 442 | 8,5 | 543 | 8,4 | 700 | 7,0 | 1.465 | 5,9 |
| 15 | 749 | 7,1 | 663 | 6,9 | 814 | 6,9 | 1.050 | 5,6 | 2.198 | 4,8 |
| 20 | 999 | 6,2 | 884 | 6,0 | 1.086 | 6,0 | 1.400 | 4,8 | 2.931 | 4,1 |
| 25 | 1.248 | 5,6 | 1.106 | 5,3 | 1.357 | 5,3 | 1.750 | 4,3 | 3.664 | 3,7 |
| 30 | 1.498 | 5,1 | 1.327 | 4,9 | 1.629 | 4,9 | 2.100 | 3,9 | 4.396 | 3,3 |
| 35 | 1.748 | 4,7 | 1.548 | 4,5 | 1.900 | 4,5 | 2.450 | 3,6 | 5.129 | 3,1 |
| 40 | 1.997 | 4,4 | 1.769 | 4,2 | 2.172 | 4,2 | 2.800 | 3,3 | 5.862 | 2,9 |
| 45 | 2.247 | 4,2 | 1.990 | 4,0 | 2.443 | 4,0 | 3.150 | 3,1 | 6.594 | 2,7 |
| 50 | 2.497 | 4,0 | 2.211 | 3,8 | 2.715 | 3,8 | 3.500 | 3,0 | 7.327 | 2,6 |

Prospetto 5 segue - Valori interpolati degli errori campionari delle stime riferite al totale dei laureati per gruppo di corsi di laurea

| PERC. (%) | Gruppo architettura | | Gruppo agrario | | Gruppo economico- statistico | | Gruppo politico-sociale | | Gruppo giuridico | |
|--------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| | Stima valore assol. | Errore relat. perc. | Stima valore assol. | Errore relat. perc. | Stima valore assol. | Errore relat. perc. | Stima valore assol. | Errore relat. perc. | Stima valore assol. | Errore relat. perc. |
| 1 | 76 | 26,1 | 26 | 30,1 | 247 | 16,5 | 101 | 23,9 | 195 | 19,4 |
| 2 | 153 | 18,5 | 52 | 21,2 | 493 | 11,8 | 201 | 16,7 | 390 | 13,5 |
| 3 | 229 | 15,1 | 78 | 17,3 | 740 | 9,6 | 302 | 13,5 | 585 | 10,9 |
| 4 | 306 | 13,1 | 104 | 15,0 | 987 | 8,4 | 402 | 11,6 | 780 | 9,4 |
| 5 | 382 | 11,8 | 130 | 13,4 | 1.233 | 7,5 | 503 | 10,4 | 975 | 8,4 |
| 6 | 459 | 10,7 | 156 | 12,2 | 1.480 | 6,9 | 603 | 9,4 | 1.169 | 7,6 |
| 7 | 535 | 10,0 | 182 | 11,3 | 1.727 | 6,4 | 704 | 8,7 | 1.364 | 7,0 |
| 8 | 612 | 9,3 | 208 | 10,6 | 1.973 | 6,0 | 804 | 8,1 | 1.559 | 6,5 |
| 9 | 688 | 8,8 | 234 | 10,0 | 2.220 | 5,6 | 905 | 7,6 | 1.754 | 6,2 |
| 10 | 765 | 8,3 | 260 | 9,5 | 2.467 | 5,3 | 1.006 | 7,2 | 1.949 | 5,8 |
| 15 | 1.147 | 6,8 | 390 | 7,7 | 3.700 | 4,4 | 1.508 | 5,9 | 2.924 | 4,7 |
| 20 | 1.530 | 5,9 | 520 | 6,7 | 4.933 | 3,8 | 2.011 | 5,0 | 3.898 | 4,1 |
| 25 | 1.912 | 5,3 | 651 | 6,0 | 6.167 | 3,4 | 2.514 | 4,5 | 4.873 | 3,6 |
| 30 | 2.294 | 4,8 | 781 | 5,5 | 7.400 | 3,1 | 3.017 | 4,1 | 5.847 | 3,3 |
| 35 | 2.677 | 4,5 | 911 | 5,0 | 8.633 | 2,9 | 3.520 | 3,8 | 6.822 | 3,0 |
| 40 | 3.059 | 4,2 | 1.041 | 4,7 | 9.866 | 2,7 | 4.022 | 3,5 | 7.796 | 2,8 |
| 45 | 3.442 | 4,0 | 1.171 | 4,4 | 11.100 | 2,6 | 4.525 | 3,3 | 8.771 | 2,7 |
| 50 | 3.824 | 3,8 | 1.301 | 4,2 | 12.333 | 2,4 | 5.028 | 3,1 | 9.745 | 2,5 |

Prospetto 5 segue - Valori interpolati degli errori campionari delle stime riferite al totale dei laureati per gruppo di corsi di laurea

| PERC. (%) | Gruppo letterario | | Gruppo linguistico | | Gruppo insegnamento | | Gruppo psicologico | |
|--------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| | Stima valore assol. | Errore relat. perc. | Stima valore assol. | Errore relat. perc. | Stima valore assol. | Errore relat. perc. | Stima valore assol. | Errore relat. perc. |
| 1 | 129 | 21,2 | 79 | 25,9 | 41 | 30,1 | 34 | 28,7 |
| 2 | 259 | 15,0 | 157 | 18,6 | 83 | 21,8 | 68 | 20,5 |
| 3 | 388 | 12,2 | 236 | 15,4 | 124 | 18,0 | 103 | 16,8 |
| 4 | 517 | 10,5 | 315 | 13,4 | 165 | 15,8 | 137 | 14,6 |
| 5 | 647 | 9,4 | 393 | 12,1 | 206 | 14,2 | 171 | 13,1 |
| 6 | 776 | 8,6 | 472 | 11,1 | 248 | 13,1 | 205 | 11,9 |
| 7 | 905 | 8,0 | 551 | 10,3 | 289 | 12,2 | 239 | 11,1 |
| 8 | 1.035 | 7,4 | 630 | 9,7 | 330 | 11,4 | 274 | 10,4 |
| 9 | 1.164 | 7,0 | 708 | 9,1 | 371 | 10,8 | 308 | 9,8 |
| 10 | 1.293 | 6,6 | 787 | 8,7 | 413 | 10,3 | 342 | 9,3 |
| 15 | 1.940 | 5,4 | 1.180 | 7,2 | 619 | 8,5 | 513 | 7,6 |
| 20 | 2.587 | 4,7 | 1.574 | 6,3 | 825 | 7,5 | 684 | 6,6 |
| 25 | 3.233 | 4,2 | 1.967 | 5,6 | 1.032 | 6,7 | 855 | 5,9 |
| 30 | 3.880 | 3,8 | 2.361 | 5,2 | 1.238 | 6,2 | 1.026 | 5,4 |
| 35 | 4.527 | 3,5 | 2.754 | 4,8 | 1.444 | 5,7 | 1.197 | 5,0 |
| 40 | 5.173 | 3,3 | 3.148 | 4,5 | 1.650 | 5,4 | 1.368 | 4,7 |
| 45 | 5.820 | 3,1 | 3.541 | 4,3 | 1.857 | 5,1 | 1.539 | 4,4 |
| 50 | 6.467 | 2,9 | 3.935 | 4,1 | 2.063 | 4,9 | 1.710 | 4,2 |

ALLEGATO E

METODOLOGIA PER LA STIMA DEL RISCHIO DI VIOLAZIONE DELLA RISERVATEZZA

Per la protezione della riservatezza dei rispondenti all'indagine sull'Inserimento professionale dei laureati nell'anno 1998, l'Istat ha adottato, come per la precedente indagine relativa all'anno 1995, una metodologia basata su una definizione di rischio di violazione della riservatezza a livello individuale. Tale approccio prevede la stima di un rischio per ogni singolo individuo coinvolto nell'indagine che si basa sulla combinazione di modalità che esso assume nell'insieme delle variabili identificative indirette (ovvero che potenzialmente permettono l'identificazione) e sulle informazioni del disegno campionario.

La definizione di un rischio a livello individuale permette di adottare una tecnica di protezione maggiormente mirata al controllo di quegli individui che presentano un più elevato rischio di identificazione. A tale scopo alla consueta tecnica di *ricodifica globale* è stata affiancata quella di *soppressione locale*.

Ricordiamo che la tecnica di ricodifica globale consiste nella ricodifica di una variabile unendo più modalità in una sola mentre, con la soppressione locale, i valori di una variabile vengono sostituiti, in una o più unità, con un valore mancante. Una differenza sostanziale tra dette tecniche di protezione sta nel fatto che mentre la prima (ricodifica globale) viene applicata sull'intero insieme di dati la seconda (soppressione locale) viene applicata ad una parte dei dati.

In particolare, stimato il rischio di identificazione per tutti gli individui coinvolti nell'indagine e fissata una soglia (al di sotto della quale il rischio è considerato accettabile) la tecnica di soppressione locale è stata applicata a tutti quei record il cui rischio supera detta soglia.

Tale tecnica è stata applicata utilizzando un procedura di ottimizzazione che permette di minimizzare il numero di soppressioni necessarie affinché il rischio di violazione di ciascun individuo non superi la soglia prefissata. Ciò ha permesso un ulteriore miglioramento della tutela della riservatezza delle unità effettivamente a rischio e il rilascio di un maggiore contenuto informativo sia per le altre unità che per le stesse unità a rischio.

I prospetti allegati presentano in maniera sintetica l'impatto delle soppressioni sul contenuto informativo del file.

Prospetto 5 - Distribuzione dei record per numero di soppressioni

| Record che presentano s soppressioni (percentuale sul totale=20.844) | | | |
|---|-------------------|----------------|--------------|
| s=0 | s=1 | s=2 | s=3 |
| 12.409 (59,96%) | 8.029 (38,52%) | 314 (1,51%) | 2 (0,01%) |

Prospetto 6 - Distribuzione del numero di soppressione per variabile coinvolta.

| Variabili | Corso di laurea | Sede universitaria | Sesso | Residenza | Età |
|---|-------------------|--------------------|----------------|------------------|-------------------|
| Numero di soppressioni (percentuale sul totale=20.844) | 2.503 (12,01%) | 160 (0,77%) | 423 (2,03%) | 1.210 (5,81%) | 4.367 (20,95%) |