

Rappresentatività campionaria delle indagini Banca d'Italia e Silc

Marco Savegnago

CAPPaper n. 46

giugno 2008

**RAPPRESENTATIVITÀ CAMPIONARIA DELLE INDAGINI BANCA
D'ITALIA E SILC.**

Marco Savegnago

Giugno 2008

INDICE

<i>INTRODUZIONE</i>	5
<i>INDICATORI</i>	6
<i>1) ANALISI PER ETÀ</i>	7
<i>2) ANALISI PER STATO CIVILE</i>	14
<i>3) ANALISI PER TITOLO DI STUDIO</i>	16
<i>4) ANALISI PER STATO OCCUPAZIONALE</i>	17
<i>5) ANALISI PER REDDITO FAMILIARE</i>	24
<i>6) CONFRONTO DEI REDDITI CON GLI AGGREGATI DELLA CONTABILITÀ NAZIONALE</i>	29
<i>7) TIPOLOGIE DI PENSIONI: PRESTAZIONI E IMPORTI</i>	31
<i>8) CONFRONTO DELLE PENSIONI CON GLI AGGREGATI DELLA CONTABILITÀ NAZIONALE (CN)</i>	33
<i>BIBLIOGRAFIA</i>	35

INDICE DELLE TABELLE

<i>Tabella 1: indici di dissomiglianza per distribuzione per età</i>	8
<i>Tabella 2: indici di dissomiglianza per classi di età quinquennali</i>	13
<i>Tabella 3: indici di struttura per area geografica</i>	14
<i>Tabella 4: indici di dissomiglianza per stato civile</i>	15
<i>Tabella 5: indici di dissomiglianza per titolo di studio</i>	16
<i>Tabella 6: indici di dissomiglianza per stato occupazionale</i>	18
<i>Tabella 7: numero di dipendenti, per sesso e classe di età</i>	20
<i>Tabella 8: numero di autonomi, per sesso e classe di età</i>	21
<i>Tabella 9: numero di disoccupati, per sesso e classe di età</i>	22
<i>Tabella 10: numero di non appartenenti alle forze lavoro, per sesso e classe d'età (15-64 anni)</i>	23
<i>Tabella 11: reddito familiare netto per tipologia</i>	24
<i>Tabella 12: confronto tra aggregati di reddito con i dati della Contabilità Nazionale (CN)</i>	30
<i>Tabella 13: numero di pensioni per tipologia. Confronto con i dati ISTAT-INPS</i>	31
<i>Tabella 14: importo medio della pensione per categoria. Confronto con i dati ISTAT-INPS</i>	32
<i>Tabella 15: confronto tra aggregati di reddito da pensione per categoria</i>	33

INDICE DELLE FIGURE

<i>Figura 1: distribuzione di frequenza per età. Banca d'Italia vs ISTAT.....</i>	<i>9</i>
<i>Figura 2: distribuzione di frequenza per età. SILC vs ISTAT</i>	<i>10</i>
<i>Figura 3: distribuzione di frequenza per classi d'età. Banca d'Italia vs ISTAT.....</i>	<i>11</i>
<i>Figura 4: distribuzione di frequenza per classi d'età. SILC vs ISTAT</i>	<i>12</i>
<i>Figura 5: distribuzione di frequenza per stato civile.....</i>	<i>15</i>
<i>Figura 6: distribuzione di frequenza per titolo di studio.....</i>	<i>17</i>
<i>Figura 7: distribuzione di frequenza per stato occupazionale</i>	<i>19</i>
<i>Figura 8: densità di frequenza del reddito familiare totale</i>	<i>25</i>
<i>Figura 9: "sfilata" Pen del reddito familiare totale</i>	<i>26</i>
<i>Figura 10: densità di frequenza del reddito familiare da lavoro dipendente.....</i>	<i>27</i>
<i>Figura 11: densità di frequenza del reddito familiare da lavoro autonomo.....</i>	<i>28</i>
<i>Figura 12: densità di frequenza del reddito familiare da pensione</i>	<i>29</i>

INTRODUZIONE

L'indagine IT-SILC, condotta a cadenza annuale dal 2004, è la componente italiana del progetto EU-SILC, volto a fornire dati ed indicatori utili allo studio, tra l'altro, della povertà e dell'esclusione sociale nei paesi membri dell'Unione. La versione definitiva dei dati relativi al 2004-2005 è stata resa nota nel novembre 2007: l'indagine campionaria ha coinvolto 22.032 nuclei familiari e 56.105 individui.

L'attenzione rivolta recentemente dai ricercatori a questa indagine riflette, tra l'altro, la speranza che essa possa essere meno affetta da distorsioni rispetto alla più nota Indagine sui Bilanci delle Famiglie Italiane, condotta a cadenza biennale dalla Banca d'Italia¹. Tali distorsioni sono legate, da un lato, al fatto che le famiglie intervistate sono estratte casualmente, dall'altro all'attendibilità delle risposte fornite.

In merito al primo punto, ad esempio, alcune caratteristiche socio-demografiche possono essere sovra o sottostimate rispetto alla popolazione "vera", per effetto delle procedure di campionamento². L'indagine SILC, intervistando più del doppio delle unità rispetto alla Banca d'Italia (8.000 famiglie e 21.000 individui), induce a sperare che il *bias* rispetto alla popolazione di riferimento sia più contenuto, grazie alla maggior dimensione campionaria: tuttavia, anticipando qualche risultato, l'analisi statistica condotta non mostra differenze di rilievo tra le due indagini.

In merito al secondo punto, rileva l'attendibilità delle risposte date agli intervistatori, ed assume particolare importanza per la stima delle variabili economiche, quali i redditi e la ricchezza. Le risposte possono essere poco precise sia per la difficoltà riscontrata nel rispondere a determinate domande, sia per una certa diffidenza nei confronti dell'intervistato (la reticenza campionaria colpisce soprattutto le stime delle attività finanziarie e delle case di non residenza, per la ricchezza, e i redditi da attività lavorative secondarie e da lavoro autonomo, per il reddito). Inoltre, non tutte le famiglie accettano di sottoporsi all'indagine: per BI si osserva come questo rifiuto sia più frequente tra le famiglie a maggior reddito. Se esse si sottraggono all'intervista, e se sono analizzate solo le famiglie meno ricche, il rischio è di disegnare un quadro distorto della distribuzione del reddito e della ricchezza. Benché questo fenomeno non sia (ancora) stato studiato per SILC, non ci dovrebbero essere ragioni

¹Un ulteriore elemento di superiorità di SILC risiede nel fatto che l'indagine è armonizzata a quelle svolte, con la medesima procedura di campionamento e con una base comune di domande, negli altri Paesi della Comunità Europea.

² Questa distorsione può essere parzialmente corretta operando (sul campione) una post-stratificazione rispetto ad alcune variabili-chiave.

per aspettarsi che tale comportamento reticente si manifesti solo nei confronti delle interviste commissionate dalla Banca d'Italia.

In questo lavoro si mettono a confronto, per alcune variabili demografiche ed economiche, le distribuzioni di frequenza ricavate dalle indagini SILC 2005 e BI 2004 con quelle di una distribuzione che fungerà da *benchmark*, e che sarà ricavata dai dati ISTAT disponibili on-line. Lo scopo finale dell'esercizio è determinare quale delle due indagini sia più rappresentativa della popolazione italiana o almeno, in mancanza di una risposta univoca, individuare i punti di forza e debolezza relativa delle due indagini.

INDICATORI

Nell'ambito della statistica descrittiva la maggiore o minore corrispondenza di due distribuzioni di frequenza può essere determinata ricorrendo ad indici di dissomiglianza. Essi possono essere relativi (variano da zero, nel caso che le due distribuzioni siano identiche, ad uno, che rappresenta il caso di massima disuguaglianza) o assoluti (in questo caso il valore massimo dipende dal numero di modalità con cui si distribuisce il carattere); possono riferirsi alle distribuzioni *non* cumulate o a quelle cumulate; possono essere semplici (se gli scarti non sono elevati ad alcun esponente) o di ordine r (laddove r è appunto la potenza applicata agli scarti).

In questo lavoro si farà riferimento ad indici *semplici, relativi*, e calcolati sia sulle funzioni di densità che su quelle cumulate. In particolare, consideriamo:

a) la differenza media

$$\Delta = \frac{1}{k} \sum_{i=1}^k |f_{Ai} - f_{Xi}|$$

b) l'indice relativo di dissomiglianza semplice calcolato sulle frequenze relative

$$z = \frac{1}{2} \sum_{i=1}^k |f_{Ai} - f_{Xi}|$$

c) l'indice relativo di dissomiglianza semplice calcolato sulle frequenze cumulate

$$Z = \frac{1}{k-1} \sum_{i=1}^{k-1} |F_{Ai} - F_{Xi}|$$

dove:

f e F identificano rispettivamente la densità di frequenza relativa e la densità di frequenza cumulata, k è il numero di modalità del carattere i , X si riferisce alla frequenza

benchmark (nel nostro caso, l'Istat) ed A è la distribuzione di frequenza di cui vogliamo conoscere la dissomiglianza (nel nostro caso sarà Banca d'Italia o SILC).

La differenza media è calcolata sommando, modalità per modalità, il valore assoluto dello scostamento tra le frequenze osservate, e poi dividendo per il numero di modalità con cui si distribuisce il carattere. Essa ci dice qual è lo scostamento medio delle due distribuzioni. Il secondo indice, molto simile al primo, è ottenuto dividendo la somma degli scostamenti per due, invece che per il numero di modalità. Questo allo scopo di ottenere un numero puro, che vari da zero ad uno, caso che si verifica se una distribuzione è tutta concentrata sulla modalità j e l'altra sulla modalità h , con $h \neq j$ ³. L'ultimo indice è concettualmente analogo al secondo, ma invece delle distribuzioni di frequenza semplici considera quelle cumulate. Esso cerca di cogliere la maggiore concentrazione della distribuzione verso le "code": il massimo teorico si ottiene se una distribuzione è tutta concentrata sulla prima modalità, e l'altra distribuzione è tutta concentrata sull'ultima modalità⁴. La scelta di questi indici si deve in parte alla loro semplicità (l'adozione d'indici complessi, con r superiore all'unità, avrebbero reso l'analisi più complicata senza aggiungere altre informazioni rilevanti) ed in parte alla loro intelligibilità (poiché sono numeri puri, possono essere interpretati in termini percentuali).

Le variabili studiate sono: età, stato civile, titolo di studio, condizione professionale o non professionale, categoria professionale e relativa retribuzione, categoria pensionistica e relativo beneficio.

1) ANALISI PER ETÀ

La struttura per età della popolazione è la prima variabile ad essere studiata. Da essa, infatti, dipende una serie d'altri fenomeni demografici e sociali.

Per ragioni di riservatezza, l'indagine SILC ha codificato tutti gli individui ultra-80enni, come aventi 80 anni. In assenza della possibilità di accedere ai dati della vera età degli ultraottantenni questo limite pregiudica fortemente la rappresentatività dell'indagine⁵. Ai soli fini della comparabilità e quindi per evidenti ragioni di coerenza la stessa ricodifica è stata eseguita sui dati BI e Istat.

³ In questo caso avremo: $z = \frac{1}{2} \sum_{i=1}^k |f_{Ai} - f_{Xi}| = \frac{1}{2} [|0-0| + \dots + |1-0| + \dots + |0-1| + \dots + |0-0|] = \frac{1}{2} \cdot 2 = 1$

⁴ In questo caso avremo: $Z = \frac{1}{k-1} \sum_{i=1}^{k-1} |F_{Ai} - F_{Xi}| = \frac{1}{k-1} [|1-0| + |1-0| + \dots + |1-1|] = \frac{1}{k-1} (k-1) = 1$

⁵ La quota di questi individui, di cui non si conosce l'età "reale, sul totale del campione è del 5%.

Il *benchmark* è costituito dalla popolazione italiana al 1° gennaio 2005, così come rilevata dagli uffici anagrafe dei Comuni⁶.

La tabella 1 mostra i risultati del confronto. In questa, come nelle successive tabelle, il secondo ed il terzo indice sono espressi in valori percentuali: essi oscillano quindi tra 0 e 100. La differenza media, invece, è un indice assoluto e non ha un valore massimo di riferimento.

Tabella 1: indici di dissomiglianza per distribuzione per età

DISSOMIGLIANZA PER ETÀ'			
INDICE		BI	SILC
Δ	Differenza media	0,156	0,165
z	Dissomiglianza semplice (%)	6,249	6,597
Z	Diss. Sempl. Cumulate (%)	0,641	0,894

Dal confronto degli indici non emergono importanti differenze tra le due distribuzioni campionarie rispetto ai dati ufficiale, sebbene quella della BI si mostri lievemente più precisa (6,2% contro il 6,6% per la dissomiglianza semplice, e 0,6% contro 0,9% per la distribuzione cumulata).

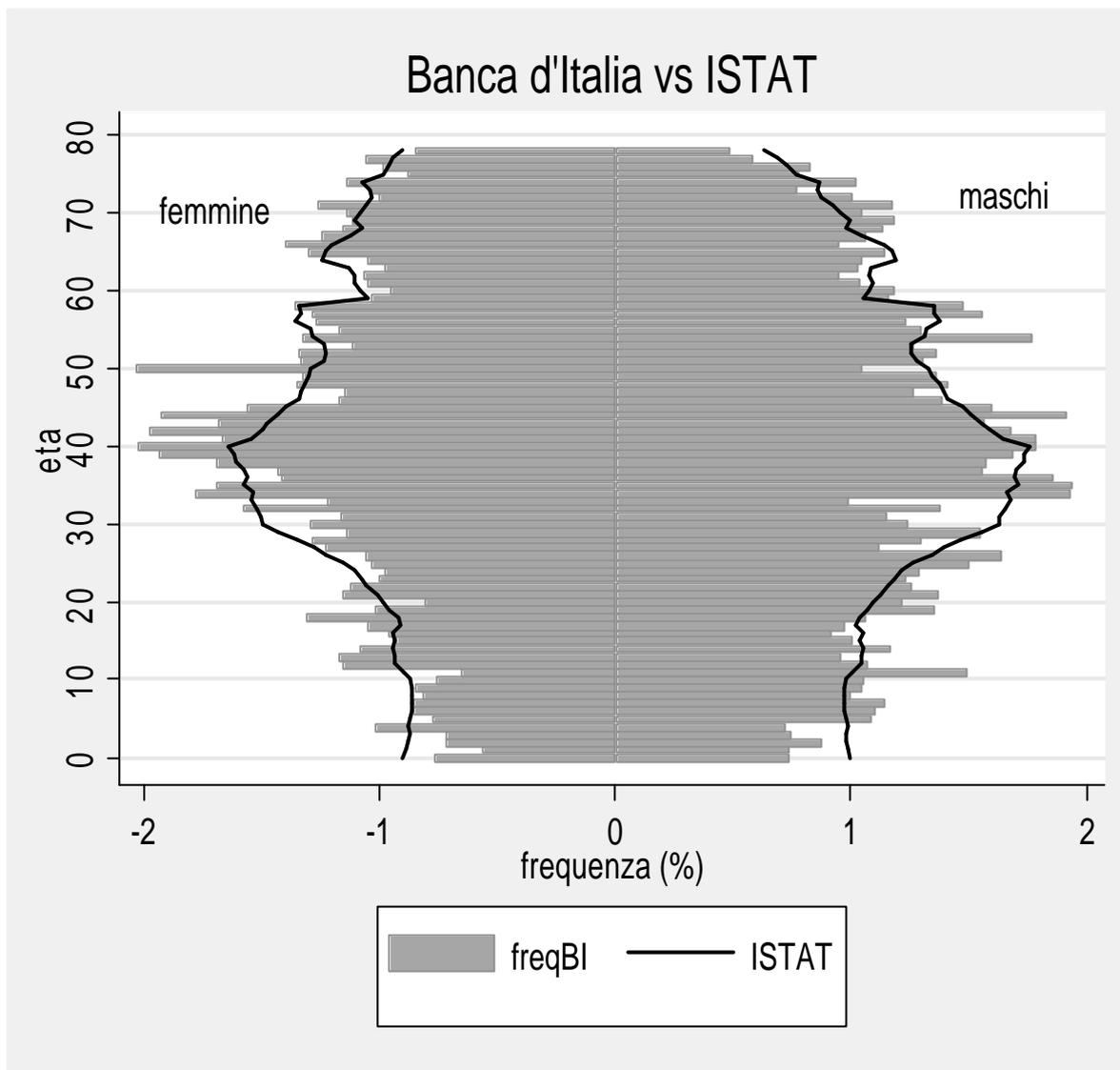
Nelle due figure successive sono riportate le distribuzioni di frequenza di BI e SILC rispetto ai dati Istat.

Iniziamo dall'indagine Bdl.

Il grafico è diviso in due sezioni: la parte destra, con valori positivi delle frequenze, si riferisce agli individui di sesso maschile; quella sinistra, con frequenze negative, agli individui di sesso femminile. Gli istogrammi grigi rappresentano le frequenze campionarie, la linea nera quelle della popolazione.

⁶ <http://demo.istat.it/pop2005/index.html>.

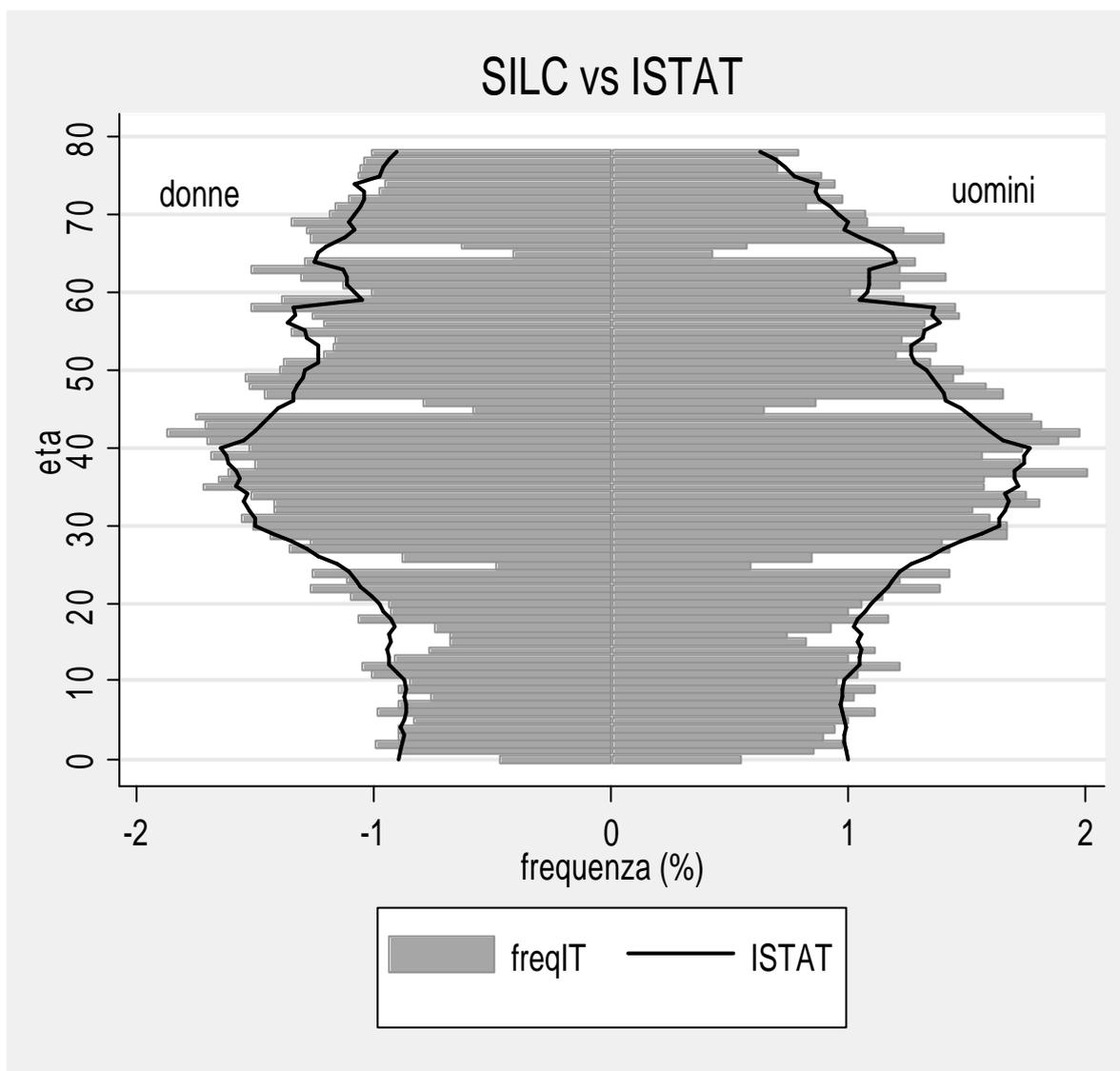
Figura 1: distribuzione di frequenza per età. Banca d'Italia vs ISTAT



E' possibile notare come per alcune età anagrafiche (soprattutto nelle fasce centrali), a causa del campionamento, le distribuzioni di frequenza campionarie siano decisamente lontane da quelle della popolazione di riferimento. Soprattutto, l'indagine BI sovrastima la frequenza delle donne intorno ai 40 anni, sottostima quella dei 30enni maschi e dei bambini fino a 10 anni di vita di entrambi i sessi. Sempre la distorsione da campionamento spiega l'andamento più "frastagliato" degli istogrammi, rispetto a quello più regolare della linea continua.

Passiamo all'indagine SILC.

Figura 2: distribuzione di frequenza per età. SILC vs ISTAT

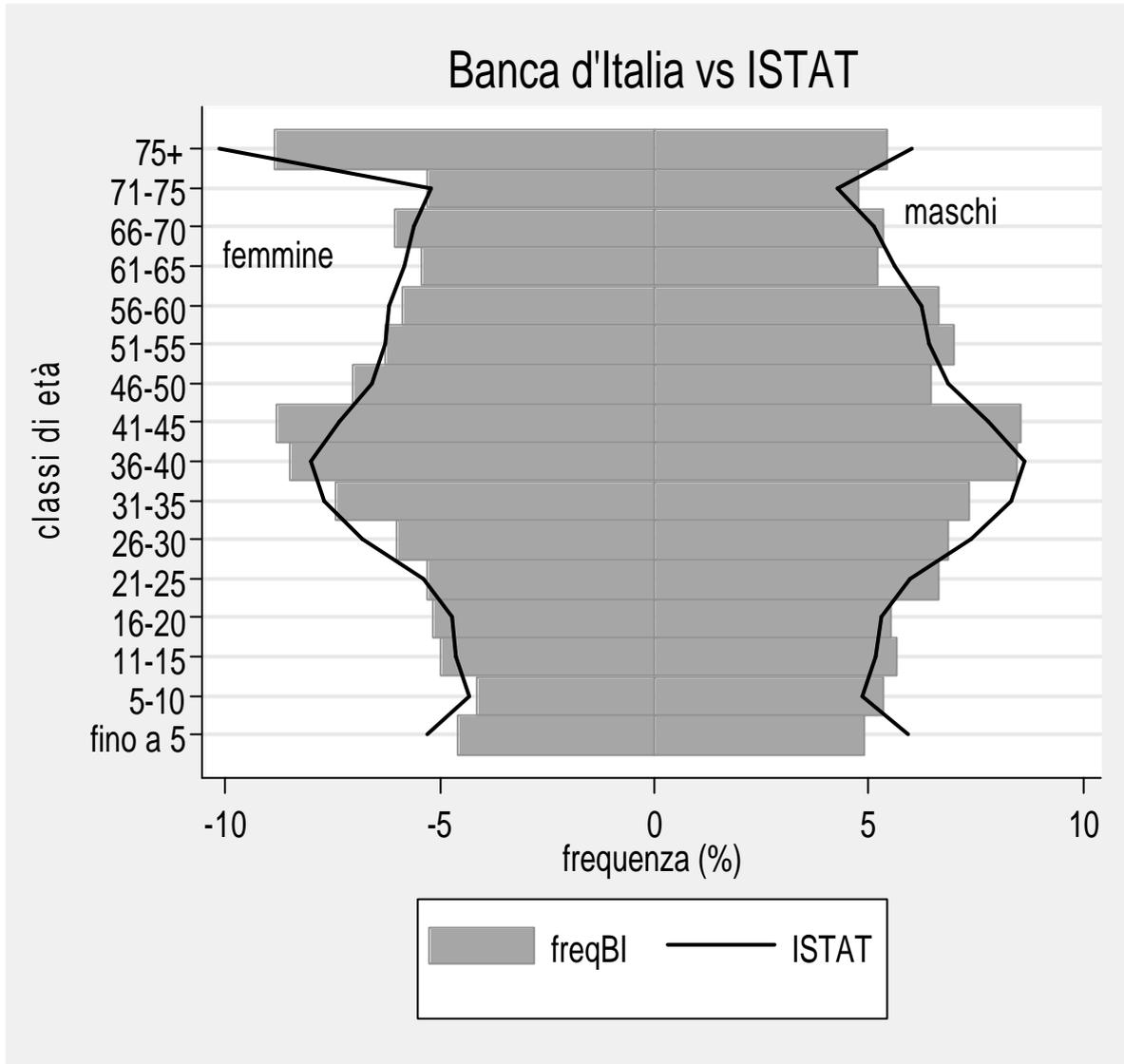


Dai due grafici precedenti non si evincono importanti differenze tra le indagini campionarie. Tuttavia, se SILC non sovrastima eccessivamente alcuna modalità, ne sottostima pesantemente altre. Si possono notare almeno tre “salti”, per entrambi i sessi, in corrispondenza dei 25, 45 e 65 anni. Questo probabilmente spiega perché, seppur di poco, gli indici di dissomiglianza di SILC siano maggiori di quelli Bdl.

Un altro modo per analizzare la struttura demografica della popolazione può essere quello di suddividere la stessa per classi di età, allo scopo di avere un’informazione più sintetica rispetto alla precedente. Così facendo, è possibile determinare quali classi di età sono sovra o sottorappresentate rispetto alla popolazione, senza dover scendere nel dettaglio di ogni singolo anno.

Come di consueto, iniziamo dall’indagine della Banca centrale.

Figura 3: distribuzione di frequenza per classi d'età. Banca d'Italia vs ISTAT

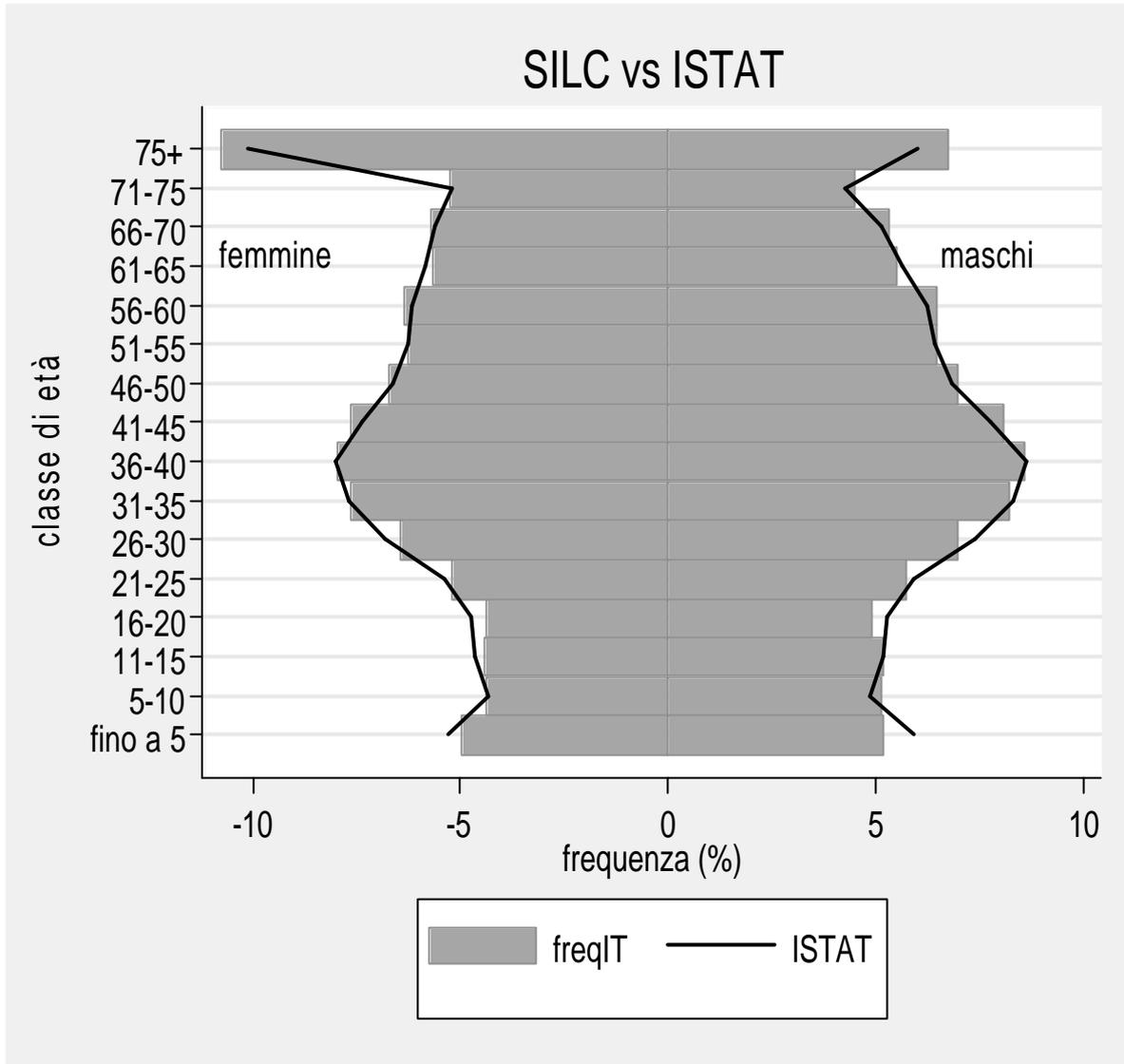


Il dato più evidente di questo grafico, che sintetizza quanto già intuito in precedenza, riguarda la sovrastima della percentuale di donne tra i 35 e i 45 anni, e la sottostima dell'ultima classe di età (quella aperta, dai 75 in poi). Per gli uomini, invece, si intuisce una certa sottostima nel gruppo di età 30-40 anni, anche se l'andamento generale del campione sembra combaciare discretamente con quello della popolazione.

Passiamo all'indagine dell'istituto di statistica⁷.

⁷ Naturalmente, l'ISTAT produce sia l'indagine SILC che le altre stime qui adottate come *benchmark*. Quando si farà riferimento ad "ISTAT", si intenderà come fornitore dei dati *benchmark*. Il riferimento come "istituto di statistica" è da intendere come autore dell'indagine campionaria SILC, analogamente alla banca centrale come autrice dell'indagine BI.

Figura 4: distribuzione di frequenza per classi d'età. SILC vs ISTAT



La suddivisione per classi rende le frequenze di questa seconda indagine molto simili a quelle dell'universo di riferimento e permette di identificare, con una certa precisione, il verso della distorsione da campionamento. Detto altrimenti, anche se di poco, si nota che la parte giovane della popolazione (cioè fino ai 40 anni) è sottorappresentata, mentre per quella più adulta accade il contrario.

Quanto intuito dagli ultimi due grafici può essere verificato con la prossima tabella, che mostra gli indici di disuguaglianza per la distribuzione in classi di età.

Tabella 2: indici di dissomiglianza per classi di età quinquennali

DISSOMOMIGLIANZA PER CLASSI DI ETA'			
INDICE		BI	SILC
Δ	Differenza media	0,506	0,225
z	Dissomiglianza semplice (%)	4,045	1,804
Z	Diss. Sempl. Cumulate (%)	0,635	0,952

Come anticipato dai grafici, dopo aver suddiviso la popolazione in classi di età quinquennali, l'indagine SILC ottiene risultati migliori in riferimento ai primi due indici: sia la differenza media che la dissomiglianza semplice sono circa la metà dei corrispondenti indici della BI, mentre la dissomiglianza sulle cumulate continua ad essere più elevata (probabilmente dovuta al fatto che SILC sovrastima l'ultima classe).

Un ultimo modo, ed è quello più sintetico, di analizzare la dimensione demografica è di far riferimento agli indici di struttura, come l'indice di vecchiaia (rapporto tra la popolazione ultra 65enne e la popolazione tra 0 e 14 anni), quello di dipendenza strutturale della popolazione (rapporto tra non attivi, cioè under 15 e over 65, e attivi, cioè con età 15-64), e quello di dipendenza degli anziani (rapporto tra popolazione con 65 anni o più e popolazione di 15-64 anni).

Il contenuto della tabella 3 appare molto interessante. Se gli indici di dissomiglianza per età lasciavano intravedere poche differenze tra le distribuzioni, gli indici di struttura fanno emergere un quadro più composito. I dati ISTAT rilevano circa 137 anziani per 100 bambini (indice di vecchiaia), laddove la Banca d'Italia ne conta 139 e SILC 141. Ancora una volta, l'indagine BI si rivelerebbe quella più precisa. Tuttavia, tale risultato finale è la media degli indici di vecchiaia calcolati nelle singole ripartizioni regionali: scendendo nel dettaglio, il campionamento della banca centrale sovrastima di molto il numero di anziani per 100 bambini al Centro (189 invece di 155) e li sottostima al Nord (140 invece di 158). Diversamente, il campionamento dell'istituto di statistica sovrastima leggermente questo indice al Centro ed al Sud.

Tabella 3: indici di struttura per area geografica

INDICI DI STRUTTURA PER L'ITALIA E PER AREA GEOGRAFICA						
INDICE	ITALIA			NORD		
	BI	SILC	ISTAT	BI	SILC	ISTAT
Vecchiaia (%)	139,2	141,4	136,9	139,9	158,2	158,4
Dipendenza Strutturale (%)	49,3	50,5	50,4	47,7	51	50,8
Dipendenza degli anziani (%)	28,7	29,5	29,1	27,8	31,3	31,2
INDICE	CENTRO			SUD		
	BI	SILC	ISTAT	BI	SILC	ISTAT
Vecchiaia (%)	189,1	165	155,5	119,4	112,6	106,5
Dipendenza Strutturale (%)	47,1	51,6	50,9	52,6	49,1	49,7
Dipendenza degli anziani (%)	30,8	32,1	31	28,6	26	25,6

Per quanto riguarda la dipendenza strutturale, l'indagine SILC replica perfettamente i dati della popolazione vera: a livello nazionale ci sono circa 50,5 inattivi (bambini ed anziani) per ogni 100 adulti, e le differenze nelle aree geografiche sono modeste. Infine, il numero di anziani per adulti. In questo caso, l'indice per la popolazione "vera" (29,1%) si colloca in mezzo a quello della Banca d'Italia (28,7%) e quello SILC (29,5%). Da notare che, entrambe le indagini, sottostimano questo indice di circa tre punti e mezzo nel Settentrione.

2) ANALISI PER STATO CIVILE

In questo caso abbiamo classificato gli individui per stato civile (celibe/nubile, coniugato, divorziato, vedovo) e per sesso. Il riferimento è sempre la popolazione residente al 1° gennaio 2005. Oltre a queste quattro modalità, SILC ne riporta altre due: separati di fatto, e separati legalmente: abbiamo ricodificato i primi come coniugati, ed i secondi come divorziati.

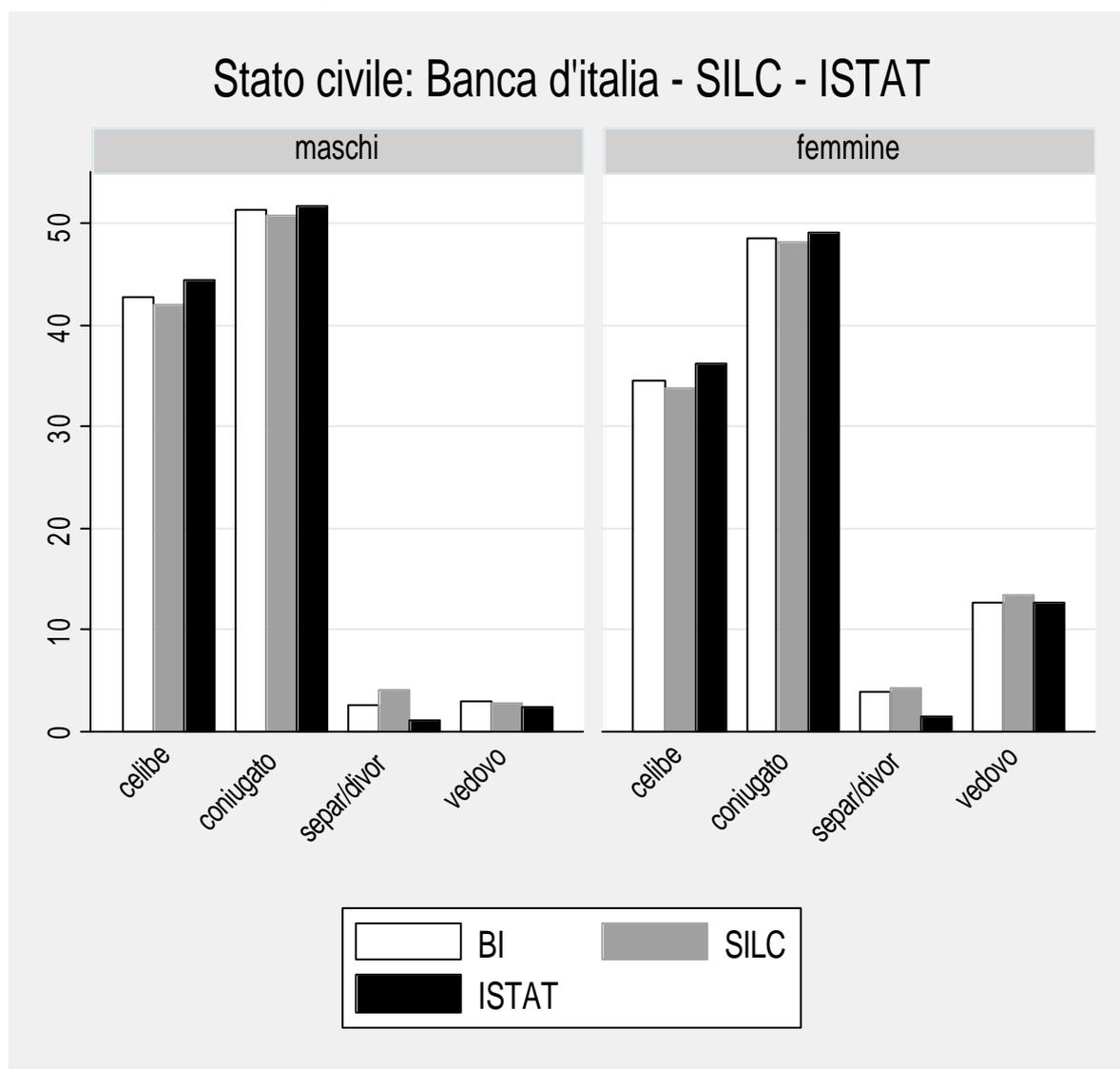
Tabella 4: indici di dissomiglianza per stato civile
DISSOMIGLIANZA PER STATO CIVILE

INDICE	BI	SILC
Differenza media	1,131	1,743
Dissomiglianza semplice (%)	2,263	3,487
Diss. Sempl. Cumulate (%)	1,431	2,178

Sebbene gli indici siano, in assoluto, molto contenuti per entrambe le indagini, è possibile individuare in BI quella più rappresentativa della popolazione: infatti, tutti e tre gli indicatori di dissomiglianza adottati sono nettamente inferiori a quelli dell'indagine SILC.

Ad un'analisi più attenta, si evince che la maggiore dissomiglianza di SILC è dovuta ad una sovrastima della quota di separati e divorziati (forse dovuta alla precedente ricodifica), sia per gli uomini che per le donne, sebbene anche BI sovrastimi questa categoria.

Figura 5: distribuzione di frequenza per stato civile



L'altra categoria per la quale si può apprezzare una certa differenza (seppur molto ridotta) tra i dati campionari e quelli dell'Istat è quella dei celibi. In questo caso le due indagini sottostimano la percentuale di questa tipologia di individui. Come in precedenza, anche se di poco, la distorsione dell'indagine SILC è inferiore a quella della banca centrale.

3) ANALISI PER TITOLO DI STUDIO

Questa variabile è stata codificata secondo cinque modalità:

- 1-senza titolo e scuola elementare
- 2-medie inferiori
- 3-medie superiori (2-3 anni, che non consentono l'iscrizione all'università)
- 4--medie superiori (4-5 anni)
- 5-laurea e titoli post-lauream

Il *benchmark* è costituito dalla Rilevazione Trimestrale sulle Forze di lavoro: i valori finali derivano dalla media dei quattro trimestri del 2005⁸, e fanno riferimento a tutti gli individui di 15 anni o più.

Tabella 5: indici di dissomiglianza per titolo di studio

DISSOMOMIGLIANZA PER TITOLO DI STUDIO			
INDICE		BI	SILC
Δ	Differenza media	0,796	0,789
z	Dissomiglianza semplice (%)	1,991	1,973
Z	Diss. Sempl. Cumulate (%)	0,960	0,558

La tabella indica indici molto simili, con valori modesti, per entrambe le indagini: in particolare, l'indice di dissomiglianza semplice è, nei due casi, inferiore al 2%. L'unica differenza di rilievo risiede nell'indice calcolato sulle cumulate: in questo caso, la dissomiglianza rispetto al *benchmark* dell'indagine SILC è circa la metà rispetto alla Banca d'Italia; tuttavia, entrambi sono inferiori all'1%, dimostrando ancora una volta come la distribuzione per titolo di studio nelle indagini campionarie sia sostanzialmente conforme a quella della popolazione di riferimento⁹.

Il grafico per gli uomini conferma quanto visto nella tabella, con istogrammi simili per ogni livello di studio. In particolare, guardando alla categoria di studio più alta, si nota una sottostima della BI rispetto al *benchmark*: questo contribuisce a spiegare perché l'indice di dissomiglianza sulle cumulate sia diverso nelle due indagini campionarie. Tale indice, infatti,

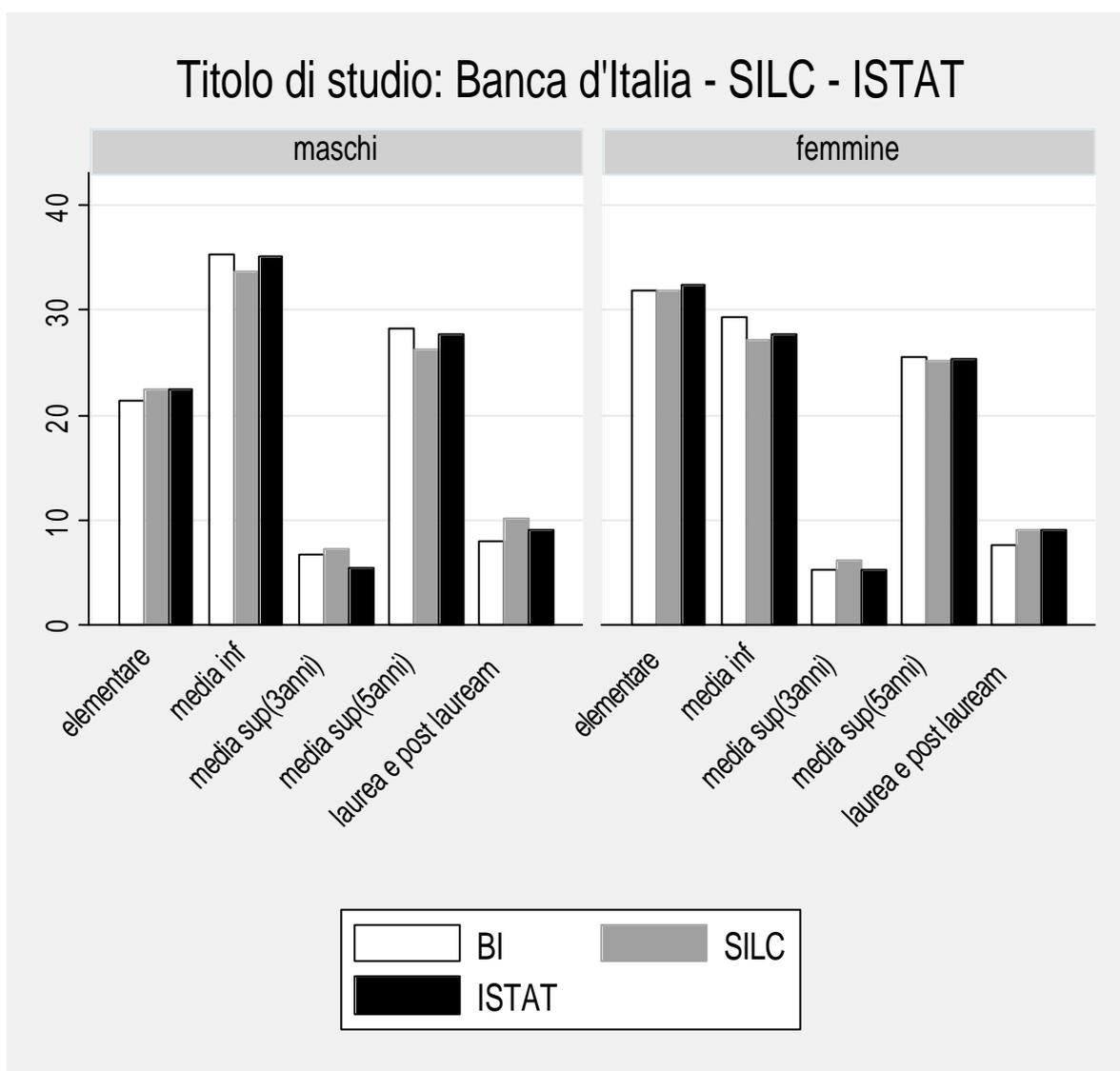
⁸ http://www.istat.it/salastampa/comunicati/non_calendario/20060418_00/.

⁹ Si noti che anche l'Indagine Trimestrale sulle Forze di lavoro è una indagine campionaria.

“esalta” la dissomiglianza quando essa si concentra sulle code della distribuzione. Come si vede nel secondo grafico, tale sottostima si ripete anche per le donne.

Un’importante differenza tra i due generi riguarda il livello di alcune frequenze: in particolare la quota di donne con licenza elementare supera il 30%, laddove quella degli uomini è di dieci punti inferiore. Tale divario è compensato dalle quote di chi detiene la licenza media e superiore (maggiori tra gli uomini), mentre la percentuale di laureati è identica per i due sessi.

Figura 6: distribuzione di frequenza per titolo di studio



4) ANALISI PER STATO OCCUPAZIONALE

Questa variabile è stata codificata secondo tre modalità:

- 1-occupato
- 2-disoccupato

3-non appartenente alla forza lavoro

Anche in questo caso, come per la variabile precedente, il *benchmark* è costituito dalla Rilevazione Trimestrale sulle Forze di lavoro. Pur essendo una classificazione piuttosto “grezza”, con tre sole modalità, essa permette confronti tra valori riferiti allo stesso periodo. L’alternativa sarebbe stata quella di adottare come punto di riferimento i dati del Censimento 2001: essi avrebbero consentito di scendere più nel dettaglio della condizione professionale, considerando altre categorie come gli studenti, le casalinghe o i pensionati; tuttavia, il confronto sarebbe stato probabilmente influenzato dai tre/quattro anni trascorsi tra il censimento e le indagini campionarie.

Tabella 6: indici di dissomiglianza per stato occupazionale
DISSOMOMIGLIANZA PER STATO OCCUPAZIONALE

INDICE		BI	SILC
Δ	Differenza media	1,504	1,220
z	Dissomiglianza semplice (%)	2,256	1,830
Z	Diss. Sempl. Cumulate (%)	1,128	0,965

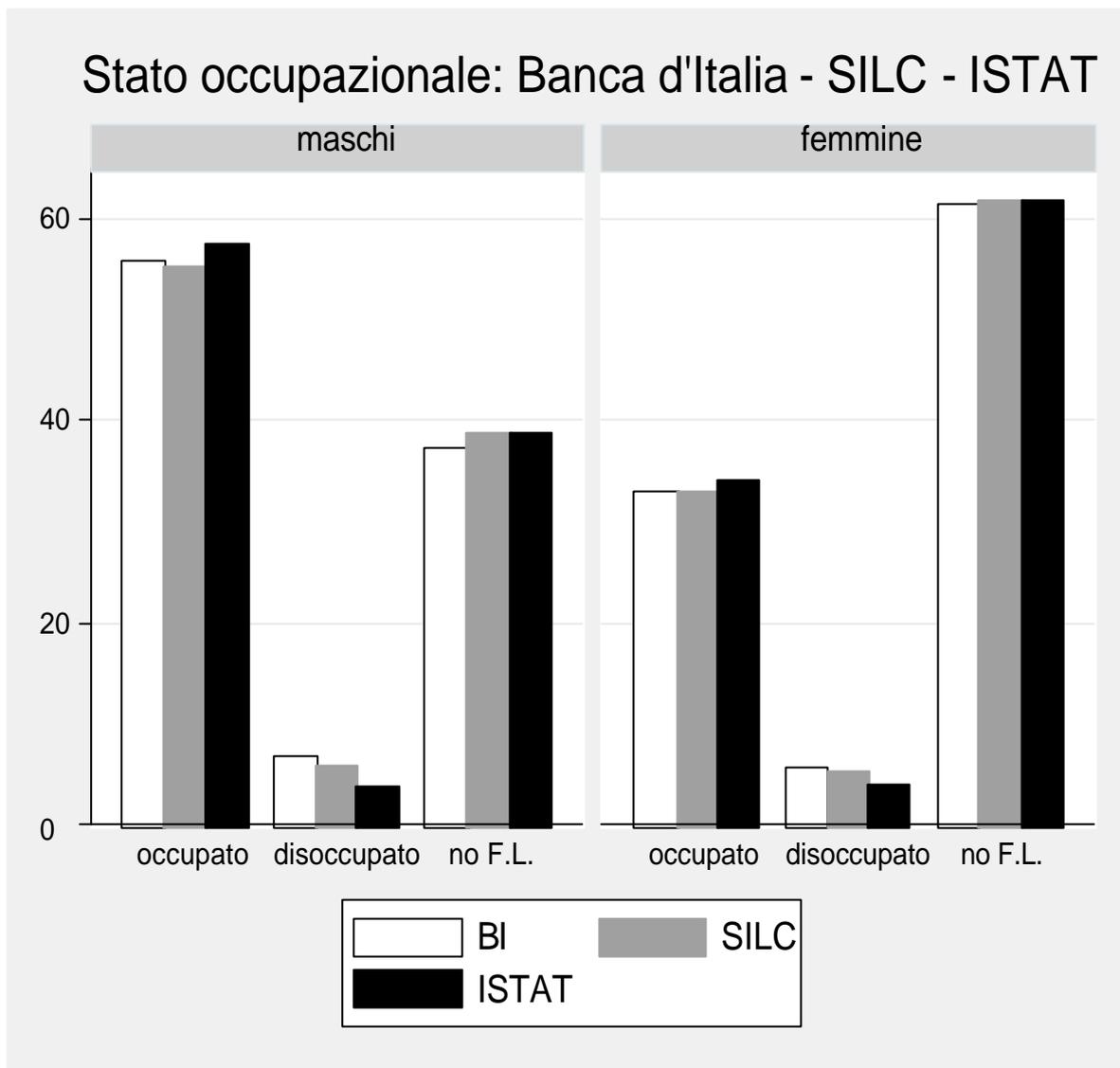
Per la questa variabile, tutti e tre gli indici considerati sono inferiori nell’indagine SILC, anche se, coerentemente con quanto visto in precedenza, le differenze non sono eclatanti. Gli indici di dissomiglianza calcolati sulle distribuzioni di frequenza e sulle cumulate sono, rispettivamente, del 2,25% e dell’1,13% per la Banca d’Italia, e dell’1,83% e dello 0,96% per IT SILC.

Attraverso l’analisi grafica è possibile appurare dove originano queste lievi distorsioni delle nostre indagini campionarie.

Rispetto agli uomini, il *bias* più evidente è relativo alla categoria dei disoccupati: se per la rilevazione sulle forze di lavoro essi sono circa il 4% della popolazione con più di 15 anni, la BdI stima una quota di circa il 7% e IT SILC circa il 6%. Inoltre, l’indagine della banca centrale sottostima di due punti la frazione di occupati (analogamente a quella dell’istituto di statistica), e di un punto quella dei non appartenenti alla forza lavoro.

Passando al grafico delle donne, emerge immediatamente come le distribuzioni di frequenza siano estremamente vicine. In particolare, la categoria delle non F.L. è perfettamente colta dalle due indagini campionarie. Le differenze si riscontrano tra le occupate (sottostima di circa un punto e mezzo) e le disoccupate (sovrastima d’identica entità).

Figura 7: distribuzione di frequenza per stato occupazionale



Naturalmente, l'altro dato molto evidente – e atteso – riguarda i valori delle frequenze nei due sessi. Così, se tra i maschi la popolazione occupata raggiunge il 57% della popolazione totale, tra le femmine essa si ferma al 34%. Poiché la quota di disoccupati è uguale in entrambi i generi (4%), tale differenza si riflette nelle quote dei non appartenenti alle forze di lavoro (39% circa per gli uomini, 62% per le donne).

Può essere utile raffinare le categorie ora descritte: ad esempio, valutare il numero di lavoratori dipendenti e autonomi nelle diverse indagini campionarie, distinguendo per sesso o classe di età. Queste informazioni sono contenute nella prossima tabella.

Tabella 7: numero di dipendenti, per sesso e classe di età

LAVORATORI DIPENDENTI: UOMINI					
CLASSE D'ETÀ	BENCHMARK¹⁰	BI	BI/RCFL	SILC	SILC/RCFL
15-24	805.920	789.520	98%	706.476	87,7%
25-34	2.543.271	2.252.120	88,6%	2.278.854	89,6%
35-44	2.971.976	2.675.280	90%	2.806.106	94,4%
45-54	2.367.631	2.692.320	113,7%	2.469.540	104,3%
55+	837.012	1.164.400	139,1%	912.792	109%
Totale	9.525.810	9.573.640	100,5%	9.173.768	96,3%
LAVORATORI DIPENDENTI: DONNE					
CLASSE D'ETÀ	BENCHMARK	BI	BI/RCFL	SILC	SILC/RCFL
15-24	543.829	403.280	74,16%	476.194	87,56%
25-34	1.987.869	1.638.680	82,43%	1.924.574	96,82%
35-44	2.270.263	2.198.160	96,82%	2.295.526	101,11%
45-54	1.668.902	2.041.960	122,35%	1.809.954	108,45%
55+	536.929	593.560	110,55%	575.184	107,12%
Totale	7.007.792	6.875.640	98,11%	7.081.432	101,05%

La tabella 7 mostra come il numero di lavoratori dipendenti¹¹, nelle due indagini in esame, sia abbastanza simile a quello dell'indagine *benchmark* (la Rilevazione Continua sulle Forze di Lavoro). Per quanto riguarda gli uomini, entrambe le indagini sottostimano il numero di dipendenti nelle prime classi di età (fino a 44 anni) per poi sovrastimarli in quelle anziane. Tuttavia, le distorsioni si compensano e l'indagine BI coglie perfettamente la numerosità totale dei dipendenti di sesso maschile, mentre SILC la sottostima di poco (circa il 4%, pari a 400.000 unità). Anche per le dipendenti donne, le indagini in esame tendono a sottostimare il numero delle giovani e a sovrastimare quello delle anziane, ma anche in questo caso le differenze si compensano e la distorsione totale è molto contenuta (circa -2% per BI e +1% per SILC).

La tabella 8 riporta gli stessi dati per i lavoratori autonomi.

¹⁰ Rilevazione Continua sulle Forze di Lavoro- Media 2005.

¹¹ Per ottenere le stime riferibili a tutta la popolazione, i dati campionari sono stati riproporzionati per un fattore, espresso come il rapporto tra la popolazione totale al 1/1/2005 (fonte Istat) e la popolazione campionaria dell'indagine di riferimento. Questa procedura è stata applicata anche per le stime degli aggregati di reddito riportate nei prossimi paragrafi.

Tabella 8: numero di autonomi, per sesso e classe di età

LAVORATORI AUTONOMI: UOMINI					
CLASSE D'ETÀ	BENCHMARK	BI	BI/RCFL	SILC	SILC/RCFL
15-24	126.730	107.920	85,2%	94.822	74,8%
25-34	899.446	497.000	55,3%	722.106	80,3%
35-44	1.315.320	877.560	66,7%	1.168.082	88,8%
45-54	983.242	897.440	91,3%	951.346	96,8%
55+	887.304	639.000	72,0%	854.440	96,3%
Totale	4.212.042	3.018.920	71,7%	3.790.796	90,0%
LAVORATORI AUTONOMI: DONNE					
CLASSE D'ETÀ	BENCHMARK	BI	BI/RCFL	SILC	SILC/RCFL
15-24	78.159	42.600	54,5%	42.722	54,7%
25-34	465.111	238.560	51,3%	334.482	71,9%
35-44	582.932	465.760	79,9%	549.134	94,2%
45-54	408.330	380.560	93,2%	407.422	99,8%
55+	282.654	184.600	65,3%	294.886	104,3%
Totale	1.817.186	1.312.080	72,2%	1.628.646	89,6%

Da questa tabella risultano evidenti le difficoltà delle due indagini campionarie di individuare correttamente il numero di lavoratori autonomi, anche se la distorsione della BI è molto più accentuata di SILC (circa il 30% contro il 10% in meno del totale, sia per gli uomini che per le donne. La sottostima della BI è comune per tutte le classi di età, ma ha dei picchi in quella dei maschi tra di 35-44 anni (-33% del totale, pari a circa 440.000 individui) e delle femmine di 25-34 anni (-49% del totale, circa 220.000 autonome).

Questa maggiore capacità dell'indagine SILC di cogliere il fenomeno del lavoro autonomo si evincerà anche nei paragrafi dedicati all'analisi delle variabili economiche.

Nella tabella 9 sono contenute le informazioni circa il numero di disoccupati, distinto per sesso e classe di età.

Tabella 9: numero di disoccupati, per sesso e classe di età.

DISOCCUPATI: UOMINI					
CLASSE D'ETÀ	BENCHMARK	BI	BI/RCFL	SILC	SILC/RCFL
15-24	254.710	536.760	210,7%	317.810	124,8%
25-34	309.051	568.000	183,8%	419.926	135,9%
35-44	172.557	252.760	146,5%	281.340	163,0%
45-54	109.114	238.560	218,6%	182.350	167,1%
55+	56.974	181.760	319,0%	105.242	184,7%
Totale	902.406	1.777.840	197,0%	1.306.668	144,8%
DISOCCUPATI: DONNE					
CLASSE D'ETÀ	BENCHMARK	BI	BI/RCFL	SILC	SILC/RCFL
15-24	234.980	454.400	193,4%	313.642	133,5%
25-34	364.629	474.280	130,1%	538.714	147,7%
35-44	248.509	181.760	73,1%	355.322	143,0%
45-54	111.862	68.160	60,9%	168.804	150,9%
55+	26.177	76.680	292,9%	41.680	159,2%
Totale	986.157	1.255.280	127,3%	1.418.162	143,8%

Nella distinzione tra occupati, disoccupati e inoccupati, i valori riferiti alla seconda modalità, così come rilevati dalle indagini campionarie, sono i più difforni rispetto al *benchmark*. Il numero di disoccupati maschi in BI è addirittura doppio rispetto a quello evidenziato dalla RCFL, mentre la sovrastima in SILC è del 145%. Anche tra le donne, il numero di disoccupate che emerge dalle indagini è più alto della “realtà”: del 27% per l’indagine BI e del 44% per SILC.

La tabella 10, l’ultima di questo paragrafo, presenta il numero di inattivi: diversamente dalle precedenti tabelle, qui l’ultima classe non è aperta ma arriva fino a 64 anni: quindi, sono esclusi gran parte dei pensionati, alla cui analisi sono dedicati gli ultimi due paragrafi.

Tabella 10: numero di non appartenenti alle forze lavoro, per sesso e classe d'età (15-64 anni).

NON APPARTENENTI ALLE FORZE DI LAVORO: UOMINI					
CLASSE D'ETÀ	BENCHMARK	BI	BI/RCFL	SILC	SILC/RCFL
15-24	1.928.562	2.047.640	106,2%	1.923.532	99,7%
25-34	544.641	335.120	61,5%	475.152	87,2%
35-44	241.792	73.840	30,5%	258.416	106,9%
45-54	344.243	266.960	77,5%	337.608	98,1%
55-64	1.887.436	2.246.440	119%	2.054.824	108,9%
Totale	4.946.675	4.970.000	100,5%	5.049.532	102,1%
NON APPARTENENTI ALLE FORZE DI LAVORO: DONNE					
CLASSE D'ETÀ	BENCHMARK	BI	BI/RCFL	SILC	SILC/RCFL
15-24	2.130.478	2.320.280	108,9%	2.076.706	97,5%
25-34	1.397.881	1.019.560	72,9%	1.203.510	86,1%
35-44	1.553.543	1.513.720	97,4%	1.527.572	98,3%
45-54	1.694.383	2.078.880	122,7%	1.647.402	97,2%
55-64	2.824.302	3.277.360	116,0%	2.997.834	106,1%
Totale	9.600.588	10.209.800	106,3%	9.453.024	98,5%

Come per i lavoratori dipendenti, e diversamente dai lavoratori autonomi e dai disoccupati, per gli inattivi le due indagini campionarie si rilevano abbastanza precise. Nonostante un'eccessiva sottostima della classe maschile di 35-44 anni (-70% del reale), l'indagine BI replica perfettamente la popolazione inattiva maschile (solo +0,5%) e discretamente quella femminile (106,3%). Anche l'indagine SILC presenta risultati soddisfacenti: solo +2% rispetto al *benchmark* per gli uomini, e -1,5% per le donne.

Per concludere l'analisi delle quattro variabili demografiche, gli indici di dissomiglianza, come abbiamo visto all'indicare una maggiore attendibilità dell'indagine BdI rispetto alla variabile stato civile e dell'indagine SILC rispetto alla condizione occupazionale, mentre non emergono differenze rilevanti rispetto alle variabili età e titolo di studio. In sintesi, nessuna delle due indagini risulta assolutamente "preferibile" in termini di somiglianza alla popolazione "vera".

La seconda parte del lavoro analizza le variabili economiche, soprattutto il reddito (e le sue componenti: da lavoro, da trasferimenti, etc.) per alcune tipologie di persone, come lavoratori dipendenti e autonomi.

5) ANALISI PER REDDITO FAMILIARE

In questo paragrafo si confrontano, *solo* per le due indagini campionarie, i valori di alcune tipologie di reddito familiare. Prima di procedere al confronto, sono state eseguite alcune operazioni per rendere le due banche dati omogenee¹².

Tabella 11: reddito familiare netto per tipologia

TIPOLOGIA DI REDDITO FAMILIARE (NETTO)	BI	SILC
Da lavoro dipendente	11.884	13.075
<i>Solo famiglie con reddito positivo</i>	21.422	23.176
Da lavoro autonomo	3.919	5.447
<i>Solo famiglie con reddito positivo</i>	23.558	19.591
Da pensione	6.741	7.669
<i>Solo famiglie con reddito positivo</i>	14.044	15.329
Da attività finanziarie	669	459
<i>Solo famiglie con reddito positivo</i>	1.282	972
Da terreni e fabbricati	329	443
<i>Solo famiglie con reddito positivo</i>	8.057	7.192
Totale	23.706	27.815

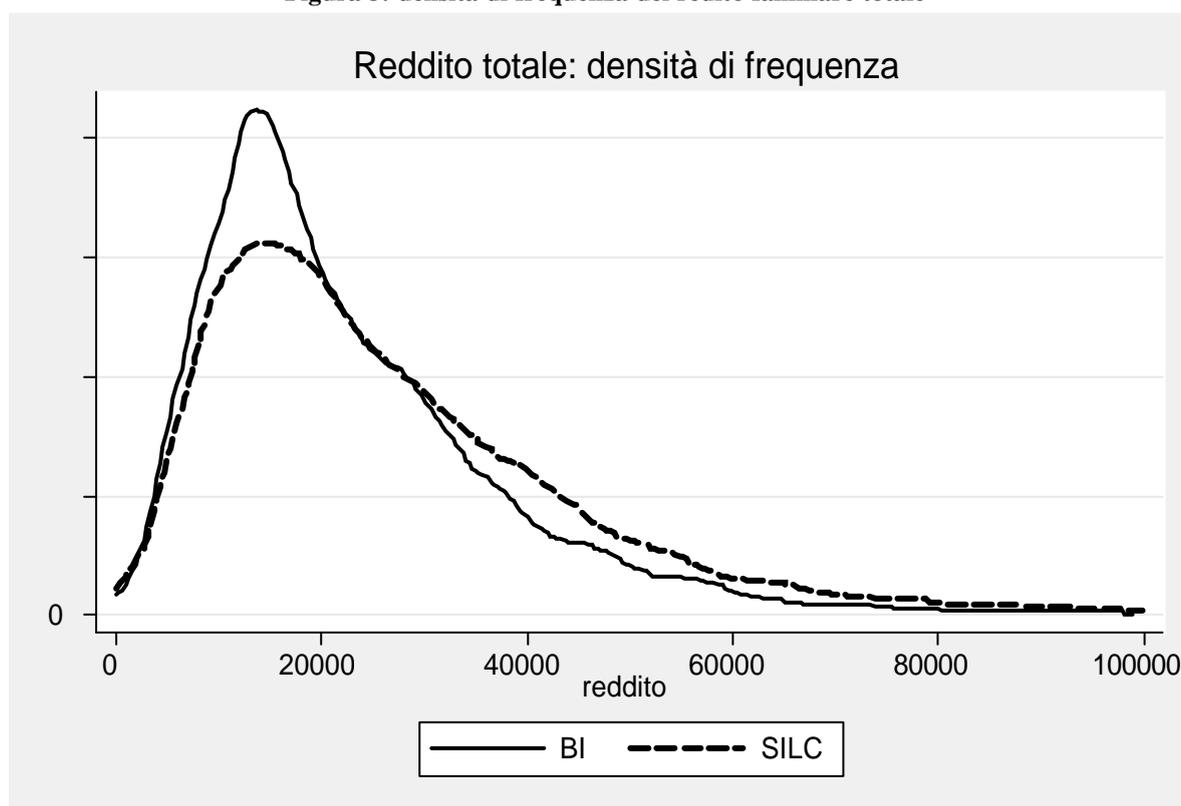
I valori medi dei redditi netti, distinti per tipologia, appaiono molto dissimili. Innanzitutto, il reddito disponibile netto, che per l'indagine SILC è di circa 4.000 euro superiore rispetto all'indagine Banca d'Italia. Esaminando i redditi per singola categoria, coerentemente con il risultato totale, quasi tutte le voci registrano una situazione più rosea nell'indagine SILC: infatti, rispetto all'indagine BI *ogni* famiglia gode di circa 1500 euro in più per i redditi da lavoro autonomo, e circa 1000 per i redditi da pensione. In controtendenza

¹² Vanno segnalate alcune definizioni particolari adottate nel confronto fra le due indagini per ragioni di omogeneità: (1) sono esclusi eventuali percettori di età inferiore ai 15 anni; (2) i redditi da lavoro dipendente sono al netto delle integrazioni non monetarie; (3) gli utili e gli altri redditi da partecipazione sono stati inclusi nei redditi da attività finanziarie ed esclusi da quelli di lavoro autonomo; (4) i redditi da pensione non includono gli arretrati; (5) i redditi da capitale reale escludono gli affitti imputati delle abitazioni occupate dai proprietari. Il reddito totale familiare riflette le convenzioni precedenti.
Cfr. Di Marco, M. (2005). "La nuova indagine sui redditi e le condizioni di vita: principali obiettivi e risultati", ISTAT, Roma, p. 10.

i redditi da attività finanziarie, ma questo può dipendere dal modo in cui sono stati rilevati¹³. Se restringiamo il campo alle famiglie con redditi positivi, categoria per categoria, le differenze più visibili emergono dalla categoria dei lavoratori autonomi, con ben 4.000 euro di divario a favore dell'indagine BI. Questo si spiega con la diversa numerosità di famiglie con reddito da lavoro autonomo positivo, nelle due indagini: in SILC esse sono 6525, pari a circa il 30% del campione; in BI sono solamente 1273, pari ad appena il 16% delle famiglie totale. Di fatto, il monte dei redditi autonomi (pur superiore in SILC), nell'indagine BdI è concentrato su una quota di popolazione molto più ristretta. Ad ulteriore riprova di quanto detto, l'indice di Gini, calcolato sui redditi in questione, è di 0,86 nel campione dell'istituto di statistica, e di 0,92 in quello della banca centrale.

La distribuzione di frequenza del reddito totale mostra la maggior densità della distribuzione BI per i redditi bassi (la coda sinistra è più "grossa" della corrispondente coda SILC), mentre nella parte alta le frequenze SILC sono più elevate, a segnalare una maggiore concentrazione di individui con un reddito più consistente.

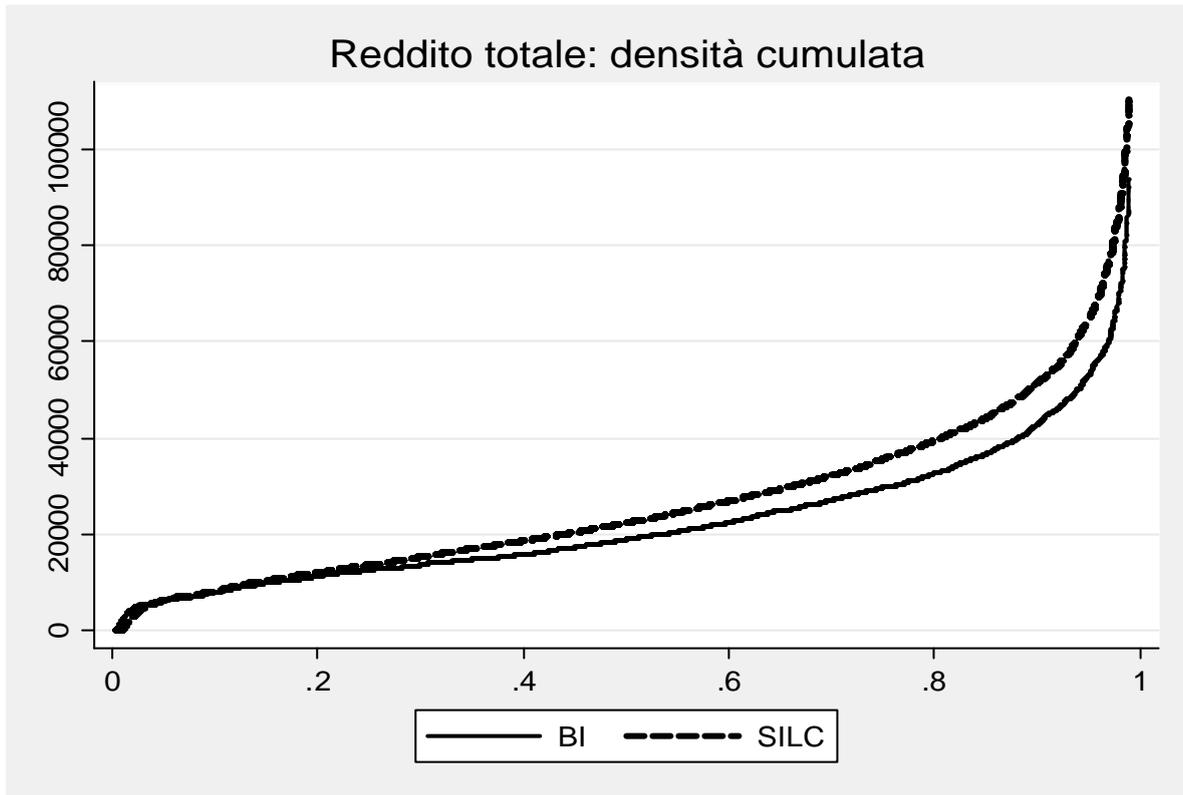
Figura 8: densità di frequenza del reddito familiare totale



Il prossimo grafico mostra la densità di frequenza cumulata delle due distribuzioni: per ragioni espositive, gli assi sono invertiti, per cui si dovrebbe più propriamente parlare di "sfilata di Pen".

¹³ Ibidem.

Figura 9: "sfilata" Pen del reddito familiare totale



La disuguaglianza delle distribuzioni emerge già dal terzo decile: oltre questa soglia, per ogni decile di popolazione, il reddito disponibile è maggiore nell'indagine SILC.

La maggior concentrazione del reddito nelle fasce alte nell'indagine SILC, rispetto a BI, ha anche l'effetto di produrre un indice di Gini più alto (0,3774 contro 0,3635¹⁴).

Può essere utile passare ad analizzare la concentrazione del reddito familiare per le più rilevanti tipologie di reddito, come quello da lavoro dipendente, da lavoro autonomo e da trasferimenti.

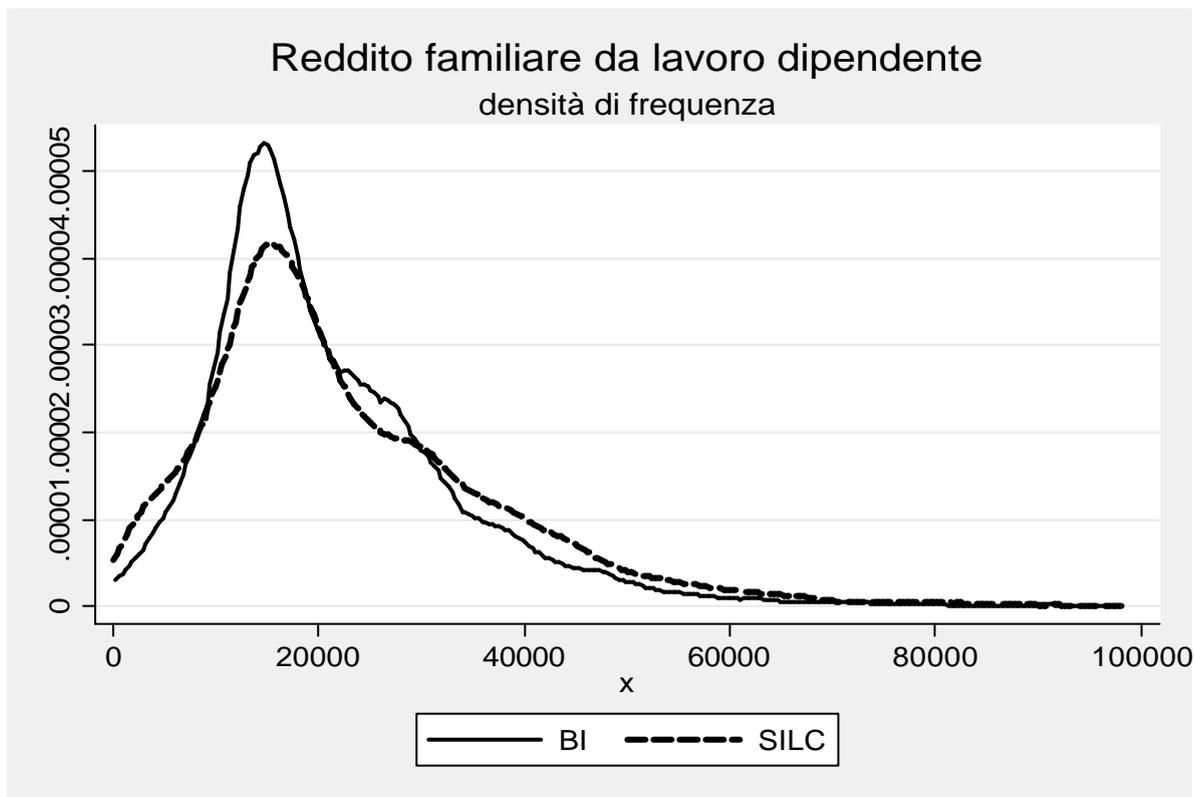
Iniziamo dai redditi da lavoro dipendente.

L'indagine BI mostra una maggiore concentrazione di famiglie intorno alla fascia medio-bassa, dai 10.000 ai 30.000 euro circa, mentre la parte iniziale della distribuzione, e quella finale, vede code più "grosse" nell'indagine SILC. Con riferimento al lavoro dipendente, l'indagine dell'istituto di statistica attribuisce delle quote maggiori ai redditi

¹⁴ Questi valori dell'indice di Gini sono calcolati sui redditi familiari escludendo i fitti imputati. Se questi ultimi sono inclusi, otteniamo dei valori, rispettivamente, di 0,3573 e 0,3512. Se, infine, i redditi sono divisi per un quoziente familiare che tiene conto della dimensione del nucleo familiare, l'indice di Gini diventa di 0,33 per BI e 0,3261 per SILC. In questo caso, la scala utilizzata è quella OCSE che assegna peso 1 al capofamiglia, 0,5 ad ogni altro componente con più di 14 anni, e 0,3 ai membri con meno di 14 anni.

bassissimi e medio-alti, mentre sottorappresenta (rispetto all'indagine della banca centrale) il cosiddetto ceto medio.

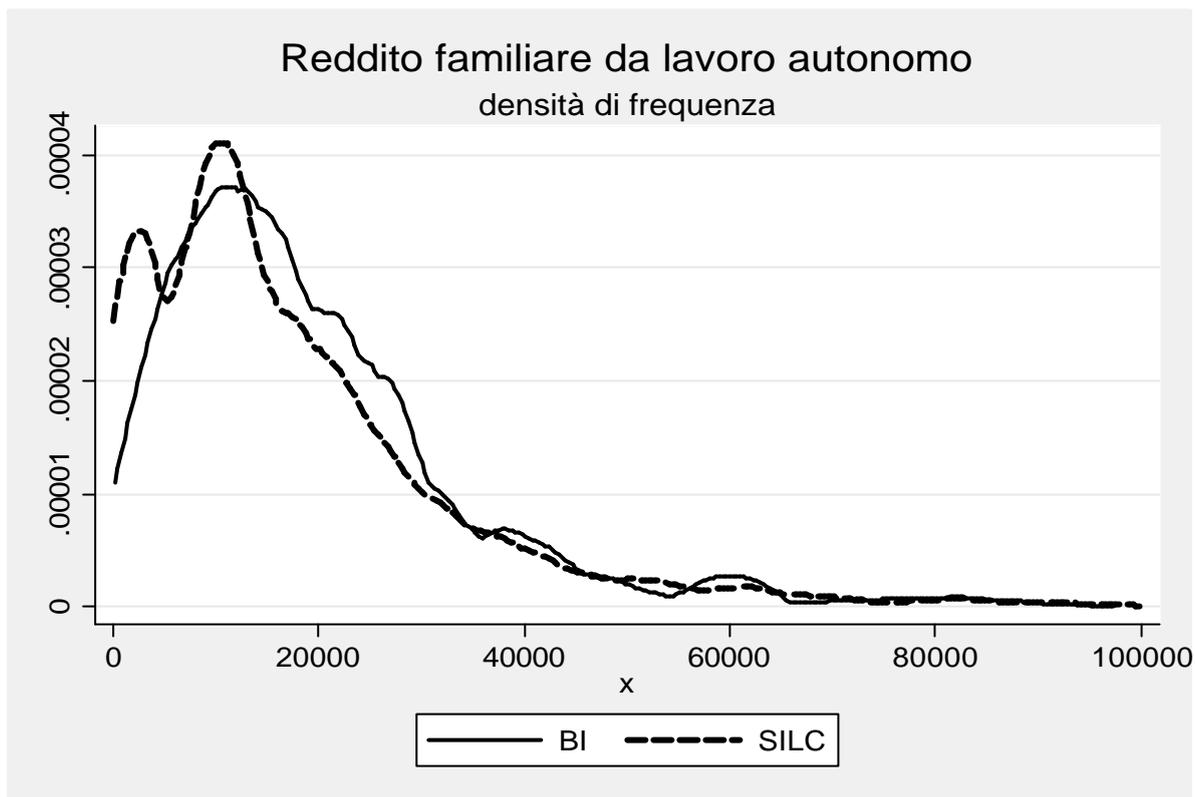
Figura 10: densità di frequenza del reddito familiare da lavoro dipendente



Passando ai redditi familiari da lavoro autonomo, si nota immediatamente la diversa forma della distribuzione di frequenza, rispetto al grafico precedente. In particolare, invece del consueto andamento log-normale, le distribuzioni mostrano una coda sinistra molto alta, e poi un andamento decrescente all'aumentare del reddito (soprattutto per SILC) . Naturalmente, in questo caso gioca il fatto che molti lavoratori autonomi chiudano in pareggio, o in un magro attivo, la loro attività (per semplicità, sono esclusi dal grafico coloro che chiudono in passivo).

Diversamente dal grafico precedente, per i redditi autonomi medio-alti è l'indagine BI a rilevare code più alte, mentre la parte sinistra della distribuzione vede il prevalere delle frequenze SILC. Come già accennato, ciò si deve alla diversa numerosità di famiglie con questa tipologia di reddito positivo, con la conseguenza che in SILC il monte dei redditi da lavoro autonomo sia “diluito” su una platea molto più ampia di individui.

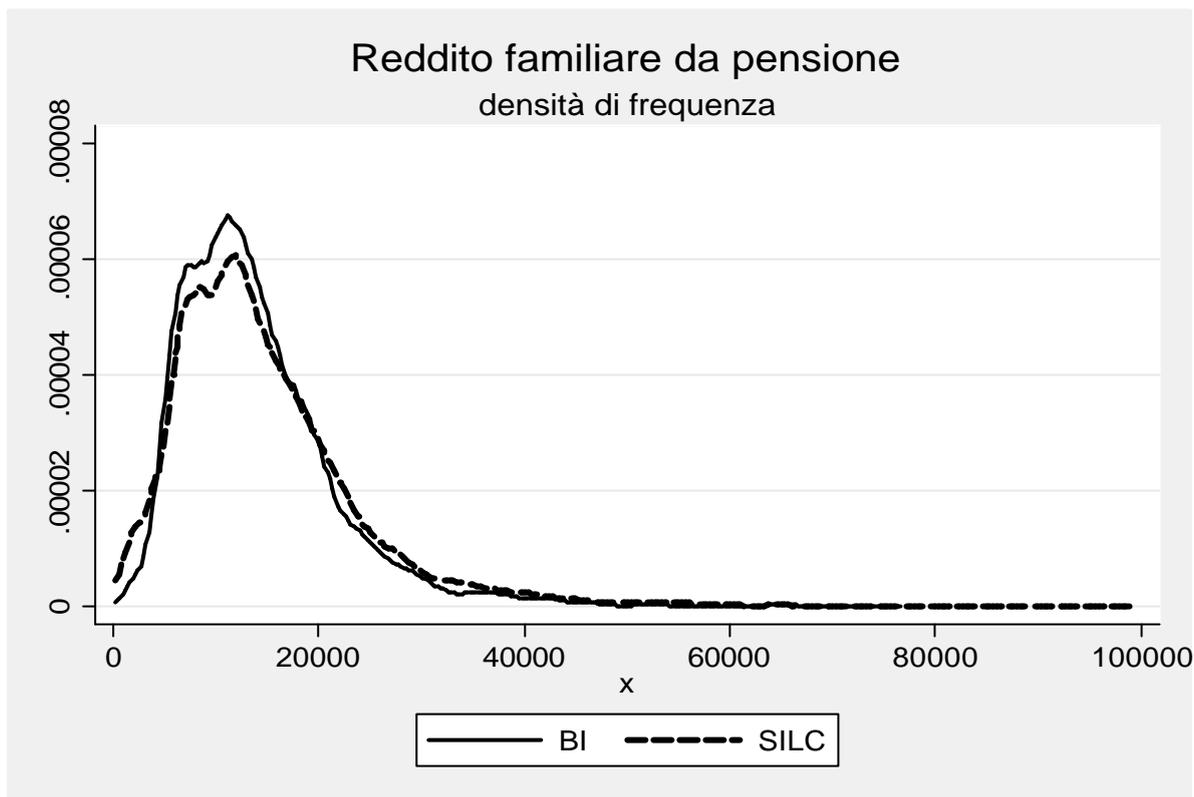
Figura 11: densità di frequenza del reddito familiare da lavoro autonomo



L'ultima categoria di interesse è quella del reddito da pensione.

L'andamento è simile a quello dei redditi dipendenti: in una primissima fase, per i redditi molto bassi, prevale l'indagine SILC; in una seconda fase, per i redditi fino a 20.000 euro l'anno, le densità di frequenza in BI sono leggermente più elevate; la relazione d'ordine si inverte per i redditi alti, che segnano, anche se non di moltissimo, la più alta concentrazione di redditi alti. Da notare che il valore massimo per la Banca d'Italia è di circa 75.000 euro l'anno, laddove il massimo in SILC supera i 300.000 (anche se questo dal grafico non si vede).

Figura 12: densità di frequenza del reddito familiare da pensione



6) CONFRONTO DEI REDDITI CON GLI AGGREGATI DELLA CONTABILITÀ NAZIONALE

Confrontare i redditi rilevati dalla indagini campionarie con i valori della contabilità nazionale pone almeno tre problemi¹⁵: riportare il campione all'universo di riferimento, rendere comparabili i redditi netti campionari con quelli lordi della CN, procedere ad una classificazione omogenea dei redditi nelle rilevazioni che si vuole confrontare. Per il primo problema, qui si è riponderato il peso familiare della Banca d'Italia per un coefficiente, ottenuto come rapporto tra la popolazione italiana al 1/1/2005 e la popolazione campionaria BI (per SILC non c'è stato bisogno, poiché questa operazione è già sussunta nel peso familiare). Il secondo problema può essere risolto in due modi: lordizzare i dati campionari netti (tuttavia questo è costoso dal punto di vista computazionale) o applicare un'aliquota media ai dati lordi della CN. Il terzo problema richiede una certa attenzione nel rendere omogenee le categorie di reddito rilevate dalla CN con quelle delle indagini campionarie.

¹⁵ Coromaldi M. e Guerrerà D. "Modello di microsimulazione ECONLAV: la costruzione del data-set di input", <http://www.dt.tesoro.it/Aree-Docum/Analisi-Pr/Modello-di/Modello-di/Working-Pa/Trattamento-preliminare-dei-dati0202.pdf>.

Fatte queste premesse, è possibile presentare una tabella dei redditi aggregati *netti* per tipologia di reddito.

I valori per la contabilità nazionale sono tratti dal lavoro di Coromaldi e Guerrera e si riferiscono al 2002. Qui sono stati riportati al 2004 applicando i coefficienti di rivalutazione monetaria e i tassi di crescita reale diffusi dall'ISTAT.

Tabella 12: confronto tra aggregati di reddito con i dati della Contabilità Nazionale (CN)

AGGREGATI PER TIPOLOGIA DI REDDITO (miliardi di euro)					
TIPOLOGIA DI REDDITO	BENCHMARK	BI	BI/CN	SILC	SILC/CN
Lavoro Dipendente	302	271	89,8%	300	99,4%
Lavoro Autonomo	269	89,2	33,1%	125	46,4%
Capitale (interessi netti e altri utili distribuiti)	45	15,2	33,8%	10,5	23,4%
Trasferimenti	221	157	71,0%	199	89,9%
Totale	888	671	75,5%	761	85,7%

Nel complesso, l'indagine SILC coglie l'85% dei redditi disponibili censiti dalla Contabilità Nazionale, a fronte del 75% della Banca d'Italia. In termini assoluti, l'indagine dell'istituto di statistica riesce a "recuperare" 90 miliardi di euro in più rispetto quella della banca d'Italia. Questi derivano dai risultati migliori ottenuti nelle categorie di reddito più importanti: dipendente, autonomo e trasferimenti.

Distinguendo per tipologia, i redditi meglio rappresentati sono quelli da lavoro dipendente: la distorsione è di 10 punti in BI e solo dello 0,4% in SILC. I redditi da lavoro autonomo, come previsto, sono largamente sottorappresentati: le percentuali di copertura sono del 33% in BI e del 46% in SILC. Dei 269 miliardi che, secondo la CN, remunerano il lavoro indipendente, ne ritroviamo solamente 89 nella prima indagine, e 125 nella seconda. Queste distorsioni sono in linea con quelle evidenziate lungo un arco di tempo più esteso da Brandolini¹⁶. L'ultima categoria di interesse (i redditi da capitale contano per solo 45 miliardi nella contabilità nazionale) è quella dei redditi tra trasferimento, comprese le pensioni. Anche in questo caso la rappresentatività di SILC è maggiore, con una quota del 90% del totale (-22 miliardi) contro il 71% di BI (-64 miliardi), e deriva soprattutto dalla diversa numerosità delle pensioni nelle due indagini (v. dopo).

¹⁶ Brandolini A. (1999). "The distribution of personal income in post-war Italy: source description, data quality, and the time pattern of income inequality", *Banca d'Italia, Temi di Discussione*, N. 350.

7) TIPOLOGIE DI PENSIONI: PRESTAZIONI E IMPORTI

Un aspetto interessante da valutare è quello della numerosità e della consistenza delle varie prestazioni pensionistiche. In questo caso il *benchmark* è costituito dalla monografia sul monitoraggio delle politiche sociali, curata dal Ministero del Lavoro nel 2006¹⁷. Da notare che in questo paragrafo l'accento non è sui pensionati, ma sulle pensioni: infatti il loro numero differisce, poiché può esserci più di una pensione per pensionato.

Tabella 13: numero di pensioni per tipologia. Confronto con i dati ISTAT-INPS

NUMERO DI PENSIONI					
TIPO DI PENSIONE	BENCHMARK (CN)	BI	BI/CN (%)	SILC	SILC/CN (%)
1. Invalidità civile	889.305	575.254	64,7%	1.030.950	115,9%
2. Accompagnamento	1.929.290	1.054.696 ¹⁸	54,7%	1.054.696	54,7%
3. Invalidità INPS	2.215.916	664.611	30,0%	1.886.800	85,1%
4. Infortunio sul lavoro INAIL	1.078.420	212.498	19,7%	923.239	85,6%
5. Sociale	755.487	567.694	75,1%	730.032	96,6%
6. Guerra	421.686	79.485	18,8%	137.959	32,7%
7. Reversibilità	4.903.213	2.779.490	56,7%	4.081.808	83,2%
8. Lavoro (vecchiaia e anzianità)	11.254.661	10.896.848	96,8%	10.186.778	90,5%
Totale	23.447.978	16.858.785	71,9%	20.032.262	85,4%

Il numero totale di trattamenti pensionistici nel 2004, così come rilevato dai dati ISTAT- INPS, è di circa 23 milioni e mezzo. Il totale di pensioni riportato dalle indagini è di circa 17 milioni per la Banca d'Italia (pari al 72% del valore "vero") e di poco più di 20 milioni per l'indagine SILC (85,4%), che si rivela quella più precisa nel rilevare il numero di pensioni.

Distinguendo per tipologia di pensione, la sottostima più evidente imputabile a BI riguarda le pensioni di invalidità civile (65% del totale) e soprattutto quelle di invalidità erogate dall'Inps (30%). Nelle prime tre categorie, l'indagine BI "perde" oltre 2,7 milioni di trattamenti, risultato che influenza negativamente la rappresentatività dell'universo dei

¹⁷ Ministero del Lavoro e delle Politiche Sociali, 2006. *Monitoraggio delle politiche sociali –II parte- I trasferimenti monetari per invalidità, pensioni sociali, integrazioni al minimo, pensioni di guerra e ai superstiti*, Roma.

¹⁸ L'indagine Banca d'Italia non riporta le indennità di accompagnamento: per questa ragione ho imputato gli stessi dati rilevati dall'indagine SILC.

pensionati. Altre categorie estremamente sottorappresentate sono quelle delle pensioni di guerra e per infortunio sul lavoro: entrambe si fermano al 20%. Tuttavia l'indagine della banca centrale rappresenta molto bene la categoria più importante, e cioè quelle delle pensioni da lavoro: in questo caso, la discrepanza con i dati reali è solo del 3%, pari a 350.000 unità.

L'indagine SILC rappresenta abbastanza bene le pensioni di invalidità ed accompagnamento: invece degli oltre 2,7 milioni di differenza, la distanza dai dati ministeriali per queste tipologie si ferma ad un milione di pensioni, dovuti perlopiù ad una sottostima delle indennità d'accompagnamento (probabilmente non agevoli da identificare da parte del beneficiario). Risultati lusinghieri si hanno anche per le pensioni sociali (96% contro il 75% di BI) e di reversibilità (83% contro 56%). Purtroppo, però, essa sottorappresenta la categoria fondamentale delle pensioni da lavoro: lo scostamento percentuale è del 10%, pari a circa un milione di pensioni. Tuttavia, ciò non toglie che, almeno in merito alla numero dei trattamenti, l'indagine SILC sia decisamente più precisa.

Passiamo ora all'analisi del reddito annuale da pensione, distinto per tipologia.

Tabella 14: importo medio della pensione per categoria. Confronto con i dati ISTAT-INPS

IMPORTO MEDIO ANNUALE DELLA PENSIONE					
TIPO DI PENSIONE	BENCHMARK (CN)	BI	BI/CN (%)	SILC	SILC/CN (%)
1. Invalidità civile	2.681	5.283	197,1%	4.675	174,4%
2. Accompagnamento	5.012	5.153	102,8%	5.153	102,8%
3. Invalidità	6.487	5.921	91,3%	6.707	103,4%
4. Infortunio sul lavoro	3.999	6.108	152,7%	4.465	111,7%
5. Sociale	4.318	5.202	120,5%	4.462	103,3%
6. Guerra	3.694	4.503	121,9%	6.688	181,1%
7. Reversibilità	6.674	7.218	108,2%	6589	98,7%
8. Lavoro (vecchiaia e anzianità)	12.531	11.110	88,7%	11.316	90,3%
Totale	8.927	9.389	105,2%	8.655	97,0%

Va sottolineato che l'importo diffuso dal casellario statistico, che qui costituisce il *benchmark*, è al lordo delle imposte sui redditi. Tuttavia, esse non dovrebbero essere elevatissime, dato il non ingente importo medio della pensione. Gli importi annui delle pensioni sono, nel complesso, abbastanza confacenti con quelli rilevati dall'indagine ministeriale: -3% per SILC e +5% per BI. I risultati peggiori registrati dalle indagini

campionarie si hanno per l'invalidità civile (anche se qui, come detto in precedenza, può risultare difficile per l'intervistato identificare la tipologia giusta e distinguerla da altri interventi *means-tested*) con una sovrastima che va dal 75% di SILC al 97% di BI. Spicca anche il 181% di SILC per le pensioni di guerra (tuttavia non preoccupante, data la relativa esiguità di questa categoria) e il 150% della Banca d'Italia per gli infortuni sul lavoro. Nuovamente, per la categoria più "sensibile", quella delle pensioni di vecchiaia e anzianità, le due indagini sembrano comportarsi abbastanza bene: la sottostima, infatti, è di circa il 10% per entrambe.

Per concludere questo paragrafo, si può affermare che SILC sia più adatta di BI a rappresentare l'universo pensionistico, soprattutto grazie alla maggiore capacità di "contare" i trattamenti; le due indagini, invece, sono abbastanza omogenee se rapportate all'importo medio della pensione distinta per tipologia.

8) CONFRONTO DELLE PENSIONI CON GLI AGGREGATI DELLA CONTABILITÀ NAZIONALE (CN)

Eseguiamo qui un confronto analogo a quello fatto sulle fonti di reddito nel sesto paragrafo.

**Tabella 15: confronto tra aggregati di reddito da pensione per categoria
AGGREGATI PER TIPOLOGIA DI PENSIONE (milioni di euro)**

TIPO DI PENSIONE	<i>BENCHMARK</i> (CN)	BI	<i>BI/CN</i> (%)	SILC	<i>SILC/CN</i> (%)
1. Invalidità civile	2.384,2	3.039,1	127,5%	4.819,7	202,1%
2. Accompagnamento	9.669,6	5.580,2	57,7%	5.434,8	56,2%
3. Invalidità	14.374,6	3.935,2	27,4%	12.654,8	88,0%
4. Infortunio sul lavoro	4.312,6	1.297,9	30,1%	4.122,3	95,6%
5. Sociale	3.262,2	2.953,1	90,5%	3.257,4	99,9%
6. Guerra	1.557,7	357,9	23,0%	922,7	59,2%
7. Reversibilità	32.724,0	20.062,4	61,3%	26.895,0	82,2%
8. Lavoro (vecchiaia e anzianità)	141.032,2	121.064,0	85,8%	115.273,6	81,7%
Totale	209.317,2	158.289,8	75,6%	173.380,3	82,8%

In generale, il monte pensioni rilevato da SILC è l'83% di quello desunto dai dati INPS, mentre la BI si ferma al 75%: nel primo caso la sottostima è di 35 miliardi di euro, nel secondo di 60¹⁹.

Trascurando le distorsioni relative alle prime due tipologie, per i problemi di rilevazione già sottolineati, notiamo subito delle grandi differenze tra le due indagini per le pensioni di invalidità e quelle per infortuni sul lavoro. Mentre l'indagine SILC si dimostra abbastanza precisa (rispettivamente, 88% e 95% dell'aggregato di riferimento), la sottostima di BI appare molto grave, non arrivando a cogliere più del 30% del fenomeno reale. In termini assoluti, la "perdita" è quantificabile in 2 miliardi di euro per SILC, ed oltre 13,5 per la Banca d'Italia (queste distorsioni, come le successive, derivano quasi esclusivamente dalla sottostima del numero di trattamenti, dato che entrambe le indagini si equivalgono nello stimare l'importo medio. Un altro risultato negativo, per lo l'indagine di via Nazionale, riguarda le pensioni di reversibilità, il cui aggregato è sottostimato del 40%, a fronte di un 18% dell'altra indagine. L'ultima categoria di rilievo è quella delle pensioni da lavoro: qui, in controtendenza, la rilevazione più rappresentativa è BI (86% a fronte di un 82% di SILC) con una differenza assoluta di 20 miliardi di euro invece di 26.

Per concludere la parte sulle pensioni, possiamo dire che l'indagine SILC offre maggiori garanzie di rappresentatività della popolazione vera, rispetto all'indagine BI. Abbiamo valutato le due indagini sulla base del numero dei trattamenti pensionistici, del loro importo medio, e dell'aggregato del monte pensioni. Se l'importo medio è simile nelle sue indagini (e molto vicino a quello "vero"), le differenze maggiori emergono dall'analisi del numero di pensioni. In quest'ultimo caso, l'indagine BI sottostima decisamente l'entità di questa variabile. Di conseguenza, poiché l'aggregato del monte pensioni non è altro che il prodotto tra importo medio e numero di pensioni, anche l'ultima variabile risulta maggiormente distorta in BI.

¹⁹ Le sottostime dovrebbero essere inferiori in entrambi i casi, visto che le indagini campionarie rilevano i dati netti, mentre quelle amministrative rilevano i redditi al lordo delle imposte.

BIBLIOGRAFIA

1. BRANDOLINI A. (1999). “The distribution of personal income in post-war Italy: source description, data quality, and the time pattern of income inequality”, *Banca d'Italia, Temi di Discussione*, N. 350.
2. BRANDOLINI A., CANNARI L., D'ALESSIO G. e FAIELLA I. (2004). “Household wealth distribution in Italy in the 1990s”, *Banca d'Italia, Temi di Discussione*, N.530.
3. COROMALDI M. e GUERRERA D. “Modello di microsimulazione ECONLAV: la costruzione del data-set di input”, <http://www.dt.tesoro.it/Aree-Docum/Analisi-Pr/Modello-di/Modello-di/Working-Pa/Trattamento-preliminare-dei-dati0202.pdf>.
4. DI MARCO, M. (2005). “La nuova indagine sui redditi e le condizioni di vita: principali obiettivi e risultati”, ISTAT, Roma.
5. ISTAT (2006). “Rapporto annuale: la situazione del Paese nel 2005”, Roma.
6. ISTAT (2006). “Rilevazione sulle forze di lavoro – Media 2005”, http://www.istat.it/salastampa/comunicati/non_calendario/20060418_00/testointegral.e.pdf.
7. ISTAT (2008). “Reddito e condizioni di vita in Italia (2005-2006)”, http://www.istat.it/salastampa/comunicati/non_calendario/20080117_01/testointegral.e20080117.pdf.
8. MINISTERO DEL LAVORO E DELLE POLITICHE SOCIALI (2006). *Monitoraggio delle politiche sociali –II parte- I trasferimenti monetari per invalidità, pensioni sociali, integrazioni al minimo, pensioni di guerra e ai superstiti*, Roma.
9. VAGLIASINDI P., MOSCHETTI F., ROMANELLI M. e BIANCHI C. (2001). “La microsimulazione delle riforme previdenziali e fiscali italiane: una nota sulla validazione delle dinamiche nazionali e regionali”, *ECOFIN Discussion Paper*, Università di Parma.
10. VAGLIASINDI P. (2004). *Effetti redistributivi dell'intervento pubblico: esperimenti di microsimulazione per l'Italia*, G. Giappichelli, Torino.