



SISTEMA STATISTICO NAZIONALE
ISTITUTO NAZIONALE DI STATISTICA

A stylized map of Italy composed of various colored squares (blue, red, orange, yellow, green, cyan, grey, white) arranged to form the geographical outline of the country. The colors likely represent different regions or data categories.

Integrazione di dati campionari Eu-Silc con dati di fonte amministrativa





SISTEMA STATISTICO NAZIONALE
ISTITUTO NAZIONALE DI STATISTICA

**Integrazione
di dati campionari Eu-Silc
con dati di fonte amministrativa**

A cura di: Paolo Consolini

Per informazioni sul contenuto della pubblicazione
rivolgersi al Cont@ct Centre dell'Istat all'indirizzo:
<https://contact.istat.it/>

Eventuali rettifiche ai dati pubblicati saranno diffuse
all'indirizzo www.istat.it nella pagina di presentazione del volume

Integrazione di dati campionari Eu-Silc con dati di fonte amministrativa

Metodi e Norme n. 38 – 2009

ISBN 978-88-458-1612-3

© 2009

Istituto nazionale di statistica
Via Cesare Balbo, 16 – Roma

Realizzazione: Istat, Servizio produzione editoriale

Stampato nel mese di marzo 2009
presso il Centro Stampa dell'Istat
Via Tuscolana, 1788 - Roma

Si autorizza la riproduzione a fini non
commerciali e con citazione della fonte

Indice

	Pagina
Prefazione	7
Capitolo 1 - L'attribuzione dei codici fiscali individuali a partire dal patrimonio informativo di Sigif	11
1.1 - Introduzione	11
1.2 - Il sistema Sigif	11
1.3 - La generazione del codice fiscale (presunto): aspetti generali	12
1.4 - La generazione del codice fiscale: implementazione della procedura di calcolo	13
Capitolo 2 - Le procedure per l'acquisizione ed il data-cleaning degli archivi amministrativi	23
2.1 - Introduzione.....	23
2.2 - Il modello Cud/770	24
2.2.1 - <i>L'acquisizione dei modelli Cud/770</i>	24
2.2.2 - <i>Le operazioni di controllo, correzione e consolidamento dei dati del Cud/770 ..</i>	27
2.3 - Il modello 730	29
2.3.1 - <i>L'acquisizione dei modelli 730</i>	29
2.3.2 - <i>Le operazioni di controllo, correzione e consolidamento dei dati del 730</i>	30
2.4 - Il modello Unico persone fisiche (Upf).....	32
2.4.1 - <i>L'acquisizione dei modelli Upf</i>	32
2.4.2 - <i>Le operazioni di controllo, correzione e consolidamento dei dati dell'Upf</i> ...	33
2.5 - Il Casellario centrale dei pensionati (CP): acquisizione dei microdati	34
Capitolo 3 - La costruzione delle variabili di reddito Eu-Silc a partire dalle variabili fiscali: armonizzazione delle definizioni	37
3.1 - Introduzione.....	37
3.2 - La costruzione del reddito da lavoro dipendente	40
3.2.1 - <i>Il reddito netto annuo da lavoro dipendente di fonte Cud/770</i>	40
3.2.2 - <i>Le mensilità retribuite nell'anno dei lavoratori dipendenti da fonte Cud/770 ..</i>	43
3.2.3 - <i>Il reddito netto annuo da lavoro dipendente di fonte 730 e Unico persone fisiche</i>	47
3.2.4 - <i>Le mensilità retribuite dei lavoratori dipendenti di fonte 730 e Upf</i>	51
3.3 - La costruzione del reddito da pensione	53
3.3.1 - <i>Premessa</i>	53
3.3.2 - <i>Il reddito lordo annuo da pensione di fonte CP</i>	54
3.3.3 - <i>Il reddito netto annuo da pensione di fonte Cud/770</i>	54
3.3.4 - <i>Il reddito netto annuo da pensione di fonte di fonte 730 e Upf</i>	55
3.4 - La costruzione dei trasferimenti non pensionistici	56
3.4.1 - <i>Premessa</i>	56
3.4.2 - <i>Gli assegni al nucleo familiare di fonte Cud/770</i>	57
3.4.3 - <i>Le liquidazioni di fine rapporto di fonte Cud/770</i>	62
3.4.4 - <i>I trattamenti non pensionistici ad erogazione diretta Inps di fonte Cud/770</i>	65
3.5 - La costruzione del reddito da lavoratore autonomo	66

	Pagina
3.5.1 - <i>Premessa</i>	66
3.5.2 - <i>Il reddito dei collaboratori coordinati e continuativi</i>	68
3.5.3 - <i>Il reddito dei soci di cooperative e dei ministri del culto</i>	69
3.5.4 - <i>Il reddito derivante dall'attività intramuraria/intramoenia di medici e operatori sanitari</i>	69
3.5.5 - <i>Le provvigioni inerenti a rapporti di commissione, agenzia, mediazione, rappresentanza di commercio e procacciamento d'affari</i>	69
3.5.6 - <i>Il reddito derivante dall'utilizzazione economica di opere d'ingegno</i>	70
3.5.7 - <i>Il reddito degli associati in partecipazione</i>	71
3.5.8 - <i>Il reddito dei lavoratori autonomi occasionali e dei venditori a domicilio</i>	72
3.5.9 - <i>Il reddito dei liberi professionisti e artisti</i>	73
3.5.10 - <i>Il reddito dei titolari d'impresa</i>	74
3.5.11 - <i>I proventi dei compartecipanti al reddito di imprese e società di persone</i>	77
Capitolo 4 - Strategie per l'integrazione dei dati campionari e di fonte amministrativa: soluzioni al problema dell'attribuzione dei redditi	79
4.1 - <i>Introduzione</i>	79
4.2 - <i>Il disegno per l'integrazione dei microdati di fonte amministrativa</i>	81
4.3 - <i>L'identificazione dei redditi da pensione da fonte amministrativa</i>	83
4.3.1 - <i>Il calcolo del reddito netto delle pensioni tassabili di sola fonte CP</i>	84
4.3.2 - <i>Riconciliazione del reddito da pensione tra fonte CP e fonti fiscali</i>	85
4.3.3 - <i>Attribuzione della tipologia di pensione ai redditi di sola fonte fiscale</i>	90
4.4 - <i>L'identificazione dei redditi da lavoro dipendente da fonte amministrativa</i>	90
4.5 - <i>Soluzioni alle incoerenze tra dati campionari e fiscali</i>	92
4.5.1 - <i>I redditi da lavoro dipendente</i>	93
4.5.2 - <i>I redditi da lavoro autonomo e parasubordinato</i>	96
4.5.3 - <i>I redditi da pensione</i>	98
4.6 - <i>Riconciliazione del reddito campionario e fiscale</i>	99
4.6.1 - <i>I redditi da lavoro dipendente</i>	99
4.6.2 - <i>I redditi da lavoro autonomo e parasubordinato</i>	100
Conclusioni e prospettive future	103
Bibliografia	105
Appendice - Criteri per la generazione automatica del codice fiscale	109

Prefazione

Il progetto Eu-Silc (European statistics on income and living conditions), istituito con apposito regolamento dell'Unione europea (n. 1177/2003), risponde alla domanda crescente di informazioni da parte delle istituzioni nazionali ed europee, della comunità scientifica e della pubblica opinione sui fenomeni della povertà, dell'esclusione sociale e, più in generale, della qualità della vita delle famiglie.

L'Italia partecipa al progetto con l'indagine "Reddito e condizioni di vita", condotta ogni anno a partire dal 2004, i cui risultati confluiscono nei rapporti periodici dell'Unione europea sulla situazione sociale e sulla diffusione della povertà nei paesi dell'Unione. I questionari dell'indagine raccolgono informazioni sia sulle diverse fonti di reddito individuale e familiare, sia su altre importanti dimensioni della vita dei cittadini (partecipazione al mercato del lavoro, condizioni di salute, caratteristiche della casa e della zona di abitazione, ecc.). Il disegno campionario prevede, inoltre, una componente panel che consente di studiare gli aspetti dinamici dei fenomeni rilevati (le famiglie permangono nel campione per quattro anni consecutivi).

Il principale obiettivo del progetto Eu-Silc consiste nella rilevazione dei redditi individuali e familiari netti sia in complesso che in forma disaggregata per componente economica¹. Come in ogni altra rilevazione mediante intervista, il presupposto fondamentale per la validazione dei risultati è costituito dalla verifica della qualità dei dati raccolti.

L'incapacità dell'intervistato di quantificare con precisione eventi distanti nel tempo o semplicemente di ricordarne l'esistenza (in linea teorica potrebbero trascorrere 12 mesi dal momento dell'intervista all'anno di riferimento dell'indagine), l'effetto di interazione intervistato-intervistatore, la reticenza a dichiarare il vero, costituiscono alcuni esempi dei potenziali errori. D'altra parte, l'indagine ammette anche l'utilizzo di risposte *proxy*, con un possibile peggioramento della qualità dell'informazione fornita dall'intervistato per conto dell'unità di osservazione. Infine, si devono ricordare altre tipologie di errori di natura non campionaria, come quelli imputabili a registrazione, codifica, autoselezione dei rispondenti, ecc.

In questo contesto, la possibilità di integrare basi di dati campionarie e amministrative rappresenta una importante strategia per migliorare la qualità dell'informazione prodotta. Questo tipo di approccio è stato dapprima sperimentato, e in seguito applicato con successo nell'ambito delle prime quattro *waves* dell'indagine Eu-Silc.

Questo lavoro costituisce il primo esempio in Italia di applicazione di procedure di abbinamento tra più fonti del dato finalizzate alla produzione di statistiche sociali su scala nazionale. La sperimentazione e la realizzazione di metodi finalizzati all'arricchimento e al completamento dell'informazione da più fonti, seppure in un contesto orientato alla produzione di dati sui redditi, rappresentano utili linee guida applicabili anche in altri settori delle statistiche sociali.

In particolare, per limitare l'impatto degli errori di misurazione sui redditi riportati nel questionario di ciascun percettore, le fasi di controllo, correzione e imputazione dei dati si sono giovate dell'utilizzo di microdati di fonte fiscale, limitatamente a talune componenti reddituali².

¹ Le principali componenti di reddito sono: da lavoro dipendente, da lavoro autonomo, da capitale reale, da capitale finanziario, da trasferimenti distinti a loro volta in pensioni e altri sussidi. A partire dall'indagine Eu-Silc 2007 sarà obbligatoria la contestuale rilevazione dei redditi familiari e individuali lordi secondo le varie componenti.

² L'ipotesi sottostante è che l'informazione da fonte fiscale o più in generale da fonte amministrativa, opportunamente corretta e validata, sia più accurata rispetto a quella dell'indagine campionaria rilevata tramite questionario, ciò in conformità con la tesi sostenuta da Atkinson et al. (1995). Per i necessari distinguo si veda comunque il lavoro di P. Consolini et al. (2006).

Nell'implementare tale piano di correzione si è dovuto tenere conto di alcuni importanti requisiti che, come sempre in questo tipo di procedure, devono essere soddisfatti dall'insieme delle varie fonti (cfr. Paul van der Laan, 2000): i) le unità statistiche devono essere definite allo stesso modo nelle diverse fonti di dati; ii) tutte le fonti devono fare riferimento allo stesso collettivo statistico (copertura); iii) il contenuto informativo della variabili della fonte "donatrice" deve essere identico a quello delle variabili "obiettivo" dell'indagine Eu-Silc (armonizzazione delle variabili e delle classificazioni).

Per collegare le unità delle fonti fiscali a quelle dell'indagine campionaria è stata utilizzata la tecnica "abbinamento esatto" (*exact matching*) che permette di abbinare informazioni relative alla stessa unità statistica da distinte fonti del dato. In realtà, esistono diverse tipologie di abbinamento esatto; nel caso in esame si è ricorsi al semplice abbinamento uno a uno, in base al quale a ogni unità statistica corrisponde un solo record in ognuna delle fonti da integrare³. L'abbinamento fra diverse fonti del dato è stata realizzata tramite l'utilizzo di una "chiave di abbinamento" (*match key*) che identifica univocamente l'unità statistica comune ai diversi giacimenti informativi. Si tratta dunque di un abbinamento di tipo deterministico realizzato attraverso la chiave del codice fiscale dell'individuo⁴.

Le fonti amministrative deputate per il *matching* con i dati individuali dei redditi Eu-Silc sono: 1) il modello Cud per le categorie di redditi da lavoro, redditi da pensione e liquidazioni di fine rapporto; 2) il modello 730, oltre alle medesime componenti di reddito da lavoro/pensione presenti nel Cud, per i redditi da fabbricati o da terreni e per i redditi da lavoro autonomo non derivanti da attività professionale; 3) il modello Unico persone fisiche, per le stesse componenti di reddito rilevate nel 730, con la differenza che i redditi da lavoro autonomo sono riferiti ad attività professionale (titolari di reddito d'impresa e redditi derivante dall'esercizio di arti e professioni), mentre per redditi derivanti da rapporti di collaborazione coordinata e continuativa si tratta anche di attività manuali e operative, senza vincolo di subordinazione e d'impiego di mezzi organizzati, e con retribuzione periodica prestabilita (cariche di amministratore, sindaco o revisore di società, ecc.); 4) il Casellario centrale dei pensionati dell'Inps, per i redditi da pensione distinti secondo la tipologia (pensioni da lavoro, di reversibilità, di inabilità, di invalidità, sociali).

Rispetto alle variabili obiettivo del progetto Eu-Silc, i dati delle dichiarazioni dei redditi sono generalmente insufficienti e devono essere integrati con altre fonti. Le fonti fiscali esaminate non forniscono, infatti, alcuna informazione riguardo alle seguenti componenti di reddito:

- redditi soggetti a imposta sostitutiva (redditi da capitale finanziario);
- redditi esenti da Irpef (pensioni sociali, pensioni di guerra, pensioni di invalidità, indennità di accompagnamento, rendite per infortunio permanente);
- altre prestazioni sociali in denaro di natura non pensionistica, con l'esclusione delle liquidazioni di fine rapporto ed equipollenti e degli assegni al nucleo familiare dei dipendenti del privato;
- assegni al nucleo familiare dei dipendenti pubblici e dei pensionati;
- redditi da lavoro autonomo di talune figure professionali: agenti e rappresentanti di commercio;
- le deduzioni illegali e legali dall'imponibile (evasione ed elusione fiscale).

³ Per ulteriori approfondimenti di questo tipo di tecnica e di altre utilizzabili in questo campo si rinvia a Winkler (2007), Scanu (2005) e Nuccitelli et al. (2003).

⁴ Una volta effettuato l'aggancio con i dati amministrativi, il codice fiscale viene eliminato dalle basi di dati integrate, in ragione delle disposizioni in materia di protezione dei dati personali (cfr. d.lgs. n. 196/2003).

Per quanto concerne i redditi soggetti a imposta sostitutiva sarà presto avviato uno studio di fattibilità per verificare il contenuto informativo della fonte fiscale modello 770-Semplificato (QUADRO SK – redditi da attività finanziarie) e, al tempo stesso, per stimare i tempi di acquisizione del dato e i carichi di lavoro conseguenti al trattamento dei dati.

Riguardo alle prestazioni pensionistiche è stato utilizzato come fonte, a decorrere dall'indagine Eu-Silc 2004, il Casellario centrale dei pensionati dell'Inps congiuntamente alle fonti fiscali. Essa raccoglie informazioni non altrimenti rilevabili, quali la tipologia di pensione percepita e le pensioni esenti dalla tassazione personale sui redditi (Irpef). La fonte amministrativa in questione viene consolidata ai fini statistici⁵ per la rilevazione dei titolari di pensione e degli importi erogati alla data del 31 dicembre di ciascun anno.

In merito alle altre prestazioni sociali in denaro non pensionistiche⁶, è in progetto la ricognizione del contenuto informativo dell'archivio fiscale Inps – Gestione fiscale delle prestazioni non pensionistiche – già impiegato per la rilevazione Istat sui trattamenti monetari non pensionistici (indennità di disoccupazione, indennità di mobilità, parte delle integrazioni salariali dei cassintegrati).

Gli assegni al nucleo familiare dei dipendenti pubblici sono rilevati in aggregato nel Conto annuale della Ragioneria Generale dello Stato. Questa sta attualmente sviluppando un sistema informativo che raccoglierà, tra le altre voci, le informazioni individuali collegate ai cedolini degli stipendi (tra cui figurano gli assegni familiari) di una quota consistente dei dipendenti del pubblico impiego. Pertanto non si esclude la possibilità di collegare in futuro questa fonte amministrativa ai dati dell'indagine campionaria.

In merito alle componenti di reddito da lavoro autonomo relative ai rappresentanti e agenti di commercio, rilevate in apposito quadro del modello 770 (non incluso nel Cud), si è previsto di ufficializzare la richiesta all'Agenzia delle entrate a decorrere dall'anno di indagine Eu-Silc 2007.

L'impianto metodologico sui cui poggia il processo di integrazione tra i dati della rilevazione Papi e i dati di fonte amministrativa in Eu-Silc può essere riassunto nelle seguenti nove fasi operative:

- 1) Recupero delle informazioni individuali generatrici del codice fiscale individuale (nome, cognome, sesso, anno e luogo di nascita) e dei codici fiscali individuali rilevati presso i Comuni dal registro familiare e individuale Sigif (Sistema integrato di gestione dell'indagine sulle famiglie)⁷;
- 2) Analisi della correttezza e compatibilità dei codici fiscali individuali presenti in Sigif (con eventuale generazione dei codici fiscali inesatti o mancanti) e preparazione del file standard per lo scambio e la trasmissione dei dati con la Sogei⁸;
- 3) Predisposizione del modulo e delle specifiche (metadati) per la richiesta dati a Sogei, tramite la Direzione centrale dati e archivi amministrativi e registri statistici dell'Istat;
- 4) Predisposizione e aggiornamento delle istruzioni di lettura e conversione dei database fiscali di Sogei da un formato compresso (codice a serrare) a un formato direttamente

⁵ Il processo di consolidamento è realizzato attraverso la collaborazione tra l'Istat e l'ente previdenziale.

⁶ Esse includono le indennità di disoccupazione e di mobilità, le integrazioni salariali dei cassintegrati, le indennità di malattia e di inabilità temporanea assoluta, le indennità di maternità.

⁷ Questo registro fornisce la lista delle unità campionarie dell'indagine Eu-Silc. Esso comprende sia informazioni a livello individuale (nome, cognome, età, sesso, stato civile, codice fiscale) che informazioni a livello familiare (luogo e indirizzo di residenza).

⁸ La Sogei è parte integrante del sistema "Fisco", per il quale svolge la funzione di ente strumentale preposto al settore dell'*Information and Communication Technology* (Ict) con la missione di sviluppare e gestire il sistema informativo della fiscalità.

leggibile da un qualunque pacchetto statistico (variabili di lunghezza e posizione prefissata);

- 5) Predisposizione del piano di identificazione e correzione dei record duplicati, con relativo consolidamento delle informazioni relative allo stesso individuo riportate su più record;
- 6) Costituzione di un archivio integrato delle informazioni individuali sui redditi rilevabili presso le varie fonti⁹, attraverso la tecnica *exact matching* che permette di abbinare informazioni relative alla stessa unità statistica da distinte fonti del dato;
- 7) Analisi della coerenza dei valori reddituali dello stesso individuo riportati sulle varie fonti amministrative, controllo/correzione dei valori inammissibili sui redditi, riconciliazione di dati incoerenti;
- 8) Analisi comparativa tra i record del file di scambio rilasciato da Sogei (comprendente informazioni anagrafiche e fiscali sulle persone fisiche incluse nella lista di campionamento) e quelli del file relativo alle unità di analisi effettivamente intervistate in Eu-Silc;
- 9) Analisi della coerenza fra i dati fiscali e campionari Eu-Silc (unità effettivamente intervistate) e riconciliazione dei valori incoerenti.

Ognuna di queste fasi costituirà oggetto di analisi e approfondimento nei prossimi capitoli, che presentano le tecniche e le metodologie impiegate per la soluzione dei vari problemi operativi.

⁹ I cinque archivi di riferimento sono dati rispettivamente da: 1) archivio dei percettori che hanno solo dichiarazioni Cud; 2) archivio dei percettori di reddito con dichiarazioni Cud e modello 730; 3) archivio delle sole dichiarazioni 730; 4) archivio delle dichiarazioni Cud e modello Unico persone fisiche; 5) archivio delle sole dichiarazioni Unico persone fisiche.

Capitolo 1 – L’attribuzione dei codici fiscali individuali a partire dal patrimonio informativo di Sigif

1.1 - Introduzione

Per legare insieme dati individuali provenienti da diverse fonti è indispensabile una procedura, deterministica o probabilistica, che consenta di identificare quali coppie di record fra due o più basi di dati afferiscono alle stessa unità¹. Nell’approccio deterministico una coppia di record si definisce “abbinata” se i due record concordano esattamente su ciascun elemento costitutivo all’interno di un collettivo di identificativi denominati *match key* (Winkler et al., 2007). Questo tipo di approccio è quello generalmente seguito quando si dispone di identificativi univoci e accurati, come è il caso dei codici fiscali. La mancanza di identificativi unici o poco affidabili fa preferire, invece, approcci di tipo probabilistico che rimandano all’individuazione di una regola ottimale che tenga sotto controllo il livello degli errori generati da una determinata scelta di abbinamento tra record². La disponibilità, nell’ambito dell’indagine Eu-Silc, dei codici fiscali delle persone fisiche rilevati presso le anagrafi dei comuni campione, oppure indirettamente calcolabili tramite le caratteristiche demografiche e personali degli individui registrati (nome, cognome, sesso, anno e luogo di nascita) ha consentito l’adozione della strategia di record linkage deterministico. Sebbene questa opzione risulti più semplice e di immediata applicazione rispetto della seconda, essa rappresenta una scelta obbligata in quanto il processo di validazione e di attribuzione finale codici fiscali è interamente gestito all’esterno dalla Sogei, cioè dal partner tecnologico dell’Agenzia delle entrate che gestisce il relativo settore Ict. I risultati conseguiti nel tempo sul numero di unità del campione che trovano un abbinamento con i codici all’anagrafe tributaria (collettivo degli identificativi) mostrano la validità dell’impianto utilizzato. Nel presente capitolo verranno affrontate le questioni relative al calcolo/generazione del codice fiscale presunto e ai criteri e alle regole seguite in caso di assenza di una o più informazioni sui dati elementari.

1.2 - Il sistema Sigif

Il sistema informatico-informativo denominato Sigif (Sistema integrato di gestione delle indagini sulle famiglie) è stato istituito presso la Direzione centrale per le indagini su condizione e qualità della vita dell’Istat per gestire e monitorare i flussi informativi tra il sistema centrale, i comuni e i rilevatori. Tale sistema, realizzato in collaborazione con il Dipartimento degli standard tecnici, è implementato su DBMS Oracle e utilizza due canali per il popolamento del relativo database. Il primo canale prevede l’utilizzo del pacchetto Gefi che permette ai comuni non dotati di anagrafe informatizzata, ma che dispongono di almeno un personal computer, di inserire i dati su un supporto magnetico. Nella fattispecie i dati giungono all’Istat su floppy disk o via e-mail. Il secondo canale, dedicato ai comuni con anagrafe informatizzata (Sirc), l’invio avviene tramite accesso remoto per mezzo di una procedura che utilizza una suite di prodotti

¹ Cfr. Newcombe (1988).

² Cfr. Fellegi e Sunter (1969), Scanu (2003).

adatti allo scopo. In entrambi i casi, una procedura di controllo evidenzia eventuali errori presenti nelle trasmissioni³.

I vantaggi che presenta l'applicativo messo a disposizione dei comuni sono molti. In particolare fornisce informazioni sulle modalità di effettuazione dell'estrazione, facilita la registrazione dei nominativi estratti e ne permette un immediato e sicuro controllo (controllo della correttezza del passo di estrazione utilizzato, controllo della numerosità campionaria, controllo dei campi obbligatori, ecc.). Riduce, inoltre, i tempi del processo lavorativo, limita i disagi relativi alla difficoltà di reperire le famiglie per eventuali errori nei dati anagrafici sia al rilevatore sia al Comune, garantisce una maggiore sicurezza della privacy dei cittadini.

L'uso dell'applicativo inizialmente utilizzato per l'indagine continua sulle Forze di Lavoro è stato applicato anche all'indagine Eu-Silc, sin dalla prima edizione, e sarà nel tempo esteso a tutte le altre indagini campionarie sulle famiglie.

Una volta ricevuti dai Comuni campione i nominativi delle famiglie selezionate, che rappresentano il campione teorico delle famiglie da intervistare, il sistema Sigif fornisce una serie di informazioni che sono indispensabili per la creazione dei codici fiscali delle persone appartenenti ai nuclei familiari di interesse.

1.3 - La generazione del codice fiscale (presunto): aspetti generali

Il codice fiscale rappresenta la chiave che identifica univocamente il soggetto o la "persona fisica" nell'archivio dell'anagrafe tributaria. La rilevazione di questa chiave in relazione a ciascuna unità campionaria consente di legare le informazioni individuali di natura amministrativa con quelle riportate nel questionario. Il presupposto per l'applicazione di una procedura di generazione automatica del codice fiscale presunto è che il sistema Sirc dei comuni consente la rilevazione diretta dei codici fiscali soltanto per una quota, seppure cospicua, delle unità campionarie (circa tre quarti) e, d'altra parte, il valore in esso riportato non è sempre corretto.

Prima di applicare una procedura di generazione automatica del codice fiscale occorre comprendere la sua struttura. Esso si compone di 16 caratteri alfanumerici, dei quali i primi 15 caratteri sono funzione di alcune caratteristiche individuali quali: <Cognome>, <Nome>, <Data_nascita>, <Sesso> e <Luogo_nascita>; mentre il sedicesimo è rappresentato da un *digit* (lettera) di controllo che è funzione degli altri quindici caratteri⁴.

Le regole con cui vengono assegnati i 16 caratteri del codice fiscale della persona fisica sono così strutturate:

- 3 caratteri alfabetici per il cognome (in posizione 1, 2, 3);
- 3 caratteri alfabetici per il nome (in posizione 4, 5, 6);
- 2 caratteri numeri per l'anno di nascita (in posizione 7, 8);
- 1 carattere alfabetico per il mese di nascita (in posizione 9);
- 2 caratteri numeri per il giorno di nascita e per il sesso (in posizione 10, 11);
- 4 caratteri (1 alfabetico e 3 numerici) per il comune italiano di nascita o lo stato estero di nascita (in posizione 12, 13, 14, 15);
- 1 carattere di controllo (funzione dei primi 15 caratteri).

³ Per ulteriori approfondimenti dell'argomento si rinvia a Broccoli et al. (2006).

⁴ Il Decreto Ministeriale del 23 dicembre 1976, n. 13813 (Supplemento ordinario alla Gazzetta Ufficiale n. 345 del 29 dicembre 1976), esplicita in una serie di articoli i criteri che presiedono al calcolo del codice fiscale e ne dà un breve presentazione.

Per la corretta generazione del codice fiscale presunto è necessario dunque disporre di informazioni corrette su: cognome, nome, sesso, data e luogo di nascita del soggetto appartenente al campione. Tale condizione nella realtà non è sempre soddisfatta, per cui il codice presunto che si ottiene, applicando una qualunque procedura automatica a partire dai dati elementari inesatti presenti in Sirc, finisce con l'essere mancante o errato. Per altro verso, non è scontato che vi sia il perfetto allineamento tra il codice fiscale riportato in Sirc e quello desumibile dalla procedura automatica applicata ai dati elementari dello stesso sistema. Nei casi in cui essi presentino valori difformi si preferisce il codice rilevato dai comuni a quello presunto, a meno che il primo non assuma al suo interno valori inammissibili. Può infine accadere che ambedue le fonti di rilevazione dei codici fiscali (archivio dei comuni e procedura automatica) non siano in grado di generare codici fiscali formalmente validi. In questa circostanza si riporta un codice mancante e si demanda alla Sogei il compito di recuperare questa informazione sulla base dei dati anagrafici del soggetto interessato (ad es. indirizzo della propria residenza). L'output della procedura di generazione automatica è comunque il codice "presunto" e non quello vero, poiché non è escluso che due o più soggetti abbiano dati anagrafici tali da generare lo stesso codice fiscale, siano cioè "omocodici". In questi casi l'Agenzia delle entrate, che dispone dell'archivio generale di tutti i codici fiscali già emessi, provvede ad attribuire a ciascun soggetto un nuovo codice fiscale, calcolato a partire dal codice fiscale "base" comune a più soggetti⁵. Fortunatamente il fenomeno della omocodia riguarda un percentuale ridotta della popolazione italiana che sia aggira attorno al 2 per mille. Per una illustrazione dei criteri di calcolo del codice fiscale presunto si rinvia all'appendice A alla fine del volume.

1.4 - Generazione del codice fiscale: implementazione della procedura di calcolo⁶

Il calcolo del codice fiscale ha previsto una fase preliminare di lettura, correzione ed imputazione dell'informazioni di base contenute in Sigif. Tale procedura è stata modificata e affinata nel corso tempo. Così ad esempio, nel primo anno d'indagine (Eu-Silc 2004) il codice fiscale è stato ricavato dalle sole informazioni contenute in Sigif, cioè dalle variabili generatrici menzionate in precedenza (cognome, nome, ecc.) e dal codice fiscale rilevato dai comuni (Sirc). Mentre, per le successive *waves* è stata utilizzata come input aggiuntivo la lista dei codici fiscali delle persone fisiche validati dalla Sogei nell'anno precedente l'indagine (*T-1*)⁷. Tale lista comprende quei soggetti con codici fiscali validati al tempo *T-1* che dovranno essere reintervistati all'anno *T*. Vediamo in dettaglio le due procedure utilizzate.

La procedura per l'anno base

La procedura per la generazione dei codici fiscali per l'anno base 2004 ha richiesto una prima selezione delle informazioni del database Sigif per l'indagine Eu-Silc 2004.

⁵ La distinzione avviene effettuando, nell'ambito dei sette caratteri numerici, sistematiche sostituzioni di una o più cifre, a partire da quella più a destra, con corrispondenti caratteri alfabetici secondo la seguente tabella:

0	=	L	5	=	R
1	=	M	6	=	S
2	=	N	7	=	T
3	=	P	8	=	U
4	=	Q	9	=	V

⁶ Particolari ringraziamenti vanno alla dott.ssa R. Ricci per la sua collaborazione nella fase preliminare di caricamento dei dati dell'archivio Sigif e per l'implementazione della procedura output del file di scambio per la Sogei.

⁷ Questa scelta si è resa necessaria in quanto Sigif non prevedeva l'aggiornamento dei codici fiscali sulla scorta delle informazioni raccolte e validate negli anni passati.

Si è quindi proceduto ad ispezionare l'archivio (ridotto nel numero di variabili) distinguendo i casi in cui fossero o meno valorizzati i codici fiscali rilevati dai comuni (Sirc). In particolare, si rileva come i record con codice valorizzato fossero in totale 58.608, mentre quelli mancanti ammontavano a 20.302. Da una prima analisi di questi codici, si è potuta identificare una tipologia di errore molto frequente, rappresentata da 16 valori "0", che sono stati opportunamente sbiancati (riportati cioè a valore mancante). I record che non contenevano informazioni sui codici fiscali (rilevati dai comuni) e sulla data nascita della persona fisica sono stati tralasciati (66 record).

Nella fase successiva sono stati identificati ed eliminati gli errori di duplicazione⁸ che, a dimostrazione della buona qualità dell'archivio di input, sono risultati pari a sole 7 unità.

La fase seguente ha richiesto l'imputazione dei valori della variabile <codice_territoriale> attraverso un procedimento che utilizza come base del dato per la donazione: *A*) l'archivio storico integrato dei comuni, *B*) l'archivio dei codici stati esteri; e come chiave di abbinamento, rispettivamente: *a*) codici_Istat_comuni, *b*) codici_Istat_stati_esteri (presenti tanto in Sigif quanto negli archivi donatori)⁹. In assenza di informazioni sui codici Istat, il codice territoriale si ricava da una sezione dello stesso codice fiscale rilevato dai comuni (posizione 12, 13, 14, 15). Allo scopo di determinare il valore del codice territoriale, si è operata una partizione dell'insieme dei record individuali dell'archivio anagrafico¹⁰ Eu-Silc_2004 (senza duplicati) in 5 sottoinsiemi $S = \{S_A, S_B, S_C, S_D, S_E\}$, sulla base dei valori assunti dalle variabili <codice_Istat_comune_nascita> <codice_Istat_stato_estero_nascita>¹¹.

Il primo sottoinsieme S_A è formato da tutti quei record dove la variabile (*attributo*) <codice_Istat_comune_nascita> (6 digit) è valorizzata ed assume un valore (*t-upla*) che è compreso nell'elenco dei valori dei Codici Istat riferiti all'archivio storico integrato, di cui al precedente sottoparagrafo (il sottoinsieme S_A contiene 72.315 record). In definitiva, il primo sottoinsieme comprende tutti quei record ove è possibile donare il <codice_territoriale> dell'Agenzia del territorio dall'archivio storico integrato (donatore) tramite la chiave di abbinamento data dal <codice_Istat_comune_nascita>. Il secondo sottoinsieme S_B corrisponde ai casi la cui variabile <codice_Istat_comune_nascita> (6 digit) è valorizzata ma non trova alcun abbinamento (*not matched*) con il collettivo dei codici Istat dell'archivio storico integrato (in questo caso il sottoinsieme è vuoto). Il terzo sottoinsieme S_C è formato, invece, da tutti quei record dove il "codice_Istat_stato_estero_nascita" è valorizzato ed assume modalità compresa nell'elenco di tutti i valori dei codici Istat degli Stati esteri presenti del database della Ragioneria dello Stato (3.954 record). Anche in questo caso si può affermare che il sottoinsieme in questione definisce un collettivo in cui è possibile donare il <codice_territoriale> dell'Agenzia del territorio dall'archivio dei codici esteri (donatore) tramite l'altra chiave di abbinamento data dal <codice_Istat_stato_estero_nascita>. Il quarto sottoinsieme S_D è costituito dai casi in cui il codice Istat dello stato estero è valorizzato non trova alcuna corrispondenza con l'archivio (donatore) dei codici stati esteri (6 record). Infine, il quinto sottoinsieme S_E è definito sui casi in cui ambedue le variabili relative ai <codici_Istat> non sono valorizzate e pertanto non è possibile associare alcun <codice_territoriale> per tale via (2.562 record).

⁸ Tale operazione è stata eseguita sulla base della chiave composta, risultante dalla concatenazione, delle variabili <Codice_Fiscale-Cognome-Nome-Data_Nascita> della persona fisica.

⁹ Questo procedimento è stato implementato in SQL, attraverso l'utilizzo dell'operatore *join* esterno sinistro ($r_1 \triangleright_{LEFT} r_2$). Si ricorda che il risultato del *join* è costituito da una relazione sull'insieme dei attributi degli operandi e le sue *t-uple* sono ottenute combinando le *t-uple* degli operandi con valori uguali sugli attributi comuni. Per una trattazione approfondita su questo argomento si rinvia a Azteni et al. (1999).

¹⁰ In esso erano presenti 78.837 record.

¹¹ Si precisa che le due variabili in questione non possono assumere simultaneamente valori pieni.

Si deve osservare che, pur nell'incapacità di stabilire una corrispondenza tra valori dei <codici_Istat> di Sigif e degli altri due archivi (insiemi S_B, S_D e S_E), è ancora possibile ricavare il <codice_territoriale> da una sezione del codice fiscale presente in Sigif. All'interno del sottoinsieme S_E si possono distinguere tre casistiche in funzione della possibilità di:

- E₁) imputare il <codice_Istat_comune_nascita> dall'archivio storico integrato (S_{E1} : 997 record);
- E₂) imputare il <codice_Istat_stato_estero_nascita> dall'archivio dei codici degli stati esteri (S_{E2} : 223 record);
- E₃) impossibilità di imputare alternativamente i Codici Istat del comune di nascita o stato esteri (S_{E3} : 1.342 record).

Nei casi E₁ e E₂ la donazione dei valori dei codici_Istat avviene tramite la chiave di abbinamento data dal codice_territoriale, con una procedura analoga a quella implementata per i sottoinsiemi S_A e S_C ma con chiavi invertite.

Nel complesso, la procedura di imputazione del codice territoriale fallisce in 1.348 casi che sono tralasciati durante la fase di generazione del codice fiscale.

La fase successiva ha riguardato l'imputazione dei valori mancanti della variabile <Sesso> rilevata in Sigif tramite il codice fiscale delle persone fisiche riportato dai comuni (Sirc). In totale Sigif presenta 254 record con valori mancanti per la variabile sesso, di essi 224 possono essere imputati dalla variabile codice fiscale. I restanti 35 record non sono imputabili e pertanto vengono tralasciati in fase di generazione del codice fiscale.

Il passo seguente è stato finalizzato all'identificazione dei caratteri indicativi del "cognome". A tale scopo si è proceduto a creare due stringhe di caratteri (*n-ple ordinate*) rispettivamente ad indicare le "vocali" (Voc_cogn) e le "consonanti" (Con_cogn) del cognome nell'ordine in cui appaiono. Per identificare i caratteri indicativi del cognome occorre, preliminarmente, normalizzare le due stringhe attraverso la conversione dei caratteri in maiuscolo e l'eliminazione di caratteri speciali o di disturbo quali: apice, spazio, punto, virgola, trattino¹². Dall'analisi del contenuto delle due stringhe è possibile, quindi, individuare la condizione che definisce i tre caratteri distinti del cognome (cfr. par. 4.1.2). A titolo esemplificativo, la prima casistica "cognome con almeno 3 consonanti" è abbinata alla condizione: "presenza di un elemento (consonante) in 3^a posizione" lungo la stringa (vettore ordinato di elementi) delle consonanti del cognome (Con_cogn)¹³.

L'identificazione dei caratteri indicativi del nome rappresenta il consecutivo passo operativo. Per esso si applicano le stesse procedure di calcolo del cognome, con la variante della casistica: "Nome con almeno 4 consonanti" identificata dalla condizione: presenza di un elemento (consonante) in 4^a posizione" lungo la stringa <Con_cogn> (vettore ordinato di elementi).

Le tre successive fasi di calcolo del codice fiscale associano una cinquina di caratteri all'anno, mese, giorno di nascita e al sesso del soggetto. In sintesi, per ognuna di queste fasi si assegnano dei valori al contenuto delle variabili di base utilizzando una corrispondenza "uno ad uno".

Il calcolo del carattere di controllo rappresenta l'ultima fase per la generazione del codice fiscale ed è strutturata in quattro passi. Nel primo si genera un contatore che cumula i punteggi

¹² Se si utilizza, ad esempio, il codice Sas di programmazione, la generazione delle due stringhe si ottiene con l'utilizzo dell'istruzione: «compress(NOME_VAR,"...;...")» che elimina selettivamente i caratteri specificati lungo la stringa.

¹³ Per meglio comprendere il meccanismo di assegnazione dei tre caratteri del Cognome, consideriamo la seconda casistica definita come: "Cognome con due consonanti e una vocale". Essa, può essere a sua volta individuata dalla seguente condizione: "presenza di una consonante in 2^a posizione e contestuale assenza di una consonante in 3^a posizione lungo la stringa <Con_cogn>" e "presenza di un elemento (vocale) in 1^a posizione lungo la stringa <Voc_cogn>".

assegnati ai caratteri che occupano posizione dispari lungo la stringa formata dai primi 15 elementi del codice fiscale, in coerenza con il criterio definito nel precedente paragrafo 1.3. Nel secondo passo si cumulano i valori assegnati ai caratteri che giacciono nelle posizioni pari della stessa stringa. Il terzo passo somma il valore dei due contatori (dispari e pari), divide il totale per 26 ed calcola il resto di tale divisione. Infine nell'ultimo passo si assegna in una lettera dell'alfabeto al resto della divisione come esposto alla fine del paragrafo 1.3.

Vediamo in sintesi i risultati delle azioni intraprese in fase di controllo e correzione dei dati di base per la generazione del codice fiscale nel primo anno di indagine Eu-Silc (2004):

Tavola 1.1 – Costruzione dell'archivio di base per la generazione dei codice fiscali – Indagine Eu-Silc 2004

CARATTERISTICHE ARCHIVIO Sigif-Eu-Silc_04 (aggiornato al 31-gen-05)	Numero di record	Numero record imputabili da Codice fiscale	Note
Data-set in ingresso	78.910	–	- 20.302 codici fiscali mancanti - 58.608 codici fiscali valorizzati
Casi con valori <i>mancanti</i> nel codice fiscale e data di nascita	(-)66	–	Eliminazione
<i>Duplicati</i> rispetto alla chiave: "codice_fiscale-cognome-nome-data_nascita"	(-)7	–	Eliminazione
Dataset potenzialmente idoneo a generare un codice fiscale	78.837		
Dati con valori <i>mancanti</i> nel: Codice_Istat_comune/Stato estero	2.562	1.515	Imputazione tramite il codice fiscale di Sigif
Dati con valori <i>mancanti</i> in Codice_Istat comune/Stato estero e codice territoriale non assegnato	(-)1.342	–	Eliminazione
Dati con valori <i>errati</i> in Codice_Istat_comune/Stato estero e codice territoriale non assegnato	(-)6	–	Eliminazione
Dati con valori <i>mancanti</i> in: "sesso dell'individuo"	259	224	Imputazione tramite il codice fiscale di Sigif
Dati con valori <i>mancanti</i> nel codice fiscale e sesso	(-)35	–	Eliminazione
Dataset con informazioni tali da generare un codice fiscale pieno	77.456	–	

La procedura di calcolo per l'anno base ha consentito di imputare i codici fiscali di 19.158 record non altrimenti rilevati dai comuni. Per altro verso 58.595 codici, tratti dagli stessi comuni, sono stati scritti come validi nella procedura di output, anche se essi differivano dai codici fiscali generati con la procedura appena descritta (1.379 casi). Infine in 1.084 casi non si è riusciti a generare un codice fiscale valido.

La procedura per gli anni successivi all'anno base

La disponibilità di informazioni sui codici fiscali generati e validati all'anno *T-1* ha permesso di migliorare la procedura di calcolo e di focalizzare l'attenzione solo su un sottoinsieme di record, segnatamente il quarto di campione entrante.

Il primo passo della procedura di attribuzione dei codici fiscali è stato quello di recuperare le informazioni valide sui codici fiscali dei soggetti presenti nelle indagini agli anni *T* e *T-1*.

L'individuazione dei soggetti presenti nelle due indagini è stata realizzata in due passi; dapprima incrociando gli archivi con la chiave identificativa dei soggetti in Eu-Silc "KEYIND (PB030)" e successivamente, in caso di mancato aggancio, utilizzando come chiave alternativa quella composta dalla concatenazione di: <Cognome-Nome-Data_Nascita (Anno-Mese-Giorno)>. Il secondo passo si è reso necessario in quanto il disegno campionario ammette che una famiglia appartenente al campione al tempo $T-1$ possa fare parte del campione al tempo T senza essere necessariamente considerata come componente *panel* (in questo caso la chiave identificativa della famiglia, o dei relativi membri, non è più la stessa).

Una volta individuati i soggetti che appartengono alle due indagini, il passo successivo è quello di identificare le persone il cui codice fiscale risulti validato in precedenza dalla Sogei. Tali soggetti rappresentano quindi il primo output della procedura di attribuzione dei codici. I soggetti il cui codice fiscale non è validato, oppure quelli appartenenti alle famiglie del quarto campione entrante, costituiscono invece l'input per la fase di generazione dei codici fiscali. Quest'ultima fase ricalca le stesse linee guida dell'anno base, con alcune piccole varianti che hanno consentito di affinare il metodo di calcolo. Ad esempio, nell'indagine 2005 si è proceduto ad imputazione la variabile "Sesso" utilizzando contestualmente il "Nome" del soggetto ed il suo "Codice fiscale". Nell'indagine 2006, diversamente dagli anni precedenti, sono stati recuperati alcuni codici fiscali dai modelli 730 (40 casi) e dal Casellario centrale dei pensionati (180 casi). Inoltre, l'incoerenza tra il codice fiscale rilevato dai comuni e quello generato dalla procedura automatica è stata risolta ricorrendo a determinate ipotesi di lavoro.

Con riferimento all'indagine Eu-Silc 2005, il problema dell'attribuzione del codice fiscale ha riguardato 66.043 soggetti appartenenti al campione teorico di quell'anno.

La procedura di recupero dei codici fiscali all'anno $T-1$, relativi cioè agli individui del campione Eu-Silc 2004, ha consentito di acquisire 43.205 codici¹⁴ certificati come validi dalla Sogei. I restanti 2.660 codici fiscali non validati¹⁵, relativi a soggetti presenti nelle due indagini Eu-Silc 2004 e 2005, sono stati inclusi nell'insieme dei record da imputare per l'anno nuovo. I soggetti appartenenti al quarto di campione entrante risultavano in totale 20.178. Di questi, 15.258 possedevano un codice fiscale rilevato dai comuni, mentre i restanti 4.920 non aveva alcuna informazione in proposito (*missing value*). Le fasi di controllo e correzione che precedono la fase di generazione del codice fiscale sono riassunti nel prospetto seguente. Da esso si evince che l'archivio Sigif_Eu-Silc'05 viene dapprima ripulito dai possibili duplicati (22.836 record) e successivamente ispezionato, limitatamente allo spazio delle variabili elementari, per la selezione dei casi idonei a generare un codice fiscale (22.159 record).

¹⁴ Di questi codici, 43.187 sono stati recuperati tramite la chiave identificativa del database Eu-Silc: "KEYIND", mentre i rimanenti 18, relativi a record con chiavi identificative diverse (KEYIND) nei due anni d'indagine, sono stati agganciati con seguente la chiave composta: <Cognome-Nome-Data_Nascita (Anno_Mese_Giorno)>.

¹⁵ In particolare 1.074 record risultavano con codici fiscali mancanti.

Tavola 1.2 – Costruzione dell’archivio di base per la generazione dei codice fiscali – Indagine Eu-Silc 2005

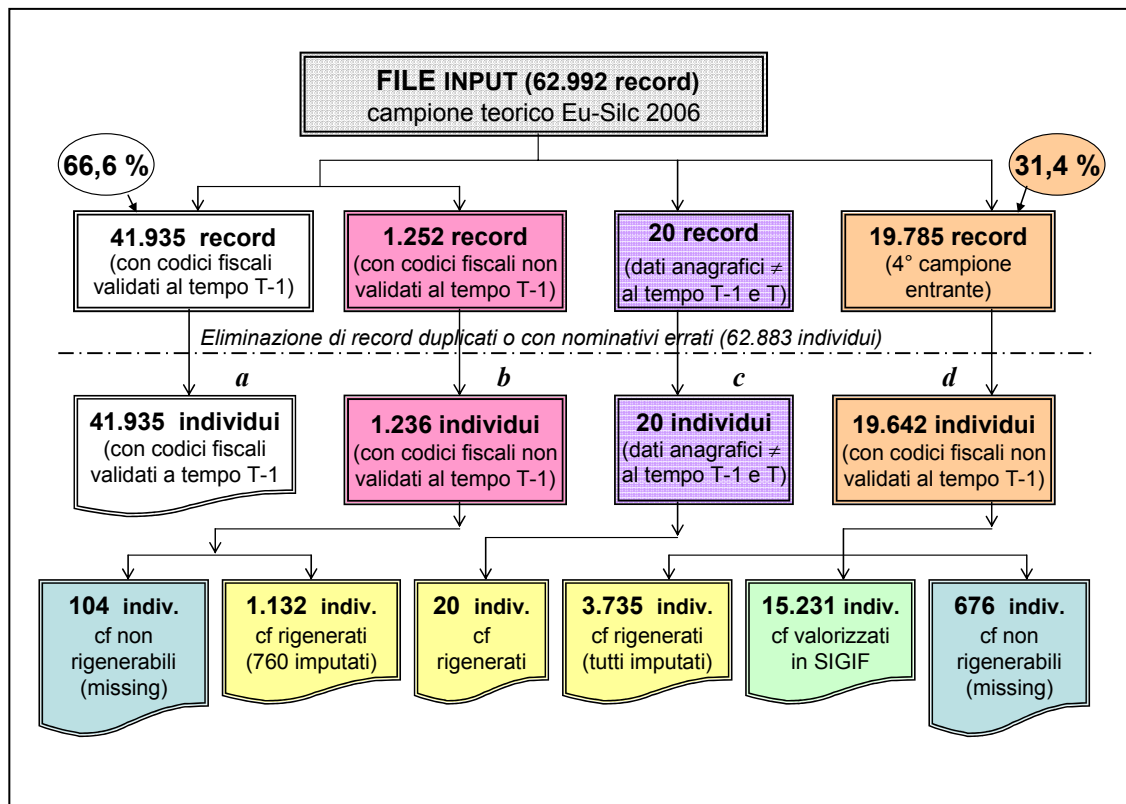
CARATTERISTICHE ARCHIVIO Sigif-Eu-Silc_05	Numero di record	Numero record imputabili da Codice fiscale	Note
Data-set in ingresso	66.043	–	
- Dati con codici fiscali validati al tempo T-1 (1° output)	43.205	–	Da includere direttamente nel file di output
- Dati con codici fiscali non validati al tempo T-1	2.660	–	Da rigenerare nuovamente
- Dati relativi al 4° di campione entrante	20.178	–	- 4.920 codici fiscali valorizzati - 15.258 codici fiscali mancanti
Data-set in ingresso per la generazione del codice fiscale	22.838	–	Input della procedura di generazione del codice
<i>Duplicati</i> rispetto alla chiave: "codice_fiscale-cognome-nome-data_nascita"	(-)2	–	Eliminazione
Data-set potenzialmente idoneo a generare un codice fiscale	22.836	–	
Dati con valori <i>mancanti</i> nel: Codice_istat_comune/Stato estero	979	331	Imputazione tramite il codice fiscale di Sigif
Dati con valori <i>mancanti</i> in Codice_istat comune/Stato estero e codice territoriale non assegnato	(-)648	–	Eliminazione
Dati con valori <i>errati</i> in Codice Codice_istat_comune/Stato estero e codice territoriale non assegnato	(-)1	–	Eliminazione
Dati con valori <i>mancanti</i> in: "sesso dell'individuo"	1.231	1.203	Imputazione tramite "codice fiscale" e "Nome" di Sigif
Dati con valori <i>mancanti</i> nel codice fiscale e sesso	(-)28	–	Eliminazione
Dataset con informazioni tali da generare un codice fiscale pieno	22.159	–	

La procedura di calcolo per l’anno 2005 (successivo all’anno base) ha consentito di imputare 5.415 codici fiscali, non altrimenti rilevabili. In altri 16.844 casi la procedura generava codici riferiti a soggetti con identificativi fiscali già presenti nell’archivio Sigif; 15.258 appartenenti al quarto campione entrante e 1.586 relativi a soggetti con codici non validati nell’anno precedente. La procedura ha emesso 1.649 codici fiscali che riportavano valori incoerenti rispetto al dato rilevato dai comuni. In questo caso, l’incoerenza è stata risolta assegnando priorità all’informazione rilevata dal sistema informativo dei comuni. Infine, in 577 casi non si è potuto attribuire alcun codice fiscale alla persona fisica, sia per l’impossibilità di generare il dato che per l’assenza di dati sul codice rilevati dai comuni.

L’attribuzione del codice fiscale dei soggetti appartenenti all’indagine Eu-Silc 2006, è stata sviluppata su quattro distinti segmenti del campione (figura 1.1): *a*) l’insieme di record riferiti a soggetti con codici fiscali validati al tempo *T-1* (41.935 casi); *b*) l’insieme di record associati a soggetti con codici fiscali non validati al tempo *T-1* con identica struttura anagrafica al tempo *T* (1.252 casi); *c*) l’insieme di record associati a soggetti con codici fiscali non validati al tempo *T-1* che cambiano i loro dati anagrafici al tempo *T* (20 casi); *d*) l’insieme di record

riguardanti i soggetti del quarto campione entrante di Eu-Silc (19.785 casi). Nel raggruppamento *a* sono stati semplicemente attribuiti i codici fiscali già validati nella precedente indagine. Per gli altri tre segmenti si è dovuto analizzare il contenuto informativo dei relativi record prima di procedere all'assegnazione dei codici. La base dati riferita a gruppi *b*, *c* e *d* era costituita da 20.898 record, dei quali solo 20.118 record contenevano informazioni sufficienti a generare un codice fiscale.

Figura 1.1 - Struttura del processo di attribuzione dei codici fiscali a partire dal campione teorico – Indagine Eu-Silc 2006



All'interno della base dati contenente informazioni tali da generare un codice fiscale sono stati imputati 4.515 codici fiscali, dei quali 760 afferenti al segmento *b* (soggetti con codici fiscali non validati al tempo *T-1* e stessi dati anagrafici nei due anni) e altri 3.735 al segmento *d* (quarto campione entrante), come da figura 1.1. La stessa procedura ha, inoltre, generato 15.603 codici fiscali per i soggetti con codice già valorizzato in Sigif; 15.231 dei quali appartenenti al quarto campione entrante ed altri 372 riferiti a persone con codici non validati nell'anno precedente. In 473 casi la procedura ha creato codici fiscali diversi rispetto a quelli rilevati dai comuni, mentre in altri 8 non è stata in grado di produrre valori a fronte di codici già presenti in Sigif¹⁶. In caso di incoerenza fra output del programma di generazione automatico e codice rilevato dai comuni (Sirc) si è stabilito di considerare valido quest'ultimo a meno di incompatibilità fra codice fiscale rilevato e caratteristiche individuali (età, sesso, nome e cognome) e di ammissibilità dei valori del codice territoriale. Infine, in 772 casi non è stato

¹⁶ Dal momento che questi codici risultavano non validati al tempo *T-1* ne è stato rimosso il relativo valore (missing).

possibile attribuire alcun codice fiscale alla persona fisica, sia per l'incapacità di generare valori che per l'assenza di informazioni rilevate dai comuni. In definitiva, il file di scambio inviato alla Sogei per la validazione dei codici fiscali inerenti ai soggetti del campione teorico Eu-Silc 2006 comprendeva 62.833 record (distinti in base alla chiave <progressivo_componente>).

Il seguente prospetto descrive le azioni intraprese nel piano di *editing* e imputazione sui dati elementari, precedente alla fase di generazione del codice fiscale. In questa fase sono stati, ad esempio, scartati 29 record, in quanto considerati duplicati sulla base della chiave "Codice_fiscale-Cognome-Nome-Data_nascita". Sempre in questo ambito stati eliminati altri 130 record poiché i nominativi dei soggetti ad essi associati erano mancanti o segnalavano situazioni di errore (ad esempio: "manca anagrafica", "senza cognome", ecc.). Per altro verso, l'imputazione dei valori mancanti sui dati elementari è stata realizzata tramite donazione delle informazioni già presenti nei codici fiscali rilevati presso i comuni. Così, ad esempio, l'imputazione dei dati mancanti sulla variabile sesso è avvenuta con successo in ben 1.245 casi su un totale di 1.249 record affetti da questo problema.

Tavola 1.3 – Costruzione dell'archivio di base per la generazione dei codice fiscali – Indagine Eu-Silc 2006

CARATTERISTICHE ARCHIVIO Sigif-Eu-Silc_06	Numero di record	Numero record imputabili da Codice fiscale	Note
Data-set in ingresso	62.992	–	
- Dati con codici fiscali validati al tempo T-1 (1° output)	41.935	–	Da includere direttamente nel file di output
- Dati con codici fiscali non validati al tempo T-1	1.252	–	Da rigenerare nuovamente
- Dati con anagrafica al tempo T differente rispetto al tempo T-1	20	–	Da rigenerare nuovamente
- Dati relativi al 4° di campione entrante	19.785	–	- 4.554 codici fiscali mancanti - 15.251 codici fiscali valorizzati
Data-set in ingresso per la generazione del codice fiscale	21.057	–	Input della procedura di generazione del codice
<i>Duplicati</i> rispetto alla chiave: "codice_fiscale-cognome-nome-data_nascita"	(-)29	–	Eliminazione
Dati con Cognome e/o Nome <i>mancanti</i> o <i>non meglio specificati</i>	(-)130	–	Eliminazione
Data-set potenzialmente idoneo a generare un codice fiscale	20.898	–	
Dati con valori <i>mancanti</i> nel: Codice_Istat_comune/Stato estero	1.080	311	Imputazione tramite il codice fiscale di Sigif
Dati con valori <i>mancanti</i> in Codice_Istat comune/Stato estero e codice territoriale non assegnato	(-)769	–	Eliminazione
Dati con valori <i>errati</i> in Codice Codice_Istat_comune/Stato estero	41	34	Imputazione codice territoriale tramite codice fiscale di Sigif
Dati con valori <i>errati</i> in Codice Codice_Istat_comune/Stato estero e codice territoriale non assegnato	(-)7	–	Eliminazione
Dati con valori <i>mancanti</i> in: "sesso dell'individuo"	1.249	1.245	Imputazione tramite "codice fiscale" e "Nome" di Sigif
Dati con valori <i>mancanti</i> nel codice fiscale e sesso	(-)4	–	Eliminazione
Dataset con informazioni tali da generare un codice fiscale pieno	20.118	–	

Nei tre anni di indagine 2004-2006 si è potuto osservare un costante miglioramento dell'accuratezza e della copertura delle informazioni relative ai codici fiscali ottenuti con la procedura di calcolo automatico o rilevati dai comuni. I risultati del processo di validazione (operato dalla Sogei) sui codici fiscali trasmessi dall'Istat hanno mostrato che la percentuale di codici coerenti cresce nel tempo, passando dal 92,1 per cento dell'anno 2004 al 95,6 per cento dell'anno 2005, per finire con il 96,2 per cento dell'anno 2006. Parallelamente nel corso degli stessi anni, risulta più che dimezzata la percentuale dei codici fiscali "errati o mancanti" non altrimenti recuperabili: dal 6 per cento di inizio indagine al 2,6 per cento del 2006. In altri termini, la percentuale di soggetti del campione che trovano un abbinamento (univoco) con i codici fiscali dell'anagrafe tributaria sale dal 94 per cento dell'anno 2004 al 97,4 dell'anno 2006.

Tavola 1.4 – Risultati del processo di validazione dei codici fiscali - Indagini Eu-Silc 2004-2006

RISULTATO DEL PROCESSO DI VALIDAZIONE DEI CODICI FISCALI INVIATI DELLA SOGEI	Anno di indagine		
	Eu-Silc 2004 % di casi	Eu-Silc 2005 % di casi	Eu-Silc 2006 % di casi
Coerenti ("0")	92,11	95,69	96,16
Errati ma recuperati dalla Sogei ("1")	0,34	0,28	0,18
Mancanti ma imputati dalla Sogei ("2")	1,59	0,79	1,08
Errati o mancanti irrecuperabili ("9")	5,96	3,24	2,58
Totale codici fiscali	100,00	100,00	100,00

Capitolo 2 – Le procedure per l’acquisizione ed il data-cleaning degli archivi amministrativi

2.1 - Introduzione

Il ricorso a tecniche di *data cleaning* è indispensabile quando si intende migliorare la qualità della propria base di dati di partenza, rendendola più adatta agli scopi degli utilizzatori finali (Chapman, 2005). Applicando queste tecniche si può ridurre, anche sensibilmente, l’errore di misurazione senza pregiudicare il contenuto informativo presente in origine. La struttura tipica di un processo di *data cleaning* comprende (Maletic and Marcus, 2000):

- La definizione/determinazione delle tipologie di errori presenti nei dati;
- La ricerca e l’identificazione dei casi che presentano errori;
- La correzione degli errori;
- La documentazione dei casi e delle tipologie di errore;
- La modifica delle procedure di generazione e strutturazione degli archivi atte a ridurre gli errori in futuro.

Applicate al caso nostro, le procedure di acquisizione e *data cleaning* sui dati di fonte amministrativa si sono sviluppate in fasi sequenziali, dando luogo: alla lettura/caricamento dei vari database rilasciati, al controllo numerico dei record in fase di lettura; all’analisi della rispondenza tra gli identificativi di record relativi all’archivio dei redditi e all’archivio anagrafico; all’individuazione e correzione dei record duplicati e più in generale degli errori di doppio conteggio; all’identificazione ed eliminazione dei valori anomali; alla generazione degli archivi ripuliti dai potenziali errori.

Quando gli archivi amministrativi comprendono più record (dichiarazioni) riferiti tutti alla stessa persona fisica, quindi tali da essere possibili fonti di errore di duplicazione, è importante verificare se le informazioni relative alle variabili di analisi presenti in essi siano in parte o in tutto replicate. Tale verifica è indispensabile per impedire che l’eliminazione di un record riferito allo stesso soggetto, ma contenente almeno una informazione univoca, finisca col produrre una perdita di informazione (*merge e purge problem*). Durante le fasi di *data cleaning* si è stabilito di mantenere comunque il database originale in modo da mantenere traccia delle modifiche apportate, garantendo la possibilità di ritornare al dato di partenza. La strategia utilizzata nella fase di identificazione degli errori di duplicazione è stata quella di osservare la presenza di ripetizioni nei valori assunti dalle variabili principali di reddito in corrispondenza di una stessa persona fisica e di un unico datore di lavoro¹. Tale tecnica riprende i principi della metodologia nota come *data profiling* che analizza le casistiche osservate sugli attributi individuali (Rahm and Do 2000). Si precisa che le procedure di *data cleaning* non si esauriscono attraverso il controllo/correzione degli errori per ogni singola fonte del dato ma si completano con l’analisi della coerenza dei valori assunti dalle varie unità statistiche su più fonti (*consistency*) che sarà oggetto di trattazione nel capitolo 4.

Nei paragrafi a seguire verranno illustrate le principali linee guida adottate nella fase di lettura ed i criteri utilizzati in fase di controllo e correzione.

¹ Per la consultazione di alcuni esempi pratici in cui trovano applicazione le tecniche di *data cleaning* si rinvia a Cody R. (1999).

2.2 - Il modello Cud/770

2.2.1 - L'acquisizione dei modelli Cud/770

L'acquisizione di archivi modelli Cud/770 da parte della Sogei ha comportato la creazione di diversi archivi sequenziali, ciascuno relativo ad ogni di acquisizione. Per la trasmissione dei dati verso l'esterno la Sogei utilizza particolari formati standard non immediatamente leggibili per i più comuni software utilizzati per l'elaborazione dati (es. Sas, Spss). Pertanto, le informazioni rilasciate devono essere successivamente ricostruite dagli utenti secondo formati più utili agli scopi dei loro processi. Inoltre per ragioni legate alla sicurezza dei dati personali, la Sogei ha disposto il rilascio dei microdati all'Istat ricorrendo a procedure di criptazione e decriptazione degli stessi.

Gli archivi fiscali sono caratterizzati da record aventi una struttura formata da una parte fissa ed una variabile; questa organizzazione è giustificata dal fatto che le informazioni effettivamente inserite durante la compilazione di un modello Cud/770 sono generalmente una piccola percentuale rispetto a quelle possibili. L'uso di una struttura a tracciato fisso produrrebbe dunque un enorme spreco di spazio, da qui la necessità di memorizzare solo i valori effettivamente presenti. Non potendo prevedere a priori quali informazioni sono effettivamente presenti in archivio, non vengono assegnate posizioni specifiche alle variabili in archivio. Per riconoscere i singoli campi si utilizza una struttura di campo virtuale o metacarpo: <codice_campo><tipo_campo><lunghezza_campo><valore_campo>. Questo metodo di rappresentazione dell'informazione viene comunemente indicata come "codici a serrare".

La procedura che ha consentito la ristrutturazione dell'archivio dei modelli Cud/770 è stata implementata per l'anno base (Eu-Silc 2004)², utilizzando criteri di "salto record" analoghi quelli implementati in precedenza in Istat. Di seguito illustriamo la struttura dell'archivio Cud/770_2004 (anno d'imposta 2003) che, rispetto all'anno precedente, presenta un'organizzazione meno complessa.

I record dell'archivio Cud/770_2004 hanno una lunghezza fisica di 2.072 caratteri e contengono informazioni fiscali (importi e tipologie di redditi, contributi, imposte ecc.) sulle persone fisiche nella sezione identificata dalle colonne 273-2072. Poiché ciascun campo corrispondente ad una variabile è formato da 24 caratteri, in linea teorica lungo ogni record vi possono essere al più 75 campi o blocchi. All'interno di un campo vi sono in testa 8 caratteri che identificano la variabile, mentre in coda ve ne sono 16 che contengono il valore (o il dato) associato a quella variabile. Gli 8 caratteri che identificano una variabile sono così strutturati: a) i primi due rappresentano la "parte" (già quadro) del modello Cud/770 (A: "DA", B: "DB", C: "DC") dove è inserita la variabile, b) i successivi tre indicano la "riga" (o linea) della dichiarazione, c) gli ultimi tre corrispondono alla "colonna" della dichiarazione e, insieme alla parte, definiscono la tipologia di variabile (ad es. reddito imponibile, imposta lorda, ecc.). Si precisa che per la gran parte delle tipologie di variabili fiscali (parte + colonna) del Cud/770 vi è al più una dichiarazione (riga). Tuttavia in taluni casi, è ammessa la presenza di più dichiarazioni (righe) a fronte della stessa tipologia di variabile. In questi casi, l'esigenza è quella di garantire un'informazione completa su una componente di reddito, su una tipologia professionale o altro, che può modificarsi lungo l'anno solare per tipo, modalità, ecc.

Per poter strutturare i dati contenuti nell'archivio Cud/770 senza perdita di informazione e secondo una logica che riflette le esigenze dell'utente, occorre individuare un criterio che consenta di associare un contenuto informativo (insieme di variabili e contenuti) all'unità logica più elementare. La scelta dell'unità logica è ricaduta sulla combinazione di cinque attributi: 1) identificativo telematico della dichiarazione, 2) numero progressivo della dichiarazione; 3)

² Il programma è stato realizzato in linguaggio Visual Basic.

identificativo del dichiarante (codice fiscale/partita IVA); 4) identificativo della persona fisica (codice fiscale), 5) progressivo della riga per la dichiarazione. L'insieme di questi cinque elementi definisce un'unità elementare, che indichiamo col termine di "dichiarazione", a cui associare il contenuto informativo presente in archivio. All'interno del tracciato record del Cud/770 è definito un campo formato da un carattere (posizionato in colonna 36) indicante la presenza o meno di informazioni fiscali del Cud per ciascun record. Tale campo prende il nome di "tipo record", per cui ogni qual volta esso assume valore "G" si procede alla lettura del relativo record.

La procedura utilizzata in fase di lettura dell'archivio è strutturata in due passi. Nel primo passo si eliminano le possibili fonti di errore, costituite da caratteri speciali che in ambiente Sas (ma più comunemente in tutti in pacchetti statistici) vengono interpretati come marcatori di fine file, successivamente si crea un vettore contenente tutte le possibili tipologie di variabili presenti in archivio. Tale vettore viene in seguito ordinato in funzione della parte del modello (DA, DB, DC) e del progressivo di colonna. Nel secondo passo, si procede ad associare il contenuto di ciascuna tipologia di variabile (elemento del vettore definito in precedenza) per ogni unità logica o dichiarazione. Pertanto, in fase di lettura (input) si definisce un punto di inizio (presenza di valore "G" e unità logica nuova) e fine lettura (segnalato dalla presenza di un'unità logica differente rispetto al rigo precedente) della sequenza di blocchi associati all'unità logica dell'archivio e quindi si memorizzano le relative informazioni. Successivamente, in fase di scrittura (output), si riporta (solo la prima volta e nel primo rigo) il nome delle variabili e (nei righe successivi) il contenuto informativo di ciascuna dichiarazione. In pratica, il contenuto di ciascuna variabile viene scritto in posizione corrispondente al campo in cui è definito il nome di quella variabile³.

Inoltre, per facilitare il compito dell'utente nella fase di definizione dei parametri e di esecuzione delle istruzioni, il programma di lettura in Visual Basic (VB) è stato sviluppato in modo tale da richiamare un'apposita maschera per l'indicazione del file di input, la digitazione del nome del file di output, la definizione della struttura del record di lettura (consentendo così di incorporare eventuali modifiche nell'organizzazione dell'archivio da un anno all'altro) e il lancio del programma.

Il file di input Cud/770_2004 rilasciato dalla Sogei conteneva 62.978 record a cui corrispondevano 41.906 codici fiscali distinti, mentre il file (output) ricavato dalla procedura di lettura conteneva 59.048, con associati 41.415 codici fiscali distinti. Si deduce, pertanto, che la procedura di lettura non era in grado di ottenere informazioni su 491 individui (codici fiscali distinti). In realtà, una parte di questi soggetti non possedevano alcuna informazione rilevante ai fini dell'analisi (dati fiscali nelle sole parti DA o DD), mentre un'altra parte, pur possedendo informazioni rilevanti, disponeva di una struttura di dati talmente complessa da non permettere la loro cattura.

La presenza di alcuni limiti del programma di lettura in VB che impedivano la cattura completa delle informazioni del Cud/770_2004 (circa l'un per cento delle persone fisiche venivano escluse) ha dato l'impulso per migliorare il medesimo per l'anno 2005.

³ Di seguito si riporta, a titolo di esempio, una sezione dell'archivio di output ottenuta dopo l'applicazione della procedura di lettura.

IdTel emati co	ProgrComCodFi scDi c	CodFi scPer c	TFNrgDB001	DB002	...
T0409281301571266700000010000001010110XXXXX		BAI CLD6XXXXXXXXXX	000100000000000161116		
T04092902500110345000000101694160970953XXXXX		BAI LGV6XXXXXXXXXX	00010000000000007281		
T04092910443784037000000100011694800787XXXXX		BAI LGV6XXXXXXXXXX	00010000000000010375		
T0409291259420357300000010000001002128XXXXX		BAI MSM5XXXXXXXXXX	00010000000000023254		
T04092813283084022000000100003798800787XXXXX		BAULGU4XXXXXXXXXX	00010000000000008606		
T04092309592709100000007700000001003373XXXXX		BAXBLY7XXXXXXXXXX	00010000000000012851		
T04092815092704613000000100843894970953XXXXX		BBTSVT3XXXXXXXXXX	00010000000000003581		
T04092919054611642000000100000027801973XXXXX		BBTSVT3XXXXXXXXXX	0001		
T04092919054611642000000100000027801973XXXXX		BBTSVT3XXXXXXXXXX	0002		

In questa sezione è possibile individuare i cinque elementi che compongono una unità logica: "IdTelematico = 1°", "ProgrCom = 2°", "CodFiscDic = 3°", "CodFiscPerc = 4°", "Nrg = 5°".

Da una prima ispezione, emergeva che il tracciato record dei modelli Cud/770_2005 differiva da quello del precedente anno, per il numero dei blocchi possibili (74 invece di 75) e per l'inclusione di una nuova informazione sull'impresa presso cui lavora il soggetto: settore Ateco.

Pertanto, si è sviluppato un programma, direttamente in ambiente Sas, che fosse in grado di caricare separatamente le tre parti del modello Cud/770: DA, DB e DC. Ciò ha portato alla creazione di tre archivi strutturati secondo le esigenze delle analisi statistiche da impostare.

Il programma di lettura ha previsto il caricamento condizionale dei dati, cioè vincolato al fatto che fosse soddisfatta la condizione che fossero presenti informazioni fiscali da modello Cud (record di tipo "G"). Per associare il "contenuto" di una variabile al suo "nome" (parte+colonna) sono stati creati appositamente due vettori, ognuno formato da 74 elementi, corrispondenti ai blocchi possibili scrivibili su ciascun record nella sezione delle colonne: 273-2049. Il primo vettore **n** (n_1, n_2, \dots, n_{74}) include come elementi i nomi delle variabili, formati dai primi otto caratteri del campo o blocco, mentre il secondo vettore **c** (c_1, c_2, \dots, c_{74}) riporta come elementi il contenuto delle precedenti variabili.

La lettura e la creazione dei vettori da sola non è sufficiente a creare le variabili con i rispettivi valori per ciascuna unità logica. Il passo successivo è quello di assegnare a ciascuna variabile il relativo contenuto. Se prendiamo ad esempio il caso delle 140 variabili incluse nella parte DB del Cud, il metodo che consente di associare il contenuto alla corrispondente variabile consiste nel creare un vettore "SA" formato da 140 elementi (nomi di variabili) e nell'associare con tecnica posizionale il valore contenuto nel database di origine. Pertanto, se nel database di origine l'*i-esimo* elemento del vettore **n** dei nomi delle variabili (74 elementi) ha come valore il nome della *j-esima* variabile fiscale si procede a creare la corrispondente variabile, assegnando il valore dell'*i-esimo* elemento del vettore **c** relativo al contenuto delle variabili (74 elementi). Il caricamento delle variabili avviene in modo sequenziale rispetto alla numerazione delle righe: DB001, DB002,...

Nell'ultima fase si fonde il contenuto informativo riferito alla stessa unità logica che nel file di input può essere riportato su più record fisici (non necessariamente contigui). Per realizzare questa fase ci si è avvalsi dell'operatore "lag" e si è ipotizzato che il contenuto informativo (colonne delle parti DA, DB, DC) relativo ad singola dichiarazione fiscale o unità logica (identificativo-progressivo-cf_datore-cf_percettore-riga) possa essere frazionato al più su cinque record. Si tratta di un'ipotesi alquanto ragionevole in quanto il numero massimo di colonne del Cud/770, ovvero dati fiscali inclusi nelle varie parti/sezioni, è pari a 265, mentre il numero massimo di campi (dati fiscali) contenuti in 5 record è pari 370 (74 campi ripetuti su 5 record).

In sintesi, nell'anno 2005 il file di input della Sogei conteneva 57.046 record associati a 36.337 codici fiscali distinti, mentre il file di output della programma di lettura in ambiente Sas era composto da 56.328 dichiarazioni (distinte) riferite a 35.704 persone fisiche (codici fiscali distinti). I 633 codici fiscali delle persone fisiche escluse in fase di generazione del dataset di output riguardavano soggetti che non possedevano alcuna informazione di interesse⁴. Pertanto a differenza del precedente programma in VB, la nuova procedura di lettura in Sas garantiva l'assenza di perdite di informazione sui percettori di reddito.

L'analisi degli archivi Sogei e dei modelli 770 ha portato alla creazione di un unico archivio comprendente le informazioni sulle parti DB e DC.

⁴ Nella generazione del file di output si è imposto il vincolo della presenza di almeno un dato fiscale (colonna) della parte DB o della parte DC – esclusa la sezione dei dati Inail.

Tavola 2.1 – Struttura del file Cud/770 utilizzato nell’indagine Eu-Silc 2005

NUMERO DI DUPLICATI	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
1	24.817	69,51	24.817	69,51
2	6.433	18,02	31.250	87,53
3	2.276	6,37	33.526	93,90
4	1.114	3,12	34.640	97,02
5+	1.064	2,98	35.704	100,00

2.2.2 - Le operazioni di controllo, correzione e consolidamento dei dati del Cud/770

Le procedure di controllo e correzione dei dati amministrativi sono state sviluppate con l’obiettivo di individuare ed eliminare gli errori di duplicazione presenti nelle principali poste di reddito del Cud/770. Il programma utilizzato a tale scopo è stato strutturato in modo da eliminare a cascata gli errori di duplicazione presenti nelle poste fiscali, ovvero nell’ordine in cui esse compaiono nel modello di dichiarazione dei redditi. Parallelamente, si è provveduto a rimuovere gli eventuali errori di doppio conteggio sui redditi conguagliati.

Tra le poste di reddito, presenti nel modello Cud/770, riveste un ruolo di grande rilevanza il reddito complessivo da lavoro dipendente e assimilati (DB1). Con riferimento all’indagine Eu-Silc 2005, si rileva che ben 686 persone (codici fiscali distinti) presentano dichiarazioni doppie su DB1, cioè mostrano importi analoghi sulle due dichiarazioni, mentre altre 1.076 persone riportano dichiarazioni plurime su DB1 (almeno tre dichiarazioni contenenti valori identici sulla stessa variabile di reddito). Per identificare le dichiarazioni doppie si è utilizzata la condizione che la varianza della variabile DB1 nelle due osservazioni ripetute (dichiarazioni riferite alla stessa persona fisica) fosse compresa tra 0 ed 1: $0 < \text{var}(\text{DB1}) < 1$. Per altro verso, l’identificazione e la rimozione delle dichiarazioni plurime su DB1 è stata realizzata in due fasi. Dapprima sono state identificate come plurime quelle dichiarazioni (almeno tre) riferite alla stessa persona e allo stesso datore di lavoro che soddisfano la condizione $0 < \text{var}(\text{DB1}) < 1$ con conseguente rimozione. In seguito sono state indicate come plurime e successivamente rimosse le tre o più dichiarazioni riferite allo stesso soggetto di identico valore (positivo). Nel primo caso (criterio della varianza) le dichiarazioni plurime erano associate a 834 persone fisiche, mentre nel secondo (criterio dell’esatto importo) riguardavano a 272 soggetti. La procedura di identificazione e correzione dei duplicati sulla variabile DB1 è stata rivista e migliorata nel corso dell’indagine Eu-Silc_2006 anche in relazione all’aumentato numero di casi coinvolti. Infatti nell’ultimo anno indagine su un totale di 35.501 persone fisiche aventi almeno una dichiarazione Cud, ben 2.290 presentavano dichiarazioni doppie su DB1 ed altre 1.741 esibivano dichiarazioni plurime sulla stessa variabile (stessa persona e allo stesso datore di lavoro e $0 < \text{var}(\text{DB1}) < 1$). Pertanto, si è stabilito di affiancare alla procedura di identificazione dei duplicati tramite “ $0 < \text{var}(\text{DB1}) < 1$ ” altre due condizioni: 1) il datore di lavoro doveva essere identico nelle due dichiarazioni; 1) l’imponibile previdenziale (DC1_4) doveva restare uguale negli stessi record. Tale affinamento ha portato al recupero di 61 dichiarazioni sulla variabile DB1 che venivano erroneamente eliminate in fase di correzione. Analogamente, in fase di correzione dei record plurimi, sono stati utilizzati ulteriori criteri per il recupero delle dichiarazioni, qualora pur in presenza di stessi importi, persone e datore le mensilità retribuite differivano tra i vari record. Tale aggiornamento ha portato al recupero di 45 dichiarazioni erroneamente eliminate. Inoltre si è voluta verificare la robustezza della procedura basata sulla chiave composita $\langle \text{Cf_perc-Cf_datore-DB1} \rangle$ rispetto alla chiave $\langle \text{Cf_perc-Cf_datore-DB1-Flag_12mesi_retr} \rangle$ e si ottenuto lo stesso risultato.

Una volta eliminati gli errori di duplicazione su DB1, il passo successivo è stato quello di identificare ed eliminare i potenziali errori di doppio conteggio sui redditi conguagliati. Un tipico errore di doppio conteggio sui redditi di questa natura sorge quando vi sono due dichiarazioni relative allo stesso individuo; una compilata dal nuovo datore di lavoro riportante il

valore del reddito conguagliato dal vecchio datore (esso rappresenta un di cui del reddito DB1 riportato della sezione principale) e l'altra compilata dal vecchio datore, riportante il reddito nella sola sezione principale dei dati fiscali (DB1). Pertanto, per individuare gli errori di doppio conteggio sui conguagli devono essere presenti tre condizioni: 1) due dichiarazioni riferite allo stesso individuo; 2) le due dichiarazioni devono riportare lo stesso datore di lavoro, una volta identificato come datore attuale (codice fiscale del dichiarante) un'altra come datore del precedente rapporto di lavoro (DB50); 3) gli importi di reddito esposti nella sezione principale: *Dati fiscali* (variabile DB1) equivalgono agli importi registrati nella sezione: *Dati relativi ai conguagli* (variabile DB53), come di seguito illustrato.

Figura 2.1 – Sezione: Dati fiscali - Parte B - Cud/770_2005

PARTE B
DATI FISCALI

DATI PER LA EVENTUALE COMPILAZIONE DELLA DICHIARAZIONE DEI REDDITI

Redditi per i quali è possibile fruire della deduzione di cui all'art.11 e delle detrazioni di cui all'art.14 del TUIR
1

Redditi per i quali è possibile fruire della deduzione di cui all'art.11, c. 1 del TUIR
2

Applicazione maggiore ritenuta
3

Richiesta di non applicazione della deduzione di cui all'art.11 del TUIR
4

Numero di giorni per i quali spettano le deduzioni di cui all'art.11 commi 2 e 3 del TUIR

Redditi derivanti dall'emersione del lavoro sommerso
5

Lavoro dipendente
6

Pensione
7

Credito d'imposta per le imposte pagate all'estero
8

Anno di percezione reddito estero
9

Reddito prodotto all'estero
10

Imposta estera definitiva
11

Ritenute Irpef
12

Addizionale regionale all'Irpef determinata sulla somma dei redditi di cui ai punti 1 e 2
13

Addizionale comunale all'Irpef determinata sulla somma dei redditi di cui ai punti 1 e 2
14

Ritenute Irpef sospese
15

Addizionale regionale all'Irpef sospesa
16

Addizionale comunale all'Irpef sospesa
17

Primo acconto Irpef trattenuto nell'anno
18

Secondo o unico acconto Irpef trattenuto nell'anno
19

Acconti sospesi
20

Credito Irpef non rimborsato
21

Credito di addizionale regionale all'Irpef non rimborsato
22

Credito di addizionale comunale all'Irpef non rimborsato
23

DATI RELATIVI AI CONGUAGLI

INCAPIENZA IN SEDE DI CONGUAGLIO

Irpef da trattenere dal sostituto successivamente al 28 febbraio
46

Irpef da versare all'erario da parte del dipendente
47

REDDITI EROGATI DA ALTRI SOGGETTI

Totale redditi conguagliato già compreso nel punto 1
48

Totale redditi conguagliato già compreso nel punto 2
49

Codice fiscale
50

Causa
51

Reddito conguagliato
53

Ritenute
54

Ritenute sospese
55

Addizionale regionale
56

Addizionale regionale sospesa
57

Addizionale comunale
58

Addizionale comunale sospesa
59

Per l'anno 2005 il numero totale di dichiarazioni generanti errori di doppio conteggio sui redditi conguagliati è pari a 312.

La fase seguente ha riguardato l'identificazione ed eliminazione degli errori di duplicazione sulla componente di reddito rappresentata dagli arretrati, cioè dai compensi relativi agli anni precedenti soggetti a tassazione separata. Per l'identificazione dei duplicati è stato utilizzato il criterio dell'esatto importo applicato a ciascun anno d'imposta, cioè sono state selezionate quelle dichiarazioni che mostravano valori ripetuti sulla terna di variabili date da: 1) codice fiscale del percettore, 2) importo dell'arretrato (DB60), 3) anno d'imposta (DB69). Il problema in esame si dimostra essere poco rilevante, coinvolgendo appena 16 titolari di reddito e 21 dichiarazioni. Tra le varie dichiarazioni duplicate, si è deciso di mantenere quella che possedeva maggiore contenuto informativo, nel senso che conteneva il maggior numero di variabili numeriche valorizzate: "NVAR_NUM_VAL".

L'investigazione sui potenziali errori di duplicazione nelle dichiarazioni relative al Tfr e indennità equipollenti si è basata, analogamente al caso degli arretrati, sul criterio dell'esatto

importo applicato, però, alla data (giorno, mese, anno) di erogazione. In pratica, si è proceduto a selezionare quelle osservazioni che mostravano valori replicati sull'insieme dei cinque seguenti elementi: il codice fiscale del lavoratore, l'importo del Tfr o indennità equipollente, giorno-mese-anno di liquidazione. La procedura descritta ha portato all'individuazione ed alla conseguente rimozione di 58 dichiarazioni duplicate sul Tfr relative a 54 soggetti. Anche in questo caso si è stabilito di inserire nel file finale quel record avente maggiore contenuto informativo (NVAR_NUM_VAL) tra le varie dichiarazioni duplicate.

Lo stesso criterio di selezione e rimozione dei duplicati è stato applicato al caso della variabile indicante le competenze correnti dei lavoratori dipendenti del privato, esposta nella sezione 1: "Dati previdenziali e assistenziali Inps" della parte C (DC9). Le variabili utilizzate nell'individuazione dei duplicati riguardano: il codice fiscale del lavoratore, l'importo delle competenze correnti "DC9" e i 12 mesi lavorati "DC15" (0='lavora', 1='non lavora'). Per l'anno 2005, le dichiarazioni eliminate ammontano a 86 e interessano 80 persone fisiche.

Il criterio citato, quando applicato alle retribuzioni particolari (DC29) della Sezione 2 - Parte C, porta alla cancellazione di 24 dichiarazioni duplicate afferenti a 23 persone fisiche. Con riferimento ai redditi co.co.co. "DC58" della Sezione 4 - Parte C (codice fiscale, importo e data erogazione), si riscontra un esiguo numero di casi duplicati: cinque dichiarazioni su altrettanti percettori. Infine, appare rilevante la presenza di errori di duplicazione sui redditi dei dipendenti pubblici, segnatamente per la componente delle retribuzioni fisse e continuative (DC85). dove figurano 715 dichiarazioni duplicate da eliminare per un totale di 557 persone fisiche.

Il file finale dopo le procedure di eliminazione dei duplicati contiene 53.312 osservazioni.

2.3 – Il modello 730

2.3.1 - L'acquisizione dei modelli 730

L'acquisizione degli archivi dei modelli 730 ha implicato la generazione di due distinti archivi: uno riportante le informazioni anagrafiche e l'altro contenente i dati fiscali sui redditi dei dichiaranti. Identicamente al caso dei modelli Cud/770, anche per i modelli 730 viene utilizzato da parte della Sogei un sistema di archiviazione e trasmissione dei dati all'esterno basato su formati compressi non immediatamente leggibili per i più comuni software utilizzati per l'elaborazione dati (ad esempio: Sas, Spss). Le informazioni di interesse devono essere ricostruite in seguito dagli utenti secondo formati più utili agli scopi dei loro processi. Il sistema utilizzato per riconoscere i singoli campi si basa su una struttura di campo virtuale o metacarpo (<codice campo><tipo campo><lunghezza campo><valore campo>) analoga al caso dei modelli 770. Per recuperare i dati reddituali relativi a ciascun dichiarante, senza perdita di informazione e senza errori di attribuzione, è stato necessario fissare i criteri per decidere quando le informazioni estratte dall'archivio appartenevano ad un determinato soggetto. Nel caso dell'archivio dei modelli 730 si è potuto verificare empiricamente che codice fiscale è un "separatore" affidabile all'interno della struttura dei dati predisposta dalla Sogei. Durante la fase di ricostruzione delle informazioni individuali sui redditi si è pertanto proceduto a caricare in modo sequenziale le informazioni di ciascun individuo presenti su più record fisici contigui tra di loro.

La procedura che ha consentito la ristrutturazione dell'archivio dei modelli 730 è stata implementata con il programma in linguaggio VB tanto per il primo anno di indagine (Eu-Silc 2004) quanto per il secondo. La procedura è strutturata in due passi. Nel primo passo si eliminano le possibili fonti di errore, costituite da caratteri speciali che in altri ambienti vengono interpretati come marcatori di fine file, successivamente si crea un vettore contenente tutte le possibili tipologie (nomi) di variabili presenti in archivio. Nel secondo passo, si procede ad associare il

contenuto di ciascuna tipologia (nome) di variabile per ogni unità logica definita dal valore del codice fiscale. Il programma di lettura utilizza quindi un criterio di “salto record” basato sull’esame dei valori assunti del codice fiscale tra una riga e all’altra. Pertanto, in fase di lettura (*input*) si memorizzano le informazioni pertinenti a ciascun codice fiscale, successivamente, in fase di scrittura (*output*), si riporta (solo la prima volta e nel primo rigo) il nome delle variabili e (nei rigli successivi) il contenuto delle variabili associato a ciascuna persona fisica.

I record dell’archivio 730_2005 hanno una lunghezza fisica di 1.900 caratteri e contengono tanto informazioni anagrafiche (record tipo “E”) quanto dati fiscali (record tipo “R”). I dati fiscali (importi e tipologie di redditi, contributi, imposte ecc.) sulle persone fisiche figurano nella sezione identificata dalle colonne 83-1874. Poiché ciascun campo corrispondente ad una variabile è formato da 32 caratteri, in linea teorica lungo ogni record vi possono essere al più 56 campi o blocchi. All’interno di un campo vi sono in testa otto caratteri che identificano la variabile, in quattordicesima posizione è indicato il tipo di variabile (*A* = campo numerico intero; *B* = campo numerico decimale o perc.; *C* = campo alfanumerico; *D* = campo indicante un flag; *E* = campo di tipo data; *F* = campo codice fiscale), mentre in coda vi sono 18 caratteri che contengono il valore (o il dato) associato a quella variabile. Per consentire all’utente di gestire autonomamente i principali parametri di lettura nonché di eseguire facilmente il programma, la procedura in VB è stata sviluppata in modo tale da richiamare un’apposita maschera per l’indicazione del file di input, la digitazione del nome del file di output, la definizione della struttura del record di lettura e il lancio del programma.

Per l’indagine Eu-Silc 2004, il file di input dei modelli 730 rilasciato dalla Sogei conteneva 41.593 record (esclusi i record di testa e di coda) a cui corrispondevano 17.643 codici fiscali distinti, mentre il file (output) ricavato dalla procedura di lettura conteneva 17.757 record con associati 17.643 codici fiscali distinti. Si deduce, pertanto, che la procedura di lettura era in grado di ottenere informazioni su tutti i dichiaranti, senza perdita di informazione. Anche per l’indagine Eu-Silc 2005 viene garantita l’assenza di perdita di informazioni durante la generazione del file di output della procedura di lettura. In questo caso, tanto per il file di input della Sogei quanto per il file output i codici fiscali distinti delle persone fisiche sono 15.741.

2.3.2 - Le operazioni di controllo, correzione e consolidamento dei dati del 730

Le procedure di controllo e correzione dei dati del 730 sono state focalizzate all’individuazione ed eliminazione degli errori di duplicazione presenti nelle rispettive dichiarazioni. Per identificare i duplicati si è calcolata la frequenza delle dichiarazioni afferenti a ciascuna coppia di valori dati da: 1) codice fiscale e 2) progressivo delle dichiarazioni. La presenza di due o più dichiarazioni in capo ad uno stesso soggetto individuano una situazione in cui sussiste un errore di duplicazione. L’impatto di tali errori nell’archivio 730 è tuttavia molto limitato; infatti nell’indagine Eu-Silc 2004 coinvolge appena 84 dichiarazioni duplicate per un totale di 81 persone fisiche; la situazione è ancora più circoscritta nell’indagine successiva (Eu-Silc 2005) dove figurano 63 duplicazioni in corrispondenza di 61 dichiaranti. Più marcato è l’impatto dell’errore insito nel mancato aggancio delle informazioni sui redditi dei “coniugi del dichiarante”, quando si utilizzano i soli record individuali (record tipo “R”) dei modelli 730. Infatti, per poter rilevare tali percettori con le rispettive componenti di reddito, occorre stabilire una corrispondenza tra i valori di reddito, esposti nelle dichiarazioni individuali sui redditi (record tipo “R”), e i codici fiscali dei coniugi presenti nella sezione dati anagrafici dei modelli 730 (record tipo “E”). A tal fine si è utilizzato, in fase di consolidamento dei dati, il progressivo della dichiarazione come chiave di aggancio tra i redditi e codici fiscali dei coniugi, permettendo così di attribuire in 2.433 casi le componenti esposte nella seconda colonna del prospetto riepilogativo al rispettivo percettore (coniuge). Di seguito si illustra il prospetto di liquidazione del modello 730 anno d’imposta 2004.

Figura 2.2 – Prospetto di liquidazione relativa alla assistenza fiscale prestata - Mod. 730_2005

730		MODELLO 730-3 redditi 2004		Modello N. <input type="text"/>		EURO	
agenzia entrate		prospetto di liquidazione relativo alla assistenza fiscale prestata		730 <input type="checkbox"/> rettificativo		730 <input type="checkbox"/> integrativo	
SOSTITUTO D'IMPOSTA O C.A.F.		CODICE FISCALE		COGNOME E NOME O DENOMINAZIONE		N. DI ISCRIZIONE ALL'ALBO	
RESPONSABILE DEL C.A.F.							
DICHIARANTE						APPLICAZIONE CLAUSOLA DI SALVAGUARDIA	
CONIUGE DICHIARANTE							
RIEPILOGO DEI REDDITI		1	DICHIARANTE	2	CONIUGE		
1	REDDITI DOMINICALI		,00		,00		
2	REDDITI AGRARI		,00		,00		
3	REDDITI DEI FABBRICATI		,00		,00		
4	REDDITI DI LAVORO DIPENDENTE E ASSIMILATI		,00		,00		
5	ALTRI REDDITI		,00		,00		
CALCOLO DELL'IRPEF		1	DICHIARANTE	2	CONIUGE	3	TOTALE
6	REDDITO COMPLESSIVO		,00		,00		
7	DEDUZIONE PER ABITAZIONE PRINCIPALE		,00		,00		
8	ONERI DEDUCIBILI		,00		,00		
9	DEDUZIONE PER LA PROGRESSIVITÀ DELL'IMPOSIZIONE (ART. 11 DEL TUIR)		,00		,00		
10	REDDITO IMPONIBILE		,00		,00		
11	Compensi per attività sportive dilettantistiche con ritenuta a titolo d'imposta		,00		,00		
12	IMPOSTA LORDA		,00		,00		
13	Detrazione per coniuge a carico		,00		,00		
14	Detrazione per familiari a carico		,00		,00		
15	Detrazioni per lavoro dipendente o pensione		,00		,00		
16	Ulteriore detrazione per redditi di lavoro dipendente di durata inferiore all'anno o pensione		,00		,00		
17	Detrazione per gli oneri di cui alla sezione I e III del quadro E		,00		,00		
18	Altre detrazioni di cui alla sezione IV del quadro E		,00		,00		
19	Credito d'imposta per il riacquisto della prima casa		,00		,00		
20	Credito d'imposta per l'incremento dell'occupazione		,00		,00		
21	TOTALE DETRAZIONI E CREDITO DI IMPOSTA		,00		,00		,00
22	IMPOSTA NETTA		,00		,00		,00
23	CREDITI D'IMPOSTA SULLE IMPOSTE PAGATE ALL'ESTERO		,00		,00		,00
24	CREDITO D'IMPOSTA PER CANONI NON PERCEPTI		,00		,00		,00
25	RITENUTE		,00		,00		,00
26	DIFFERENZA		,00		,00		,00
27	ECCEDENZA DELL'IRPEF RISULTANTE DALLA PRECEDENTE DICHIARAZIONE		,00		,00		,00
28	ACCONTI VERSATI						,00
CALCOLO DELLE ADDIZIONALI REGIONALE E COMUNALE ALL'IRPEF		1	DICHIARANTE	2	CONIUGE		
29	REDDITO IMPONIBILE		,00		,00		
30	ADDIZIONALE REGIONALE ALL'IRPEF DOVUTA		,00		,00		
31	ADDIZIONALE REGIONALE ALL'IRPEF RISULTANTE DALLA CERTIFICAZIONE		,00		,00		
32	ECCEDENZA DELL'ADDIZIONALE REGIONALE ALL'IRPEF RISULTANTE DALLA PRECEDENTE DICHIARAZIONE		,00		,00		
33	ADDIZIONALE COMUNALE ALL'IRPEF DOVUTA		,00		,00		
34	ADDIZIONALE COMUNALE ALL'IRPEF RISULTANTE DALLA CERTIFICAZIONE		,00		,00		
35	ECCEDENZA DELL'ADDIZIONALE COMUNALE ALL'IRPEF RISULTANTE DALLA PRECEDENTE DICHIARAZIONE		,00		,00		
VERSAMENTI O RIMBORSI DA EFFETTUARE A CURA DEL SOSTITUTO D'IMPOSTA		1	DICHIARANTE	2	CONIUGE	3	TOTALE
36	IMPORTO IRPEF DA VERSARE						,00
37	IMPORTO IRPEF DA RIMBORSARE						,00
38	IMPORTO ADDIZIONALE REGIONALE IRPEF DA VERSARE		,00		,00		,00
39	IMPORTO ADDIZIONALE REGIONALE IRPEF DA RIMBORSARE		,00		,00		,00
40	IMPORTO ADDIZIONALE COMUNALE IRPEF DA VERSARE		,00		,00		,00
41	IMPORTO ADDIZIONALE COMUNALE IRPEF DA RIMBORSARE		,00		,00		,00
42	PRIMA RATA DI ACCONTO IRPEF PER IL 2005		,00		,00		,00
43	SECONDA O UNICA RATA DI ACCONTO IRPEF PER IL 2005		,00		,00		,00
44	ACCONTO DEL 20% SUI REDDITI SOGGETTI A TASSAZIONE SEPARATA		,00		,00		,00
45	RICHIESTA DI RATEIZZAZIONE DEI VERSAMENTI DI SALDO E DEGLI EVENTUALI ACCONTI IN RATE						
DATI DA RIPORTARE NELLA SUCCESSIVA DICHIARAZIONE DEI REDDITI							
46	Residuo del credito d'imposta per il riacquisto della prima casa da utilizzare in compensazione		,00		,00		,00
47	Residuo del credito d'imposta per l'incremento dell'occupazione da utilizzare in compensazione		,00		,00		,00
48	Totale spese sanitarie rateizzate nella presente dichiarazione (righe E1, E2 e E3)		,00		,00		,00
49	Crediti per imposte pagate all'estero	DICHIARANTE	Anno	Credito utilizzato	,00	Anno	Credito utilizzato
50		CONIUGE DIC.	Anno	Credito utilizzato	,00	Anno	Credito utilizzato

2.4 - Il modello Unico persone fisiche (Upf)

2.4.1 - L'acquisizione dei modelli Upf

Per l'acquisizione dei modelli Unico persone fisiche, rilasciati anch'essi con struttura a formato compresso, è stato necessario sviluppare un apposito programma di ricostruzione delle informazioni secondo gli stessi canoni utilizzati per i modelli 730. Anche in questo caso si è provveduto a generare di due distinti archivi: uno riportante le informazioni anagrafiche e l'altro contenente i dati fiscali sui redditi dei dichiaranti. Il sistema utilizzato per riconoscere i singoli campi si basa su una struttura di campo virtuale o metacarpo (<codice campo><tipo campo><lunghezza campo><valore campo>) analoga la caso dei modelli fiscali analizzati in precedenza. Anche in questo caso è stato necessario identificare un "separatore di record" che consentisse la corretta attribuzione delle informazioni sui redditi al corrispettivo percettore, senza perdita di informazioni. Al pari degli archivi dei modelli 730, si è potuto verificare empiricamente che i codici fiscali delle persone fisiche rappresentano un separatore affidabile all'interno della struttura dei dati dei modelli Upf predisposta dalla Sogei. La procedura che ha consentito la ristrutturazione dell'archivio dei modelli Upf è stata implementata con il programma in linguaggio VB per il primo anno di indagine (Eu-Silc 2004), mentre per il secondo si è realizzato un programma in Sas simile a quello già implementato per i modelli Cud/770. La procedura in VB è identica alla procedura implementata per i modelli 730. Essa si compone di due fasi, nella prima si eliminano le possibili fonti di errore, costituite da caratteri speciali che in ambiente Sas vengono interpretati come marcatori di fine file, nella seconda si crea un vettore contenente tutte le possibili tipologie (nomi) di variabili presenti in archivio. I record dell'archivio 730_2004 hanno una lunghezza fisica di 1.900 caratteri e contengono: informazioni anagrafiche (record tipo "A"), dettagli relativi ai redditi (record tipo "R"), dettagli contabili relativi ai sostituti (record tipo "S"), dettagli contabili relativi all'Irap (record tipo "Q") e dettagli contabili relativi all'Iva (record tipo "V"). Di queste tipologie di dati solo le prime due sono di interesse ai fini dell'indagine Eu-Silc.

Il programma in Visual Basic testato con successo sui dati delle dichiarazioni fiscali Upf dell'anno 2001, presentava per l'anno dei redditi 2004 forti limitazioni in termini di capacità di memorizzazione delle variabili fiscali da scrivere nel file di output. Infatti, da un primo test, si è osservato che il file di output generato per il complesso delle osservazioni conteneva un numero di variabili fiscali di gran lunga inferiore alle quelle scritte dal medesimo programma in relazione alla prima metà dei record dell'archivio di origine: 702 variabili contro 1.198. Questo stesso risultato ha fornito la traccia per una possibile soluzione al problema della perdita di informazione sulle variabili fiscali scritte dal programma. Infatti, da un'analisi empirica del file di output generato sul primo blocco di record dell'archivio di base si è potuto inferire che le informazioni in esso caricate erano complete⁵. Si è provveduto, quindi, a caricare le informazioni del file di origine in due momenti successivi, lanciando rispettivamente il programma per la prima e la seconda metà dei record in esso contenuti.

Il due file di output generati dalla procedura di VB sono stati successivamente trattati in ambiente Sas con due differenti procedure di input, in quanto il posizionamento delle variabili differiva nei due casi. Il dataset finale, contenente le dichiarazioni di reddito Upf, si è ottenuto con un *merge* dei due file intermedi.

Per quanto attiene il controllo quantitativo del numero delle persone fisiche presenti nel dataset finale, si è potuto accertare l'assenza di perdita di informazioni sulle unità di analisi rispetto all'archivio di partenza. Infatti, il numero di codici fiscali distinti nel dataset finale dei

⁵ In particolare è stata analizzata la congruenza tra le informazioni del file di origine e quelle del file di output, relativamente alle prime 20 persone fisiche a cui erano associate dichiarazioni di redditi disposte su almeno 4 record contigui.

redditi, pari a 14.528, coincideva con il numero di codici fiscali distinti presenti nel file dell'anagrafica Upf.

Se nell'indagine 2004, i limiti incontrati dal programma di caricamento risultavano superabili con opportune strategie, nell'indagine 2005 essi rappresentavano un ostacolo insormontabile. Come soluzione al problema è stata implementata una nuova procedura di caricamento delle variabili direttamente in ambiente Sas, alla stregua di quanto realizzato per i modelli Cud/770 del 2005.

Il programma di lettura in questione effettua, dunque, un caricamento condizionale dei record, cioè sotto il vincolo che sono presenti dati contabili sui redditi (record tipo "R"). Per associare il contenuto di un variabile al relativo nome (parte+colonna) sono stati creati appositamente due vettori, ognuno formato da 56 elementi, corrispondenti ai blocchi possibili scrivibili su ciascun record nella sezione delle colonne: 38-1874. Il primo vettore \mathbf{n} (n_1, n_2, \dots, n_{56}) includeva come elementi i nomi delle variabili, formati dai primi otto caratteri del campo o blocco, mentre il secondo vettore \mathbf{c} (c_1, c_2, \dots, c_{56}) riportava come elementi il contenuto delle precedenti variabili.

La lettura e la creazione dei vettori, da sola, non è sufficiente a creare la struttura della matrice dei dati, formata da variabili in colonna che assumono determinati valori in corrispondenza di ciascuna riga o unità di analisi (codice fiscale). Il passo successivo è quello di assegnare a ciascuna variabile il relativo contenuto. Se prendiamo ad esempio il caso delle variabili del Quadro RN dell'Upf, si evince che per associare il contenuto alla corrispondente variabile è sufficiente creare di una matrice "SA" (33x4) formata da 132 elementi (nomi di variabili)⁶ ed assegnare il valori con tecnica posizionale allo stesso modo della procedura esposta nel paragrafo 2.2.1 (sebbene quest'ultima sia riferita ad una sola dimensione: riga).

Il vantaggio di questa tecnica è quello di consentire il caricamento selettivo delle sole variabili utili allo scopo della ricostruzione del reddito del lavoratore autonomo o di altro percettore, determinando così un risparmio nei tempi di calcolo e una riduzione consistente delle dimensioni del dataset finale. Le variabili di reddito selezionate dall'archivio Upf ammontano a 327 e rappresentato circa un quarto del complesso delle variabili presenti nel tracciato record dell'archivio Upf. Dal controllo numerico dei codici fiscali delle persone fisiche, presenti rispettivamente nell'archivio anagrafico Upf (11.967) e nel dataset finale dei redditi (11.905), si nota che in fase di generazione del file di output vi è una perdita di informazioni inerente a 62 soggetti. Tuttavia, ad un'approfondita analisi delle dichiarazioni afferenti a questi ultimi, si osserva che in nessun caso sono ivi presenti variabili di reddito di interesse per Eu-Silc. Pertanto, l'esclusione di questi soggetti è riconducibile esclusivamente al filtro posto nella selezione delle variabili di studio, a conferma della correttezza delle istruzioni utilizzate nel programma di lettura.

2.4.2 - Le operazioni di controllo, correzione e consolidamento dei dati dell'Upf

Le procedure di controllo e correzione dei dati dell'Upf sono state focalizzate all'individuazione ed eliminazione degli errori di duplicazione presenti nelle rispettive dichiarazioni. Per il 2005 sono state identificate ed eliminate 133 dichiarazioni duplicate in riferimento a 125 persone fisiche, riportanti diversi progressivi della dichiarazione ma identici codici fiscali e dati reddituali. Come criterio per l'identificazione dei duplicati si è utilizzata la frequenza (≥ 2) delle dichiarazioni (record) corrispondente ai diversi codici fiscali. Tra le varie dichiarazioni duplicate è stata considerata valida quella dichiarazione riportante il progressivo più alto (cioè quella relativa all'ultimo invio telematico).

⁶ In realtà le variabili presenti nel Quadro N del modello Upf sono in numero inferiore a 132, poiché una parte delle 33 righe (tipologie di dato fiscale), di cui si compone il quadro, possiede meno di 4 tipologie di informazioni.

2.5 – Il Casellario centrale dei pensionati (CP): acquisizione dei microdati

L'acquisizione dei dati del Casellario centrale dei pensionati dell'Inps è stata realizzata con tecniche sempre più affinate nel corso dei due anni di indagine Eu-Silc 2004-2005. Nel primo anno di indagine è stata sviluppata una procedura che effettuava il *merge* tra i titolari di pensione del CP (presenti al 31 dicembre 2003) e la lista dei potenziali percettori di reddito in Eu-Silc, cioè con gli intervistati di età superiore ai 15 anni. La chiave di aggancio utilizzata era rappresentata dal codice fiscale. Tale procedura implementata in modo sperimentale, con l'intento di testare la possibilità di aggancio con i dati del CP, ha generato risultati molto soddisfacenti in termini di copertura, tali da suggerire l'impiego dei dati amministrativi nella fase di controllo e correzione dei dati campionari sui trasferimenti pensionistici. In questa fase, infatti, l'87,3 per cento dei soggetti che in Eu-Silc 2004 dichiaravano di percepire almeno una pensione (16.072) trovavano un abbinamento con i pensionati del CP, tramite codice fiscale. Inoltre, 1.257 intervistati che in Eu-Silc risultavano titolari di redditi diversi dalle pensione trovavano una corrispondenza nel CP, e dovevano quindi essere imputati come pensionati di fase di *editing*. I dati del CP utilizzati per il *merging* hanno avuto come riferimento temporale: i pensionati rilevati al 31 dicembre del 2003. Ai fini di un controllo sulla qualità dei dati del CP, sono stati successivamente agganciati i pensionati di fonte campionaria che non figuravano nell'archivio Inps 2003, ma che erano presenti nei due anni contigui: 2002 e 2004. Da una prima analisi emerge che i codici fiscali del CP sono sufficientemente affidabili nel tempo, poiché in appena 17 casi, su un totale di 15.288 accoppiamenti, l'abbinamento tra le due fonti si realizza nei due anni contigui e non nell'anno 2003 di riferimento. Dal momento che l'unità di analisi del CP è il trattamento individuale⁷ e non il percettore, ovvero il codice fiscale del titolare, è stato necessario consolidare le informazioni a livello del singolo titolare, seguendo un appropriato schema di classificazione delle prestazioni. Trattandosi di prestazioni pensionistiche si è utilizzato il sistema di classificazione SPCP sviluppato in Istat⁸, coerente con il sistema di classificazione Esspros⁹ adottato in sede europea tanto per il conto della protezione sociale, quanto per la classificazione dei redditi da trasferimento di Eu-Silc (variabili target). Pertanto, l'archivio del CP conteneva in origine 49.133 trattamenti pensionistici individuali riferibili a 15.636 titolari aventi codice fiscale comune a quello dei potenziali titolari di reddito presenti in Eu-Silc 2004. Di questi pensionati, 348 erano in realtà fuori dal campo di osservazione, in quanto percettori di trattamenti aventi decorrenza successiva all'anno 2003 di riferimento dei redditi.

Nell'anno 2005 di indagine, si è sviluppata una nuova procedura di acquisizione dati che, accanto al tradizionale *merge* delle due fonti per mezzo del codice fiscale, applicava la stessa tecnica dell'abbinamento esatto su altre chiavi. In particolare, si è utilizzata come chiave di abbinamento alternativa quella formata dai primi 11 caratteri dello stesso codice fiscale. Inoltre come ulteriore affinamento, si è esteso il procedimento anche ai casi in cui i codici fiscali dei soggetti intervistati in Eu-Silc non figuravano in anagrafe tributaria ma erano tuttavia valorizzati in Sigif.

La procedura prevedeva la seguente sequenza di operazioni:

- 1) individuazione degli abbinamenti esatti tra i codici fiscali (*16 digit*) del CP 2004 e quelli dei soggetti presenti in Eu-Silc 2005 e rilevati nell'anagrafe tributaria;

⁷ Si definisce "trattamento individuale" un trasferimento di risorse monetarie, da una istituzione o da un'impresa ad un determinato beneficiario, identificabile secondo un appropriato sistema di classificazione delle prestazioni e collocabile in un istante temporale definito (cfr. P. Consolini (2002)).

⁸ Per approfondimenti si rimanda a Ceccato F. (1998).

⁹ Si veda Eurostat (1996).

- 2) selezione dei record del CP 2004 che si abbinano secondo la nuova chiave costituita dai primi *11 digit* dello stesso codice fiscale nell'insieme residuale costituito dai record non accoppiati al punto 1);
- 3) individuazione degli abbinamenti esatti tra i codici fiscali (*16 digit*) del CP 2003 e quelli dei soggetti presenti in Eu-Silc 2005 (rilevati nell'anagrafe tributaria) non precedentemente agganciati nei punti 1 e 2;
- 4) selezione dei record del CP 2003 che si abbinano secondo la nuova chiave costituita dai primi *11 digit* dello stesso codice fiscale, non accoppiati ai punti 1, 2, 3;
- 5) individuazione degli abbinamenti esatti tra i primi *11 digit* del codice fiscale del CP 2004 e quelli dei soggetti in Eu-Silc 2005, presenti in Sigif ma non in anagrafe tributaria;

In sintesi, la procedura di acquisizione delle informazioni del Casellario centrale dei pensionati utili all'indagine Eu-Silc 2005 produceva i seguenti risultati.

Tavola 2.2 – Struttura del file Cud/770 utilizzato nell'indagine Eu-Silc 2005

Fonte amministrativa	Fonte campionaria	Chiave di aggancio	Fonte Derivata	Totale abbinamenti
CP_2004	Eu-Silc_2005	Codice fiscale (16 digit) ANAGR. TRIB.	{CP ₀₄ ^{at} ^EUS ₀₅ }CF _{at}	17.389
CP_2004 ^(a)	Eu-Silc_2005	Codice fiscale 11 digit ANAGR. TRIB.	{CP ₀₄ ^(a) ^EUS ₀₅ }CF11 _{at}	228
CP_2003 ^(b)	Eu-Silc_2005	Codice fiscale ANAGR. TRIB.	{CP ₀₃ ^(b) ^EUS ₀₅ }CF _{at}	98
CP_2003 ^(c)	Eu-Silc_2005	Codice fiscale 11 digit ANAGR. TRIB.	{CP ₀₃ ^(c) ^EUS ₀₅ }CF11 _{at}	7
CP_2004	Eu-Silc_2005	Codice fiscale 11 digit Sigif	{CP ₀₄ ^s ^EUS ₀₅ }CF11 _s	210
CP_2004/2003	Eu-Silc_2005	K={CF_{at}, CF11_{at}, CF11_s}	{CP_{03/04}^{at}^EUS₀₅}K	17.932

(a) per l'insieme dei codici del Casellario centrale dei pensionati che non trova alcun abbinamento al passo precedente.

(b) per l'insieme dei codici del Casellario centrale dei pensionati che non trova alcun abbinamento nei due passi precedenti.

(c) per l'insieme dei codici del Casellario centrale dei pensionati che non trova alcun abbinamento nei tre passi precedenti

L'utilizzo della chiave costituita dai primi 11 caratteri del codice fiscale, non essendo univoca nell'universo di riferimento dell'anagrafe tributaria, accresce la probabilità di falsi abbinamenti; cioè di agganciare erroneamente le informazioni di individui che in realtà sono diversi. Per limitare tale inconveniente è stato imposto un vincolo in fase di aggancio dei record, consistente nel fatto che le unità dovessero avere in comune anche la provincia di residenza. Tale restrizione ha portato all'esclusione di 158 abbinamenti e quindi al caricamento delle informazioni relative a 17.774 percettori di pensione da CP. Di questi, 14.548 rappresentano le unità di analisi di interesse per l'indagine Eu-Silc 2005, cioè corrispondono a soggetti intervistati con almeno 15 anni di età. Il numero delle persone che dichiarano almeno una pensione in Eu-Silc e che non trovano una corrispondenza con il CP, è pari a 718. Si può pertanto affermare che i record abbinati con il CP coprono una quota del 95,3 per cento del totale dei pensionati presenti in Eu-Silc 2005.

Capitolo 3 – La costruzione delle variabili di reddito Eu-Silc a partire dalle variabili fiscali: armonizzazione delle definizioni

3.1 - Introduzione

Tra le variabili obiettivo di Eu-Silc sono particolarmente rilevanti: il reddito netto monetario da lavoro dipendente (*PY010N*), il reddito netto da lavoro autonomo (*PY050N*), le “prestazioni sociali nette in denaro” associate alle funzioni vecchiaia (*PY100N*), superstiti (*PY110N*) e invalidità (*PY130N*). Con riferimento a dati dell’indagine sulle condizioni di vita, anno 2003, queste tre componenti di reddito comprendono il 94,7 per cento del reddito totale individuale. La prima variabile *target*¹ è definita come l’insieme delle “componenti monetarie, al netto delle trattenute fiscali e previdenziali, erogate dai datori di lavoro a remunerazione delle prestazioni lavorative dei propri dipendenti.

La seconda variabile *target* è definita come l’ammontare dei guadagni conseguiti, durante il periodo di riferimento, dal soggetto (per sé o per conto dei propri familiari) in relazione all’attività di lavoratore autonomo. La remunerazione del lavoro autonomo è generalmente legata ai profitti derivanti dell’attività produzione di beni e servizi (o agli autoconsumi quando sono ritenuti parte dei profitti).

Le ultime tre sono definite come l’insieme delle prestazioni in denaro di protezione sociale rispettivamente contro i rischi legati alla funzione vecchiaia (*PY100N*), alla funzione superstiti (*PY110N*) e alla funzione invalidità (*PY130N*). Queste prestazioni in denaro possono essere periodiche e continuative (pensioni in senso stretto) oppure una tantum (liquidazioni).

Gli elementi che concorrono alla determinazione del reddito monetario da lavoro dipendente, definito da Eurostat, sono rappresentati nel contesto italiano da: paga base o minimo tabellare, provvigioni, aumenti periodici per anzianità (scatti di anzianità), superminimi (per merito del lavoratore o per onerosità del mansione svolta), premi aziendali (produzione, rendimento e fedeltà), compensi per lavoro straordinario, indennità compensative (connesse al rischio della mansione o alla funzione, alla reperibilità, alla sede disagiata, ecc.), gratifiche natalizie o mensilità aggiuntive (tredicesima, quattordicesima), patto di non concorrenza, mance, contributi spese per la frequenza asili nido dei figli del lavoratore. Tutti questi elementi, con la sola eccezione dei contributi per le rette dell’asilo², costituiscono base imponibile ai fini fiscali e sono pertanto rilevati nel modello Cud/770. In particolare, questi redditi figurano al punto 1 (redditi per i quali è possibile fruire della deduzione di cui art. 10-bis e art. 13 del Tuir) e al punto 2 (redditi per i quali è possibile fruire della deduzione di cui art. 10-bis, comma 1) della Parte B – Sezione: “*Dati fiscali*”, segnatamente al lordo delle trattenute fiscali e al netto delle ritenute contributive. Per ottenere i redditi netti da lavoro dipendente (*PY010N*) occorre che dai redditi imponibili (punti 1 e 2) siano defalcati le ritenute Irpef (punto 12 del Cud_2004), l’addizionale regionale (punto 13 del Cud_2004) e l’addizionale comunale (punto 14 del Cud_2004).

Per quanto riguarda le prestazioni sociali in denaro aventi carattere periodico e continuativo (pensioni), vi è una parte esente da imposta (pensioni sociali, pensioni di guerra, pensioni di invalidità, indennità di accompagnamento, rendite per infortunio permanente) e

¹ Secondo la definizione adottata dall’Eurostat (cfr. Eurostat 2004).

² In base all’art. 48 c. 2, lett. f, del Tuir i contributi del datore lavoro a favore della generalità dei dipendenti o a relative categorie per frequenza di asili nido e di colonie climatiche, per borse di studio dei loro familiari non costituiscono base imponibile ai fini fiscali e contributivi.

quindi non coperta dall'archivio dei modelli Cud/770. Mentre le prestazioni sociali erogate una tantum (liquidazioni) sono rilevabili in un'altra Sezione: *Trattamenti di fine rapporto e indennità equipollenti*. I dati sui trattamenti pensionistici al lordo (tassabili) sono riportati negli stessi punti 1 e 2, Parte B – Sezione: “*Dati fiscali*”, occupati dai redditi di lavoro dipendente e assimilati. Le ritenute fiscali applicate ai titolari di pensione sono indicate, come per i lavoratori dipendenti, ai punti 12-13-14 della medesima Sezione. Sebbene il Cud/770 non preveda campi distinti per i redditi da lavoro e quelli da pensione, è ugualmente possibile tracciare alcune linee guida per la loro separazione. Così ad esempio, al punto 9 della Parte A – Sezione: “*Dati relativi al dipendente, pensionato o altro percettore*” è specificata la modalità che contraddistingue il pensionato (lettera A o B) dal resto. Inoltre, ai punti 6 e 7 della Parte B – Sezione: “*Dati fiscali*” è possibile rilevare con precisione il numero dei giorni di pensione, in corrispondenza dei quali spettano le deduzioni di cui all'art. 10-bis, commi 2 e 3, del Tuir. L'identificazione del sostituto d'imposta erogatore di pensioni (tramite la ragione sociale o il codice fiscale) fornisce un altro criterio per individuare il titolare di reddito da pensione.

Il computo del “reddito netto da lavoro o da pensione” dal Cud/770, come differenza tra l'imponibile lordo e le ritenute fiscali applicate nell'anno (imposta provvisoria) dal sostituto d'imposta, non tiene conto di eventuali aggiustamenti in positivo in sede di dichiarazione dei redditi (modello 730) e pertanto potrebbe fornire un valore in difetto. Se il contribuente non possiede altri redditi nell'anno e fa valere il diritto ad eventuali deduzioni o detrazioni di imposta (per spese sanitarie, per interessi passivi legati a mutui ipotecari per acquisto dell'abitazione principale, premi per l'assicurazione vita, ecc.), in sede di conguaglio fiscale riceverà un rimborso fiscale. L'imposta definitiva in questo caso sarà inferiore a quella provvisoria (ritenuta) applicata dal datore di lavoro.

In realtà, l'aggiunta di eventuali altri redditi non tassati anticipatamente (ad esempio, i redditi da fabbricati o i redditi da terreni) o tassati su un'aliquota convenzionale (compensi di attività occasionale di lavoro autonomo) determina un aumento della base imponibile e produce un saldo d'imposta a debito (differenza tra l'imposta dovuta e le trattenute o le ritenute d'acconto), rendendo più difficile il calcolo del reddito netto di ciascuna componente di reddito. In questi casi per ottenere il reddito netto di ogni tipologia, si procede a calcolare il contributo di ciascuna componente rispetto al reddito complessivo e ad applicare pro-quota l'imposta definitiva (al netto delle detrazioni e deduzioni comuni a tutti i redditi). Quest'ultima operazione dovrà tener conto delle detrazioni specifiche per tipologia di reddito (p.es. la detrazione per lavoro dipendente) che si aggiungono alle detrazioni comuni, cioè indipendenti dalla fonte di reddito (p.es. la detrazione per spese sanitarie). Se il contribuente è registrato nell'archivio dei modelli Cud/770 ma non in quello dei modelli 730 (o modello Unico persone fisiche), si può ragionevolmente ipotizzare che l'imposta provvisoria, applicata dal datore di lavoro nel Cud/770, coincida con quella definitiva e che il reddito netto desumibile da quest'ultima fonte sia corretto. Si ricorda, inoltre, che nel Cud/770 è presente una sezione sui redditi derivanti dalle collaborazioni coordinate e continuative. Sebbene assimilati fiscalmente ai redditi da lavoro dipendente questi compensi sono classificati in Eu-Silc come redditi da lavoro autonomo.

Le componenti di reddito da lavoro e da pensione, sinora esaminate, sono rilevabili anche presso il modello 730, nel caso in cui percettore intenda assolvere all'obbligo di presentazione e di pagamento delle imposte, per i redditi che si aggiungono a quelli dichiarati nel Cud, ovvero ottenere in tempi rapidi rimborsi d'imposta spettanti (deduzioni e detrazioni). I redditi imponibili da lavoro dipendente o da pensione, di cui al punto 1 della Parte B – Sezione: “*Dati fiscali* del Cud, sono ricompresi nel Quadro C – Sezione I: “*Redditi da lavoro dipendente e assimilati*” del modello 730 e distinti in tre categorie: 1) redditi da pensione; 2) redditi da lavoro e assimilati; 3) compensi ai lavoratori socialmente utili che hanno raggiunto l'età pensionabile. I redditi

imponibili da lavoro dipendente parzialmente deducibili, di cui al punto 2 della Parte B – Sezione: “*Dati fiscali del Cud, sono indicati nel Quadro C – Sezione I “Redditi da lavoro dipendente e assimilati per i quali non spettano le deduzioni”*”. Il complesso delle ritenute applicate dal sostituto d’imposta per questi redditi è riportato nella Sezione III: “*Ritenute sui redditi da lavoro dipendente e addizionali regionali e comunali Irpef*”. Dal modello 730 si può ottenere una stima dei redditi netti da lavoro dipendente, defalcando le ritenute operate dal sostituto (imposta provvisoria) ai redditi imponibili da lavoro o da pensione, come per il Cud. Se il contribuente è presente contestualmente negli archivi dei modelli Cud e dei modelli 730, allora è necessario verificare la coerenza delle informazioni ivi riportate e procedere al calcolo esatto del reddito netto. Le informazioni raccolte nel prospetto di liquidazione del modello 730 forniscono tutti gli elementi per il computo di questo aggregato. Il modello 730 consente, tra l’altro, di rilevare altre variabili reddito obbiettivo, quali i redditi da capitale reale (affitto di terreni e/o fabbricati) e talune componenti del reddito da lavoro autonomo (compensi derivanti dall’esercizio di attività occasionali da lavoro autonomo).

Il modello Unico persone fisiche viene regolarmente utilizzato da percettori di compensi da lavoro autonomo, riferiti ad attività professionale, in sede di dichiarazione dei redditi. Esso comprende, tuttavia, un apposito Quadro – RC – sui redditi da lavoro dipendente e assimilati, per includere le dichiarazioni dei lavoratori autonomi che percepiscono contestualmente redditi da lavoro dipendente e assimilati. Il modello Upf, in talune circostanze³, può essere utilizzato dai percettori di soli redditi da lavoro dipendente e assimilati. Nel quadro RC dell’Upf si riporta, come da modello 730, la distinzione per tipologia di reddito: 1) reddito da lavoro dipendente; 2) reddito da pensione; 3) compensi ai lavoratori socialmente utili che hanno raggiunto l’età pensionabile. Il modello Upf, al pari del modello 730, classifica le tipologie di redditi (imponibile) da lavoro dipendente e assimilati sulla base alla possibilità di applicare o meno le specifiche deduzioni. Le ritenute applicate dai sostituti d’imposta sui redditi da lavoro dipendente sono indicate nel complesso, cosicché non vi è possibilità di associarle alle fonti di reddito. Il modello Upf permette inoltre di rilevare una serie di componenti di reddito da capitale reale (redditi da fabbricati e da terreni), alla stregua del modello 730. Tale modello costituisce il punto di riferimento per la validazione dei dati campionari sulle componenti di redditi da lavoro autonomo. Il reddito che deriva da lavoro autonomo è definito come quello di un’impresa: ricavi meno costi. Esso è quindi costituito dalla differenza tra i compensi percepiti e le spese sostenute per l’esercizio del proprio lavoro durante il periodo d’imposta. Gran parte di questi redditi sono anch’essi soggetti al sistema della ritenuta alla fonte a titolo di acconto. Seguendo la struttura del modello Upf i redditi da lavoro autonomo si compongono di: redditi derivanti dall’esercizio delle arti e professioni (Quadro RE); redditi dei titolari d’impresa in contabilità ordinaria (Quadro RF) e semplificata (Quadro RG); redditi di partecipazione in società di persone e assimilate (Quadro RH); redditi di allevamenti (Quadro RD), compensi derivanti da attività di lavoro autonomo, anche se svolte all’estero, non esercitate abitualmente (Quadro RL). La somma di questi redditi al netto delle ritenute d’acconto versate fornisce la stima del reddito netto da lavoro autonomo.

³ Nella fattispecie trattasi di: 1) lavoratori dipendenti che hanno cambiato datore di lavoro e sono in possesso di più certificazioni di lavoro dipendente o assimilati (Cud), dove l’imposta corrispondente al reddito complessivo supera il totale delle detrazioni spettanti (ovvero se non è richiesto il conguaglio fiscale all’ultimo datore di lavoro); 2) lavoratori dipendenti che hanno percepito direttamente dall’Inps o da altri enti indennità e somme a titolo di integrazione salariale o ad altro titolo, se erroneamente non sono state effettuate le ritenute e se non ricorrono le condizioni di esonero; 3) lavoratori dipendenti che hanno percepito retribuzioni e/o redditi da privati non obbligati per legge ad effettuare ritenute d’acconto (per esempio collaboratori familiari, autisti ed altri addetti della casa).

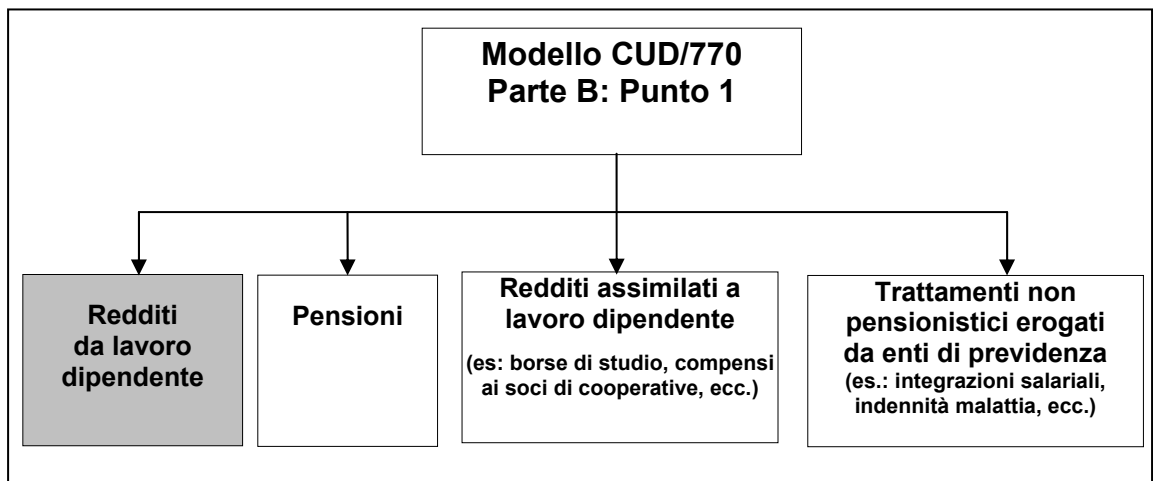
3.2 - La costruzione del reddito da lavoro dipendente

3.2.1 - Il reddito netto annuo da lavoro dipendente di fonte Cud/770

Le informazioni contenute nel modello Cud consentono la ricostruzione del reddito netto da lavoro dipendente, definito come “retribuzione lorda al netto dei contributi sociali obbligatori e delle ritenute fiscali alla fonte applicate dal datore di lavoro”. In realtà questa definizione se da un lato è coerente con la definizione Eu-Silc di reddito netto⁴, dall’altro rappresenta un’ approssimazione del concetto di “reddito disponibile” da lavoro dipendente. Infatti, se si escludono altre componenti aggiuntive di reddito a tassazione ordinaria, il calcolo provvisorio del reddito netto da lavoro dipendente da modello Cud coincide con il calcolo finale, cioè quello applicato in sede di dichiarazione finale dei redditi⁵, solo nel caso particolare in cui il contribuente non abbia particolari benefici da fare valere nel calcolo dell’imposta finale. Questi benefici consistono, generalmente, in deduzioni dell’imponibile o detrazioni dall’imposta come sgravio degli oneri sostenuti dalle famiglie, quali ad esempio le spese sanitarie, gli interessi passivi su mutui ipotecari e l’assegno periodo corrisposto al coniuge separato/divorziato. In Eu-Silc si stabilisce che i benefici fiscali sopra menzionati vadano riportati nella variabile target “HY0145N” che definisce il saldo fiscale a livello familiare.

Nella parte B del modello Cud, al punto 1 vengono rilevati i “redditi per i quali è possibile fruire della deduzione di cui all’art. 11 e delle detrazioni di cui all’art. 14 del Tuir”. Essi comprendono stipendi, pensioni, mensilità aggiuntive, indennità di trasferta, borse di studio, compensi derivanti da rapporti di collaborazione coordinata e continuativa, compensi ai soci di cooperative di produzione e lavoro, ecc. Si tratta, in sostanza, di componenti di reddito eterogenee e classificabili secondo lo schema riportato nella seguente figura 3.1.

Figura 3.1 – Classificazione delle componenti di reddito di cui al punto 1, Parte B, modelli Cud/770



Ai fini, dunque, di una corretta attribuzione degli importi individuali di reddito da lavoro dipendente è stato necessario scorporare quest’ultima componente di reddito, evidenziata in grassetto, dalle restanti tipologie di reddito. Per realizzare la separazione tra le fonti di reddito si

⁴ La variabile target Eurostat, definita come “reddito monetario netto da lavoro dipendente (PY010N)”, si riferisce alle componenti in denaro versate dal datore di lavoro ai propri dipendenti, come contropartita per la loro prestazione lavorativa, al netto delle ritenute alla fonte dovute per il prelievo fiscale e per i contributi sociali.

⁵ Generalmente avviene nel maggio/giugno dell’anno successivo a quello in cui si sono percepiti.

è proceduto, dapprima, a scorporare il reddito da lavoro dipendente dai trasferimenti pensionistici. Come linea guida sono state utilizzate le variabili, di cui al punto 6 e 7 della Parte B del Cud, indicanti il numero di giorni per i quali spettano le deduzioni, di cui all'art. 11 del Tuir punto, rispettivamente per il lavoro dipendente e la pensione⁶. Quando tali informazioni risultano mancanti si è ricorso all'analisi dei dati retributivi esposti nella Parte C (Sezioni: *Dati previdenziali ed assistenziali Inps e Inpdap-Ipost*), relativi ai lavoratori dipendenti del settore privato o pubblico. In presenza di elementi retributivi (specificati ai punti: 9, 10, 29, 58, 85, 86, 90, 91 - parte C) si è ipotizzato che i redditi esposti al punto 1 della parte B fossero ascrivibili alla tipologia di reddito da lavoro dipendente. Infine, come ultimo criterio di assegnazione, in alternativa ai precedenti, si è considerata la tipologia del sostituto d'imposta (identificato in base al codice fiscale/ partita Iva): 1) ente di previdenza; 2) datore di lavoro privato o pubblico (diverso dall'Ente di previdenza). I sostituti d'imposta che erogano pensioni sono tipicamente enti di previdenza che, tuttavia, possono essere a loro volta datori di lavoro (cioè erogatori di reddito da lavoro dipendente) nei confronti del personale alle proprie dipendenze. Come regola, si è stabilito di classificare come redditi da lavoro dipendente quei compensi erogati dai datori di lavoro privati o pubblici, diversi dagli enti di previdenza. Al termine della procedura di assegnazione realizzata per l'indagine Eu-Silc 2005 si osserva che appena 42 soggetti, su un totale di 35.544 titolari di reddito, possiedono una forma di reddito non assimilabile a nessuna delle due tipologie di reddito: lavoro dipendente/pensione. Nell'indagine 2005 è stata impiegata, seppur limitatamente alla fase di selezione dei redditi dei soci di cooperative e delle borse di studio, la variabile di cui punto 9 della Parte A – *Dati relativi al dipendente, pensionato o altro percettore delle somme* – che definisce la “categoria di percettore di reddito”. Tale variabile si rivela di grande ausilio per l'attribuzione delle componenti di reddito dopo la prima fase di assegnazione. Sulla base dei valori da essa assunti, si scopre così che i 42 casi di cui sopra sono, in realtà, rappresentati da eredi del lavoratore dante causa. Per l'indagine 2006 si prevede un utilizzo più esteso ed integrato di questa variabile ai fini della classificazione delle poste di reddito.

Una volta fissato il procedimento di estrazione della componente di reddito da lavoro dipendente, il passo successivo consiste nella ricostruzione della variabile target Eurostat *PY010N* definita come reddito lordo da lavoro dipendente al netto del prelievo fiscale e contributivo. Il criterio adottato considera la differenza tra il reddito al punto 1 (retribuzione lorda al netto dei contributi sociali) e le ritenute Irpef (punto 12), le addizionali regionali (punto 13) e le addizionali comunali (punto 14). Di seguito si illustra la formula adottata nel calcolo del reddito netto da lavoro dipendente sulla base degli elementi rilevati nel modello Cud/770.

$$[1] \quad NY^E = \underbrace{GY^E - SC^E}_{\substack{\text{Punto 1} \\ \text{Parte B} \\ \text{CUD}}} - T^E \quad \leftarrow \substack{\text{Punti 12-13-14} \\ \text{Parte B CUD}}$$

dove:

NY^E : reddito netto da lavoro dipendente (al netto dei contributi sociali e delle ritenute fiscali alla fonte);

GY^E : retribuzione lorda (esclude i contributi sociali a carico dei datori di lavoro);

SC^E : contributi sociali (obbligatorie) a carico dei lavoratori dipendenti;

T^E : ritenute fiscali alla fonte applicate al reddito da lavoro dipendente (Irpef e addizionali regionali e comunali).

⁶ In nessuna dichiarazione del modello Cud/770 vi è la compresenza di valori sui punti 6 e 7, a conferma del fatto che i modelli Cud riportano al più una tipologia di reddito.

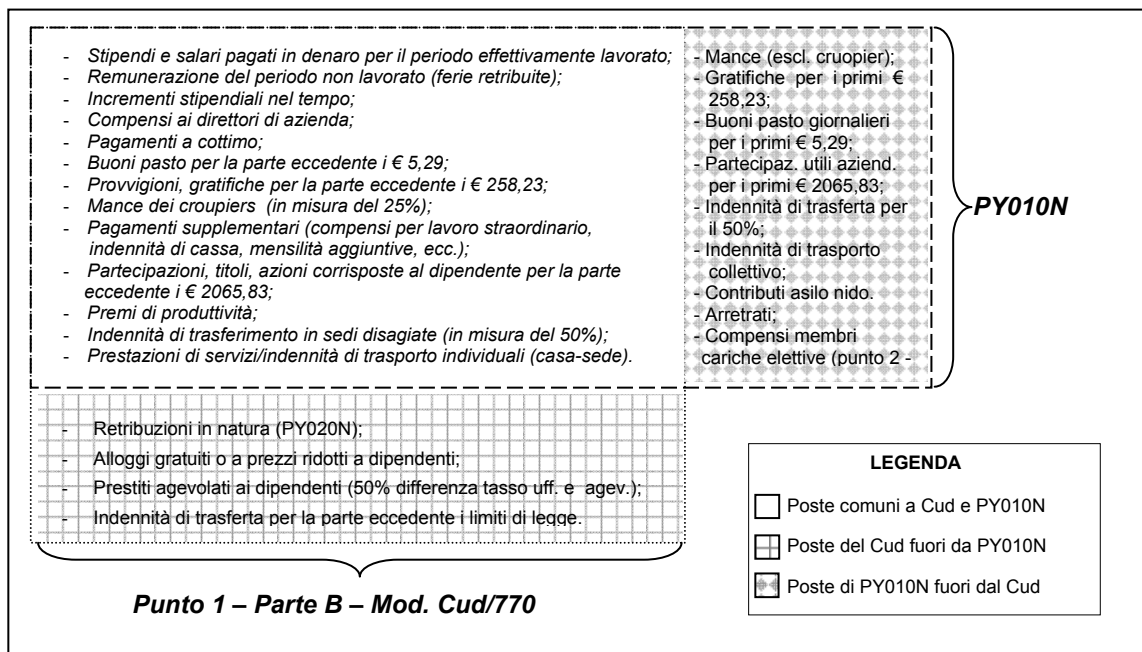
Il calcolo del reddito netto risulta più complesso in compresenza di redditi al punto 1 e al punto 2 della parte B, in quanto le ritenute fiscali, esposte nei successivi punti (12, 13, 14), sono comuni alle due tipologie di reddito⁷. Al punto 2 sono, infatti, riportate alcune componenti non tutte collocabili nella sfera dei redditi da lavoro dipendente, ma pertinenti anche alla tipologia del lavoro autonomo. Tra queste ultime vi sono i compensi per l'attività libero professionale intramuraria del personale dipendente del Servizio sanitario nazionale ed i compensi corrisposti ai membri delle commissioni tributarie. A causa delle difficoltà incontrate nella ripartizione delle componenti di reddito da lavoro dipendente e da lavoro autonomo comprese al punto 2, si è stabilito di tralasciare questa voce dal calcolo finale dei redditi di fonte amministrativa per l'anno di indagine 2005. Tuttavia, per l'edizione Eu-Silc 2006, si prevede l'uso della variabile indicatrice della categoria di percettore (punto 9 della parte A) al fine di recuperare quegli elementi retributivi specificati al punto 2 (ad es. indennità percepite dai membri del Parlamento nazionale e del Parlamento europeo e indennità, comunque denominate, percepite per le cariche elettive e per l'esercizio di pubbliche funzioni) che concorrono al determinare il reddito da lavoro dipendente.

In base alle normative fiscali (D.p.r. n. 917/86 e succ. disposizioni) gli elementi costitutivi del reddito da lavoro dipendente (base imponibile), esposti al punto 1 della parte B, includono tanto gli stipendi, emolumenti, premi di produttività, mensilità aggiuntive quanto una serie benefici economici concessi al lavoratore. Tra questi benefici che determinano la base imponibile si possono annoverare le indennità per ferie non godute, le indennità di trasferta o diaria fuori dal comune sede di lavoro per la parte eccedente i limiti di legge (ad esempio per la parte eccedente 46,48 euro delle trasferte in Italia e 77,47 euro all'estero in caso di rimborso delle sole spese di viaggio), le indennità di prima sistemazione per una quota del 50 per cento, ed una lista di *fringe benefits*. Questi ultimi comprendono, al loro volta, le erogazioni liberali per parte eccedente 258,23 euro; le prestazioni sostitutive della mensa quali buoni pasto per parte eccedente i 5,29 euro); prestiti agevolati (costituisce base imponibile il 50 per cento della differenza tra il tasso di interesse applicato dal datore e il tasso ufficiale di riferimento); auto aziendali (il 30 per cento della tariffa Aci corrispondente ad una percorrenza convenzionale di 15 mila chilometri); indennità di trasporto alla singola persona (quelle collettive non concorrono a determinare la base imponibile).

Non tutti i benefici economici, appena citati, rientrano nella definizione Eurostat di "reddito in denaro da lavoro dipendente in denaro" relativa alla variabile target *PY010N*. Alcuni di essi sono ricompresi nella complementare definizione Eurostat di "retribuzione in natura *PY020N*"; come ad esempio l'auto aziendale, i buoni pasto, le erogazioni liberali, che tuttavia ai fini fiscali sono calcolate in modo convenzionale ed in misura inferiore all'effettivo valore. Altri ancora, non fanno parte della componente del reddito da lavoro dipendente (in denaro o in natura), quali le indennità di trasferta (costi di produzione) e gli alloggi gratuiti o a prezzo inferiore di mercato offerti dal datore di lavoro (affitti imputati *HY030G*). La seguente figura 3.2, sintetizza in modo esaustivo l'insieme delle differenze o corrispondenze tra le poste rilevate al punto 1 – parte B – del modello Cud/770 e le voci che compongono la variabile target *PY010N*.

⁷ Il problema incontrato si risolve con la ripartizione pro-quota delle ritenute fiscali totali rispetto alle relative componenti retributive lorde: $T^E = T^{E+O} \cdot (G^E / G^{E+O})$; dove T^{E+O} e G^{E+O} indicano rispettivamente l'ammontare complessivo delle ritenute fiscali e delle retribuzioni lorde per la somma della componente da lavoro dipendente (E) e della restante componente non attribuibile (O).

Figura 3.2 – Corrispondenze/differenze tra le poste economiche incluse nella variabile target PY010N e nella variabile fiscale di cui al punto 1 della parte B del mod. Cud/770



3.2.2 - Le mensilità retribuite nell'anno dei lavoratori dipendenti da fonte Cud/770

Il calcolo delle mensilità retribuite è stato realizzato separatamente per i lavoratori del privato e del settore pubblico, in quanto le variabili indicatrici sono diversamente articolate. Le variabili utilizzate per la stima delle mensilità retribuite a favore dei lavoratori del privato comprendono: i giorni retribuiti (punto 13), le settimane retribuite (punto 11), i mesi retribuiti da calendario (punti 14 e 15), le settimane "Sett. I"⁸ relative alla malattia/infortuni (punto 38), maternità (punto 40) e della CIG (punto 46); esposti nelle Sezioni 1-3 *Dati Previdenziali e Assistenziali Inps - Parte C -* del modello Cud/770 (figura 3.3).

⁸ Si tratta delle settimane intere di calendario (da domenica a sabato) per le quali il lavoratore non ha percepito alcuna retribuzione dal datore di lavoro in funzione di malattia/infortunio, maternità e congedi parentali, Cig.

Figura 3.3 – Sezioni 1, 2 e 3: Dati previdenziali ed assistenziali Inps - Parte C, Cud/770_2005

Codice fiscale del percipiente _____ Progressivo Comunicazione _____

PARTE C
DATI PREVIDENZIALI ED ASSISTENZIALI INPS

1 Qualifica 2 3 Matricola azienda 4 Prov. Lax. 5 NS 6 DS 7 ALTRE 8 Competenze correnti 9 Altre competenze 10

Settimane retrib. 11 Settimane utili 12 Giorni retrib. 13

MESI RETRIBUITI NELL'ANNO
tutti con esclusione di
14 T 15 G F M A M G L A S O N D

CONTRATTO
Codice 16 Tipo 17 Livello inquadramento 18

SEZIONE 1

Data cessazione 19 giorno mese Tipo rapporto 20 Trasn. rapporto 21 Accantonamento T.F.R. spettante 22

COORD. ASSEGNI FAMILIARI
Tabella 23 Num. Fam. 24 Classe 25

SEZIONE 2
RETRIBUZIONI PARTICOLARI

	1) 26	27 giorno mese anno	28 giorno mese anno	29	30	31	32	33	34
2) 26	27 giorno mese anno	28 giorno mese anno	29	30	31	32	33	34	
3) 26	27 giorno mese anno	28 giorno mese anno	29	30	31	32	33	34	
4) 26	27 giorno mese anno	28 giorno mese anno	29	30	31	32	33	34	

SEZIONE 3
CONTRIBUZIONE FIGURATIVA

	35	36	37	38	39	40	41	42	43	
Tot. annuo settiman. retribuita	Retribuzioni ridotte	Differenza retribuzioni da accreditare per CIG	MALATTIA E/O INFORTUNI Sett. 1	Sett. 2	MAT e CONG. D. Lgs. 151/01 Accreditato retrib. off.	Sett. 1	Sett. 2	CONG. D. Lgs. 151/01 Accreditato retrib. convenzionale	Sett. 1	Sett. 2
L88/87 Sett. 1	Sett. 2	CIG Sett. 1	Sett. 2	Congedi art. 42 c. 5 D. Lgs. 151/01 Sett. 1	Sett. 2	Donat. sangue L107/90				
44	45	46	47	48	49	50				

Il criterio impiegato per il calcolo delle mensilità dei lavoratori privati è strutturato in diverse fasi. Nel primo passo si è proceduto alla scomposizione della variabile “Mesi retribuiti nell’anno” (punto 15) nelle dodici variabili indicatrici mensili (0=mese retribuito, 1=mese non retribuito) e al relativo consolidamento in presenza di dichiarazioni ripetute.

Figura 3.4 – Consolidamento (SOMMA) della variabile “Mesi retribuiti nell’anno” in presenza di dichiarazioni ripetute

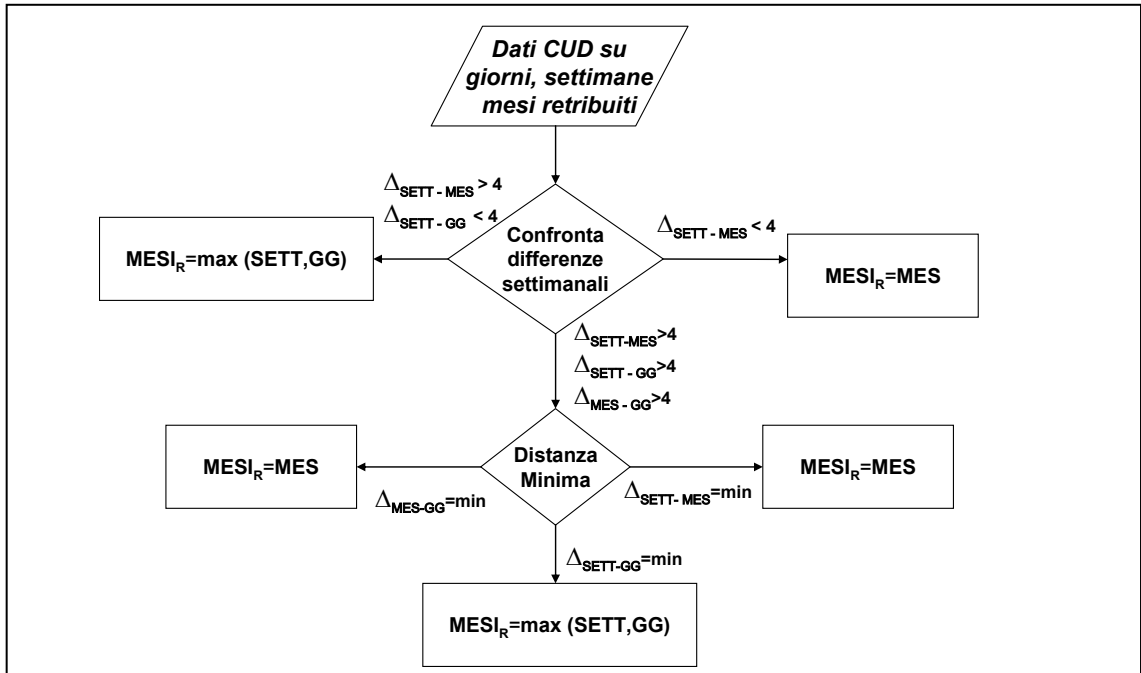
Record	Codice fiscale	MESI RETRIBUITI NELL'ANNO											
		mese 1	mese 2	mese 3	mese 4	mese 5	mese 6	mese...					
rec 1	XXX111AAA	0	1	1	1	1	1						
rec 2	XXX111AAA	0	1	1	0	0	1						
rec 3	XXX111AAA	1	0	1	0	1	1						
SOMMA		0	0	1	0	0	1						

Casi sovrapposti
Casi non sovrapposti

Successivamente, si è calcolata la somma dei mesi riportanti valore “0” (mese retribuito) lungo il vettore di 12 elementi costituiti dalle mensilità consolidate al passo precedente. Tale somma ha fornito il valore iniziale della variabile derivata: “stima del numero dei mesi retribuiti”. Poiché da un’analisi empirica è emerso che la variabile “Mesi retribuiti nell’anno”, al punto 15, conteneva un numero di mensilità pagate inferiori alla somma dei mesi ricavati dall’insieme delle settimane retribuite (11) e delle settimane “Sett. I” interamente non retribuite legate eventi di malattia, maternità e Cig, si è stabilito di utilizzare in questi casi le due variabili

settimanali per l'*inizializzazione* della stima delle mensilità⁹. Il valore finale della stima si ricava da un confronto incrociato su base settimanale dei valori rilevati rispettivamente nella variabile giorni retribuiti, al punto 13, trasformati in settimane (*GG_sett*), nella variabile di sintesi delle settimane retribuite e non retribuite (*SETT*) per gli eventi di cui sopra (punti 11 e 38, 40, 46) e nella variabile mesi retribuiti (*MES_sett*) nell'anno, di cui al punto 15, trasformata in settimane (*MES_sett*).

Figura 3.5 – Albero delle decisioni relativo alla stima finale dei mesi retribuiti ($MESI_R$)



La procedura di calcolo indica che nel 82,3 per cento dei casi non vi sono differenze rilevanti¹⁰ fra le informazioni riportate nelle tre variabili indicatrici per le settimane retribuite, per cui si assume come valido il valore fissato al passo precedentemente (casistica 1 in tavola 1). Per un restante 9,8 per cento si presentano differenze contenute tra le variabili associate ai mesi retribuiti (*MES_sett*) e alle settimane retribuite e non retribuite (*SETT*) e, allo stesso tempo, differenze rilevanti tra la variabile *SETT* e la variabile dei giorni tradotti in settimane *GG_sett*. Anche in questa circostanza si stabilisce che il valore inizializzato in precedenza fornisca la stima finale. In una percentuale ridotta di casi (4,5 per cento) la divergenza è da attribuire alle variabili mesi retribuiti (*MES_sett*) e settimane retribuite e non (*SETT*), mentre non vi sono differenze di rilievo tra quest'ultima i giorni trasformati in base settimanale (*GG_sett*). La stima finale dei mesi retribuiti si ricava dal valore massimo, tradotto in mensilità, tra la variabile *SETT* e giorni *GG_sett*.

Se tutte le differenze assolute, calcolate per ogni coppia di variabili indicatrici, superano la soglia delle cinque settimane si procede ad identificare la coppia con minore distanza assoluta. Come stima finale e si sceglie il valore massimo tra queste due variabili.

⁹ Si tratta di 394 dichiarazioni errate su un totale di 15.086 dichiarazioni che presentavano valori positivi nelle competenze correnti.

¹⁰ Col termine "non rilevante" si intende che la differenza in valore assoluto tra due variabili indicatrici è contenuta nell'intervallo 0-4 settimane.

Si rileva, infine, come nell'1,3 per cento dei casi la procedura non assegna alcun valore positivo. Da un'ispezione sui dati si evince che la mancata assegnazione si associa a valori di basso importo e a trattamenti di fine rapporto (licenziamento o quiescenza). Si può pertanto ipotizzare realisticamente che si tratta di redditi erogati per rapporti di lavoro pregressi e non più in essere.

Tavola 3.1 – Esiti del confronto incrociato tra le variabili indicatrici delle settimane retribuite. Valori assoluti e percentuali – Indagine Eu-Silc 2005

ESITI DEL CONFRONTO INCROCIATO TRA LE TRE VARIABILI INDICATRICI:	Casistica	Azione	Frequenza assolute	Percentuali
Mesi non assegnati	(0)	$MESl_R = 0$	196	1,30
Differenze poco rilevanti sulle tre variabili indicatrici SETT,GG,MES	(1)	$MESl_R = MES$	12.415	82,29
Differenze: - contenute tra SETT=0 e GG - rilevanti tra SETT=0 e MES	(2)	$MESl_R = \max(SETT, GG)$	15	0,10
Differenze: - contenute tra SETT>0 e MES - rilevanti tra SETT>0 e GG	(3)	$MESl_R = MES$	683	4,53
Differenze: - contenute tra SETT e MES>0 - rilevanti tra SETT e GG	(4)	$MESl_R = MES$	1.478	9,80
Differenze: - contenute tra MES>0 e GG - rilevanti tra SETT e GG/MES	(5)	$MESl_R = \max(SETT, GG)$	116	0,77
Differenze tutte rilevanti: distanza minima SET-MES	(6)	$MESl_R = MES$	94	0,62
Differenze tutte rilevanti: distanza minima SET-GG	(7)	$MESl_R = \max(SETT, GG)$	81	0,54
Differenze tutte rilevanti: distanza minima MES-GG	(8)	$MESl_R = MES$	8	0,05
Totale	-	-	15.086	100,00

In caso di mancata corrispondenza tra le tre variabili citate (casistiche 6, 7, 8) si utilizza, come ulteriore controllo, la variabile “numero di giorni per i quali spettano le deduzioni di cui art. 11, commi 2 e 3, del Tuir per il lavoro dipendente” indicata al punto 6 della Sezione *Dati fiscali* – Parte B.

Nel caso in cui vi fossero lavoratori di categorie particolari (minatori, doganieri, ecc.), inquadrati come assicurati ai Fondi sostitutivi gestiti dall'Inps e gestioni contabili separate (Dazieri, Ferrovieri, Pubblici creditizi, personale di volo), ai Fondi integrativi gestiti dall'Inps (Fondo Esattoriali, Fondo Gas); si ricorrerà all'utilizzo delle variabili al punto di cui al 29 (settimane retribuite) e 30 (giorni retribuiti) Parte C (Particolari categorie di lavoratori o di retribuzioni). Nell'indagine Eu-Silc 2005, appena sei lavoratori appartenenti a categorie particolari non hanno assegnate le corrispondenti mensilità retribuite.

Per giungere ad una stima dei mesi retribuiti in favore dei lavoratori pubblici si è utilizzata una tecnica differente rispetto a quella descritta. Innanzitutto, sono state impiegate le date di inizio (punto 65) e di fine (punto 66) periodo di servizio prestato nell'anno di riferimento dei redditi, esposte ambedue nella Sezione: *Dati Previdenziali ed Assistenziali Inpdap e Ipost* – Parte C – del Modello Cud/770 (figura 3.6).

Figura 3.6 – Sezione: Dati previdenziali ed assistenziali Inpdap e Ipost - Parte C, Cud/770_2005

Inquadramento			Codice fiscale Amm.ne			Decorrenza giuridica			Data inizio					
62	Codice comparto	Codice sottocomparto	Qualifica	63		64	giorno	me	anno	65	giorno	me	anno	
Data fine			N. giorni utili ai fini del TFR	Causa cessazione	Gestioni assicurative			Tipo impiego	Tipo servizio	Cassa/Comp	Giorni utili			
66	giorno	me	anno	67	68	69	70	71	72	73	74	75	76	
Maggiorazioni				Retribuzioni fisse e continuative				Retribuzioni accessorie						
77	78	n. giorni	79	80	n. giorni	81	82	n. giorni	83	84	n. giorni	85	86	
Retribuzioni di base per il 18%			Retribuzioni ai fini TFS			Retribuzioni utili ai fini TFR			Premio di produzione			Indennità non annualizzabili		
87			88			89			90			91		
D.Lgs. 165/97, art. 4			Tredicesima mensilità			Data opzione per il TFR			Codice fiscale Amm.ne versante retribuzione fissa e continuativa					
92			93			94			95					
Codice fiscale Amm.ne versante retribuzione accessoria			CONTRIBUTI A CARICO DEL LAVORATORE TRATTENUTI:						Contributi ai fini del TFR			L. 388/2000, art. 78, comma 6		
96			Ai fini pensionistici			Ai fini T.F.S.			99			100		
Indennità integrativa speciale annua conglobata			Imponibile contributivo ai fini TFR			CONTRIBUTI SOSPESI:			Percentuale aspettativa sindacale			Ente		
101			102			Anno 2002			Anno 2003			Anno 2004		
						103			104			105		
												106		
												107		

Le date di inizio e fine servizio, inizialmente lette in formato “giorno-mese-anno” (come da figura 3.6), sono state successivamente trasformate in vettore di 12 elementi (DCMM12_66) , per rappresentare i 12 mesi del calendario in cui il lavoratore ha percepito o meno un reddito (0=mese retribuito, 1=mese non retribuito). Ciò in analogia con il vettore dei mesi retribuiti dei lavoratori del privato (cfr. punto 15 figura 3.3). Il numero di mensilità retribuite si ottiene contando il numero di elementi che riportano valore “0” lungo il vettore dei mesi retribuiti (N_ZERI_DC66).

3.2.3 - Il reddito netto annuo da lavoro dipendente di fonte 730 e Unico persone fisiche

La variabile target Eurostat *PY010N*, indicante il reddito lordo da lavoro dipendente al netto del prelievo fiscale e contributivo, può essere rilevata, alternativamente¹¹ o contestualmente alla fonte Cud/770, presso gli archivi fiscali relativi al modello 730 e al modello Upf.

Si rammenta che per ottenere una stima di *PY010N* dalle informazioni del Cud si utilizza il saldo tra la retribuzione lorda al netto dei contributi sociali (punto 1 - Parte B) e le ritenute Irpef (punto 12 - Parte B), le addizionali regionali (punto 13 - Parte B) e le addizionali comunali (punto 14 - Parte B). Queste stesse informazioni figurano nelle Sezioni I e III - Quadro C - del modello 730 qualora il percettore di tali somme intenda assolvere all’obbligo di presentazione e di pagamento delle imposte (figura 3.7). Questo obbligo sorge nel momento in cui il percettore è tenuto a dichiarare al fisco redditi¹² che si aggiungono a quelli segnalati nel Cud, ovvero quando intenda avvalersi delle deduzioni/detrazioni (spese mediche, interessi passivi sul mutuo, ecc.) spettanti in sede di rimborso d’imposta. Così, ad esempio, il dato di cui al punto 1 del modello Cud/770 è identicamente riportato nel rigo C1 alla colonna 3 (C1_3) o, in presenza di più

¹¹ In linea di principio è impossibile che un reddito da lavoro dipendente sia esclusivamente dichiarato nel modello 730 o nel modello Unico persone fisiche, cioè senza che vi sia una contestuale emissione di un modello Cud. In questo caso ci riferiamo alla reale possibilità di deficit informativi, ossia di mancata copertura, degli archivi del Cud/770.

¹² Diversi da quelli per cui sorge l’obbligo di dichiarazione del modello Unico persone fisiche.

dichiarazioni Cud/770, nei successivi punti C2_3, C3_3. Alla stessa stregua, le ritenute fiscali dichiarate nel Cud (punti 12, 13 e 14) sono conformemente registrate nei righi C8, C9 e C10 del medesimo modello 730.

Figura 3.7 – Sezioni I, II, III - Quadro C - Redditi di lavoro dipendente e assimilati, mod. 730_2005

QUADRO C REDDITI DI LAVORO DIPENDENTE E ASSIMILATI											
SEZ. I - REDDITI DI LAVORO DIPENDENTE E ASSIMILATI											
C1	1	2	3	C2	1	2	3	C3	1	2	3
	Vedere istruzioni	Barrare se TFR assente	REDDITO		Vedere istruzioni	Barrare se TFR assente	REDDITO		Vedere istruzioni	Barrare se TFR assente	REDDITO
			,00				,00				,00
C4	PERIODO DI LAVORO (giorni per i quali spettano le deduzioni)						Lavoro dipendente	1	Pensione		2
SEZ. II - REDDITI ASSIMILATI PER I QUALI NON SPETTANO LE DEDUZIONI											
C5	Vedere istruzioni		REDDITO	C6	Vedere istruzioni		REDDITO	C7	Vedere istruzioni		REDDITO
			,00				,00				,00
SEZ. III - RITENUTE SU REDDITI DI LAVORO DIPENDENTE E ASSIMILATI, ADDIZIONALE REGIONALE E COMUNALE ALL'IRPEF											
C8	RITENUTE IRPEF		,00	C9	ADDIZIONALE REGIONALE		,00	C10	ADDIZIONALE COMUNALE		,00

Anche con riferimento al modello 730 sussiste il problema di scorporare il reddito da lavoro dipendente dalle restanti componenti, siano esse pensioni o altri redditi assimilati. In questo caso, le informazioni espresse in colonna 1 (Vedere istruzioni) e colonna 2 (Barrare se Tfr assente) lungo i rigi C1, C2 e C3 e quelle riportate al rigo C4 - colonna 1 e 2 (rispettivamente: *Lavoro dipendente* e *Pensione*), forniscono le linee guida per la ripartizione di tali redditi (figura 3.7). In particolare la colonna 1 relativa ai rigi C1, C2, C3 denota una serie di codici che classificano la tipologia di reddito. In totale sono contemplate tre tipologie di redditi associate ai seguenti codici:

- “1” redditi da pensione;
- “2” redditi da lavoro dipendente o assimilati derivanti da contratti a tempo determinato;
- “3” compensi percepiti da soggetti, in età pensionabile, impegnati in lavori socialmente utili.

La colonna 2 (rigi C1, C23, C3) è, a sua volta, utilizzata per individuare quella categoria di percettori di reddito da lavoro dipendente o assimilati per i quali non si applica il Tfr (ad esempio i dipendenti pubblici oppure i soggetti con redditi derivanti da collaborazioni coordinate e continuative compresi i lavoratori a progetto). I rimanenti punti C4_1 e C4_2 individuano il numero di giorni per i quali spettano le deduzioni di cui art. 11, commi 2 e 3, del Tuir, rispettivamente per il reddito da lavoro dipendente e per la pensione. Le regole utilizzate per stimare la variabile target *PY010N*, indicativa reddito netto da lavoro dipendente, sono contenute in sei punti. Per ciascuna di queste si definisce l'insieme delle “condizioni sottostanti” (*if*) e le relative “operazioni di calcolo” (*then*):

- 1) Condizione: {presenza di almeno un giorno utile per le deduzioni di cui all'art. 11 del Tuir riferite al solo reddito da lavoro dipendente} \cap {assenza di giorni utili per le deduzioni di cui all'art. 11 del Tuir riferite al solo reddito da pensione} \cap {assenza di redditi assimilati per i quali non spettano le deduzioni¹³} \cap {presenza di indicatori di reddito tutti diversi dalla pensione¹⁴}.

Calcolo di PY010N: Differenza tra *retribuzione lorda*, definita come il valore massimo tra la “sommatoria dei redditi ai punti C1_3, C2_3 e C3_3” e il “totale dei redditi lordi da lavoro dipendente” esposti nel prospetto di liquidazione (punto 4 alla colonna 1), e *ritenute fiscali* esposte nel prospetto di liquidazione.

¹³ Si riferiscono ai punti C5, C6, C7 del modello 730 equivalenti ai redditi esposti al punto 2, Parte B, del Cud.

¹⁴ Cioè i valori in colonna 1 dei rigi C1, C2, C3 modello 730 devono essere tutti diversi da “1”.

- 2) Condizione: {presenza di almeno un giorno utile per le deduzioni di cui all'art. 11 del Tuir riferite al solo reddito da lavoro dipendente} \cap {assenza di giorni utili per le deduzioni di cui all'art. 11 del Tuir riferite al solo reddito da pensione} \cap {presenza di redditi assimilati per i quali non spettano le deduzioni} \cap {presenza di indicatori di reddito tutti diversi dalla pensione}.

Calcolo di PY010N: Differenza tra *retribuzione lorda*, definita come il valore massimo tra la “sommatoria dei redditi ai punti C1_3, C2_3 e C3_3” e il “totale dei redditi lordi da lavoro dipendente esposti nel prospetto di liquidazione al netto dei redditi assimilati non deducibili (punti C5_2, C6_2, C7_2 figura 3.7)”, e la *quota di ritenute fiscali* (il cui totale è esposto nel prospetto di liquidazione) corrispondenti al solo reddito da lavoro dipendente¹⁵.

- 3) Condizione: {assenza di giorni utili per le deduzioni di cui all'art. 11 del Tuir riferite al solo reddito da lavoro dipendente} \cap {presenza di almeno un giorno utile per le deduzioni di cui all'art. 11 del Tuir riferite al solo reddito da pensione} \cap {assenza di redditi assimilati per i quali non spettano le deduzioni} \cap {presenza di indicatori di reddito tutti diversi dal reddito da lavoro dipendente}.

Calcolo di PY010N: ha valore nullo. Mentre si attribuisce valore positivo al reddito da pensione utilizzando lo stesso criterio di calcolo definito nel precedente punto 1.

- 4) Condizione: {assenza di giorni utili per le deduzioni di cui all'art. 11 del Tuir riferite al solo reddito da lavoro dipendente} \cap {presenza di almeno un giorno utile per le deduzioni di cui all'art. 11 del Tuir riferite al solo reddito da pensione} \cap {presenza di redditi assimilati per i quali non spettano le deduzioni} \cap {presenza di indicatori di reddito tutti diversi dal reddito da lavoro dipendente}.

Calcolo di PY010N: ha valore nullo. Mentre si attribuisce valore positivo al reddito da pensione utilizzando lo stesso criterio di calcolo definito nel precedente punto 2.

- 5) Condizione: {presenza di almeno un giorno utile per le deduzioni di cui all'art. 11 del Tuir riferite al solo reddito da lavoro dipendente} \cap {presenza di almeno un giorno utile per le deduzioni di cui all'art. 11 del Tuir riferite al solo reddito da pensione}.

Calcolo di PY010N: Differenza tra *retribuzione lorda*, definita come la “sommatoria dei redditi diversi da pensione esposti ai punti C1_3, C2_3 e C3_3”, e la *quota di ritenute fiscali* ascrivibili al reddito da lavoro dipendente sul totale dei redditi riportati ai punti C2, C3, C4.

- 6) Condizione (residuale): {presenza di redditi da lavoro dipendente/pensione o assimilati ($R > 0$)} \cap {inapplicabilità delle condizioni ai punti 1-5}:

$\{\{Condizione_1\} \cup \{Condizione_2\} \cup \{Condizione_3\} \cup \{Condizione_4\} \cup \{Condizione_5\}\} \cap \{R > 0\}$.

Calcolo di PY010N: identico al caso 5.

Le condizioni 1 e 2 definiscono l'ambito in cui è presente la sola componente da lavoro dipendente, con o senza la contestuale presenza di redditi non deducibili. Viceversa le condizioni 3 e 4 rappresentano la situazione opposta, dove sono presenti esclusivamente i redditi da pensione, con o senza altre componenti aggiuntive non deducibili. Infine le condizioni 5 e 6 indicano la presenza di redditi misti da lavoro dipendente e pensione, per i quali necessita una separazione dei relativi importi (tavola 3.2).

¹⁵ Per un esempio di ripartizione pro-quota delle ritenute fiscali si rinvia alla precedente nota 7.

Tavola 3.2 – Criteri di calcolo applicati alla stima della variabile target PY010N da fonte mod.730_2005. Valori assoluti e percentuali delle dichiarazioni del Quadro C – Indagine Eu-Silc 2005

CRITERIO DI CALCOLO DI PY010N	Casistica	Frequenza	Percentuale
Non applicabile	(0)	2	0,01
Condizione 1	(1)	9.059	57,94
Condizione 2	(2)	203	1,3
Condizione 3	(3)	5.871	37,55
Condizione 4	(4)	38	0,24
Condizione 5	(5)	226	1,45
Condizione 6	(6)	237	1,52
Totale	-	15.636	100,00

Si fa notare che nell'anno d'imposta 2004 (redditi 2005) vi erano in totale 15.741 dichiarazioni 730, di cui 2.461 in forma congiunta (dichiarante e coniuge). Rispetto al totale delle dichiarazioni, 15.636 mostravano valori positivi sul reddito da lavoro dipendente, pensione o redditi assimilati, del dichiarante. Delle 2.461 dichiarazioni congiunte, 1.853 mostravano valori positivi del reddito lavoro dipendente, pensione o redditi assimilati, afferenti al coniuge. In riferimento ai coniugi non è stata tuttavia possibile alcuna ripartizione tra le varie componenti di reddito, in quanto essi non erano tenuti alla compilazione del Quadro C ma soltanto alla dichiarazione del totale dei loro redditi - da lavoro dipendente, pensione e assimilati - nel prospetto di liquidazione.

Alternativamente al modello 730, il percettore di redditi da lavoro dipendente può utilizzare per gli stessi scopi il modello Unico persone fisiche. In realtà quest'ultimo modello trova un'applicazione più estesa. Infatti, tipicamente utilizzato dai percettori di compensi da lavoro autonomo (redditi da libera professione e redditi d'impresa), permette di includere in un'unica dichiarazione tanto le componenti di reddito autonomo quanto quelle da lavoro dipendente e assimilate. Il modello Upf, in talune circostanze, può essere utilizzato dai percettori di soli redditi da lavoro dipendente e assimilati¹⁶.

Il modello Upf presenta un apposito Quadro – RC – che alla Sezione I espone i redditi da lavoro dipendente e assimilati di cui al punto 1 del modello Cud/770 (figura 3.8). La struttura dei dati di questa sezione del modello Upf è identica a quella della Sezione I del Quadro C (Redditi da lavoro dipendente e assimilati) del modello 730 esaminato in precedenza¹⁷. In essa vengono riportate, per ciascun rigo (RC1, RC2, RC3 e RC4), tre colonne che denotano rispettivamente la tipologia di reddito (stessi codici del modello 730), la mancata applicazione del Tfr e l'importo del reddito corrispondente. Le informazioni relative alle ritenute fiscali dei redditi per il complesso delle Sezioni I e II del quadro RC sono riportate nella successiva Sezione III. Anche

¹⁶ Nella fattispecie trattasi di: 1) lavoratori dipendenti che hanno cambiato datore di lavoro e sono in possesso di più certificazioni di lavoro dipendente o assimilati (Cud), dove l'imposta corrispondente al reddito complessivo supera il totale delle detrazioni spettanti (ovvero se non è richiesto il conguaglio fiscale all'ultimo datore di lavoro); 2) lavoratori dipendenti che hanno percepito direttamente dall'Inps o da altri enti indennità e somme a titolo di integrazione salariale o ad altro titolo, se erroneamente non sono state effettuate le ritenute e se non ricorrono le condizioni di esonero; 3) lavoratori dipendenti che hanno percepito retribuzioni e/o redditi da privati non obbligati per legge ad effettuare ritenute d'acconto (per esempio collaboratori familiari, autisti ed altri addetti della casa).

¹⁷ L'unica differenza è che il modello Upf può accogliere 4 diverse dichiarazioni di redditi da lavoro dipendente in luogo delle 3 dichiarazioni del modello 730. Inoltre il modello Upf a differenza del modello 730 non consente la dichiarazione congiunta dei redditi, cioè del dichiarante unitamente al coniuge.

in questo caso vi è perfetta corrispondenza tra la Sezione III delle ritenute fiscali del modello Upf (righe RC10, RC11 e RC12) e la stessa Sezione III del modello 730 (righe C8,C9, C10). Dal momento che la struttura dei dati relativi al reddito da lavoro dipendente del modello Upf è omogenea a quella del modello 730, si è deciso di applicare al modello Upf le stesse regole di calcolo della variabile target *PY010N* (reddito da lavoro dipendente al netto delle ritenute fiscali e contributive) definite sul modello 730.

Figura 3.8 – Sezioni I,II,III,IV Quadro RC - Redditi di lavoro dipendente e assimilati, mod.Upf_2005

QUADRO RC	RC1	1	(Vedere istruzioni)	2	Barrare se TFR assente	Redditi	3				
REDDITI DI LAVORO DIPENDENTE E ASSIMILATI	RC2									,00	
	RC3									,00	
	RC4									,00	
Sezione I Lavoro dipendente e assimilati	RC5	Sommare gli importi di col. 3 dei righe da RC1 a RC4; (di cui 1 ,00) sommare l'importo del rigo RC5 agli altri redditi Irpef e riportare il totale al rigo RN1 col. 4							2	TOTALE	,00
	RC6	Periodo di lavoro (giorni per i quali spettano le ulteriori deduzioni)								Lavoro dipendente	Pensione
Sezione II Redditi assimilati per i quali non spettano le ulteriori deduzioni per lavoro dipendente	RC7	1	(Vedere istruzioni)			Redditi	2				,00
	RC8										,00
	RC9	Sommare gli importi di col. 2 dei righe RC7 e RC8; sommare l'importo del rigo RC9 agli altri redditi Irpef e riportare il totale al rigo RN1 col. 4									,00
Sezione III Ritenute su redditi di lavoro dipendente e assimilati, addizionale regionale all'IRPEF e addizionale comunale all'IRPEF	RC10	Ritenute subite (punto 12 del CUD 2005). Sommare l'importo di rigo RC10 e le altre ritenute e riportare il totale al rigo RN23 col. 2									,00
	RC11	Addizionale regionale all'IRPEF trattenuta dal datore di lavoro (punto 13 del CUD 2005)									,00
	RC12	Addizionale comunale all'IRPEF trattenuta dal datore di lavoro (punto 14 del CUD 2005)									,00
Sezione IV Ritenute su redditi derivanti da lavori socialmente utili in regime agevolato	RC13	Ritenute per lavori socialmente utili trattenute dal datore di lavoro. Vedere istruzioni.									,00
	RC14	Addizionale regionale all'IRPEF trattenuta dal datore di lavoro. Vedere istruzioni.									,00

I risultati dell'applicazione delle sei regole di calcolo sono illustrati nella seguente tavola 3.3

Tavola 3.3 – Criteri di calcolo applicati alla stima della variabile target PY010N Upf_2005. Valori assoluti e percentuali delle dichiarazioni del Quadro RC – Indagine Eu-Silc 2005

CRITERIO DI CALCOLO DI PY010N	Casistica	Frequenza	Percentuale
Non applicabile	(0)	94	1,46
Condizione 1	(1)	3.029	47,04
Condizione 2	(2)	97	1,51
Condizione 3	(3)	2.882	44,76
Condizione 4	(4)	50	0,78
Condizione 5	(5)	78	1,21
Condizione 6	(6)	209	3,25
Totale	-	6.439	100,00

3.2.4 - Le mensilità retribuite dei lavoratori dipendenti di fonte 730 e Unico persone fisiche

Le mensilità retribuite nell'anno si possono ricavare da ambedue le fonti fiscali, utilizzando il dato relativo al periodo (numero di giorni) di lavoro per il quale spettano le deduzioni di cui art. 11, commi 2 e 3, del Tuir per il lavoro dipendente”, corrispondente al dato riportato al punto 6 della Sezione *Dati fiscali* – Parte B del Cud/770. In questo caso non è possibile ricostruire il vettore dei mesi del calendario per i quali si percepisce una retribuzione ma solo calcolarne il numero.

Tavola 3.4 – Le meta-informazioni sul reddito da lavoro dipendente: definizioni, percettori, voci economiche e fonti del dato

Variabile target EU-SILC	Definizione	Percettori	Voci economiche	Fonte del dato				ISTAT Rilevazione EU-SILC
				Agenzia delle entrate		INPS Casellario Pensioni	Modello UPF	
				Modello 730	Modello 770			
PY010N <i>Reddito netto in denaro da lavoro dipendente</i>	Insieme delle componenti monetarie pagate dal datore di lavoro ai dipendenti a remunerazione dell'attività prestata al netto del prelievo contributivo e fiscale	Lavoratori dipendenti del settore privato e pubblico	<ul style="list-style-type: none"> - Stipendi e salari pagati in denaro per il periodo effettivamente lavorato; - Remunerazione del periodo non lavorato (ferie retribuite); - Incrementi stipendiali nel tempo; - Compensi ai direttori di azienda; - Pagamenti a cottimo; - Rimborsi spese per l'accudimento dei figli (asili nido e colonie climatiche); - Provvigioni, gratifiche, mance; - Pagamenti supplementari (compensi per lavoro straordinario, indennità di cassa, mensilità aggiuntive, ecc.); - Partecipazioni, titoli, azioni corrisposte al dipendente; - Premi di produttività; - Indennità di trasferimento in sedi disagiate; - Prestazioni di servizi/indennità di trasporto, rimborsi viaggio (casa-sede); 	✓	✓	✓	✓	✓
				P	P	P	P	✓ ⁰⁷
PY020N <i>Reddito netto in natura da lavoro dipendente</i>	Insieme delle componenti non-monetarie pagate dal datore di lavoro ai dipendenti a remunerazione dell'attività prestata al netto del prelievo contributivo e fiscale	Lavoratori dipendenti del settore privato e pubblico	<ul style="list-style-type: none"> - Concessione al lavoratore di veicoli aziendali in suo privato e sgravi dei relativi costi (assicurazione, bolli, rifornimenti, ecc.); - Mensa aziendale o prestazioni sostitutive (buoni pasto o ticket restaurant); - Altri beni o servizi forniti a titolo gratuito o a prezzo ridotto 	P	P	P	P	✓

Legenda: ✓ = completa rilevazione;
 ✓⁰⁷ = completa rilevazione dal 2007;
 P = parziale rilevazione;

3.3 - La costruzione del reddito da pensione

3.3.1 - Premessa

Prima di ricostruire la componente di reddito relativa ai trasferimenti pensionistici, occorre preliminarmente definire cosa si intende per pensione. Al riguardo si può fare riferimento alla definizione di F. Ceccato (1998): “per pensione si intende la prestazione periodica e continuativa in denaro erogata individualmente da amministrazioni pubbliche ed enti pubblici e privati al verificarsi di determinate condizioni ...”. Tra le condizioni che danno diritto al beneficio dei trattamenti pensionistici possiamo distinguere quelle connesse all’attività lavorativa del soggetto assicurato, comprendenti l’infortunio sul lavoro e le malattie professionali, e quelle relative ad eventi che, indipendentemente dall’attività lavorativa o dalla condizione professionale, compromettono o rendono impossibile lo svolgimento di un’attività lavorativa: inabilità, invalidità, vecchiaia, anzianità e morte. Ai fini della costruzione delle variabili target Eu-Silc legate al concetto di prestazione sociale vi sono tre categorie di trattamenti pensionistici associabili alle seguenti funzioni: *vecchiaia* “PY100N” (comprensiva dell’anzianità), *superstiti* “PY110N” e “PY130N” *invalidità* (comprensiva della sottofunzione inabilità, infortuni sul lavoro e malattie professionali). In base al Esspros la funzione *vecchiaia* si traduce in un provvedimento di protezione sociale a copertura dei rischi connessi al sopraggiungere di un’età avanzata: perdita della capacità di procurarsi guadagno connessa all’invecchiamento, perdita di autosufficienza, ecc.. La funzione *superstiti* include i trattamenti corrisposti a persone che, perdendo il coniuge o altro parente, non hanno più la fonte primaria di sostentamento. La funzione *invalidità* include le prestazioni che garantiscono un sostegno economico alle persone con capacità lavorativa ridotta o compromessa a causa di invalidità fisica o mentale. I trattamenti pensionistici erogati dal sistema previdenziale ed assistenziale italiano non esauriscono le tre funzioni legate alle variabili target Eu-Silc ma rappresentano tuttavia una quota preponderante. Infatti, alla formazione dei redditi compresi variabili target Eu-Silc concorrono tanto le prestazioni di natura pensionistica quanto quelle non pensionistiche¹⁸. Le fonti amministrative utilizzate allo scopo sono il Casellario centrale dei pensionati, il modello Cud/770, il modello 730 e Unico persone fisiche. Come si evince dalla seguente tavola, sintesi del contenuto informativo di ciascuna fonte, solo grazie all’utilizzo congiunto delle diverse fonti è possibile stimare le variabili target Eu-Silc: importo netto annuo delle pensioni e le relative mensilità distinte per categoria o funzione.

Tavola 3.5 – Il contenuto informativo delle fonti amministrative sui pensionati e sui redditi da pensioni

FONTE DEL DATO	Principali caratteri rilevabili					Univ. di riferimento	Unità di analisi	
	Importo lordo reddito da pensione		Importo netto reddito da pensione		Mensilità			Tipologia di pensione (funzione)
	Mensile	Annuale	Mensile	Annuale				
Casellario pensioni	✓(a)	✓(c)	-	-	✓(c)	✓(a)	Totale percettori di pensioni del sistema previdenza italiano	Pensionato e/o Pensione
Modello Cud/770	✓(b)	✓(a)	✓(b)	✓	✓(b)	-	Percettori di pensioni tassabili dal fisco	Pensionato
Modello 730	✓(b)	✓(a)	✓(b)	✓	✓(b)	-	Percettori di pensioni tassabili con Md.730	Pensionato
Modello Upf	✓(b)	✓(a)	✓(b)	✓	✓(b)	-	Percettori di pensioni tassabili con Md.Upf	Pensionato

(a) dato rilevato.

(b) dato derivato dall’insieme di più variabili rilevate.

(c) dato in parte stimato (neopensionati degli Archivi CP 2003-2004). Per l’Archivio centrale dei pensionati 2005 esso è invece rilevato.

¹⁸ Per una disamina di questo argomento si rinvia a Consolini P. (2000, 2002), Guerrucci D. (1996).

3.3.2 - Il reddito lordo annuo da pensione di fonte CP

Il Casellario centrale dei pensionati dell'Inps¹⁹ costituisce l'unica fonte amministrativa nel nostro paese in grado di censire il numero dei percettori di trattamenti pensionistici e il tipo di pensioni erogate o percepite in un determinato anno. Sebbene la fonte rilevi l'importo lordo mensile delle pensioni individuali erogate al 31 dicembre di ciascun anno, che non risponde alla definizione Eu-Silc delle variabili target *PY100N-PY110N-PY130N* (reddito netto annuo della prestazione sociale in denaro -pensionistica e non- per la corrispondente funzione Esspros), essa permette tuttavia di identificare correttamente i pensionati e le varie tipologie di pensioni legate al nostro sistema di protezione sociale. Per ciascuna prestazione l'archivio contiene, oltre alle caratteristiche della pensione e al suo importo lordo in pagamento, una serie di informazioni sul beneficiario²⁰. Sulla base di tali elementi è possibile valutare non solo il numero delle pensioni vigenti alla fine dell'anno, ma anche il numero dei soggetti percettori delle stesse. Il numero dei beneficiari non coincide necessariamente con quello delle pensioni, in quanto ogni pensionato può percepire più trattamenti pensionistici. Ad esempio, il titolare di una pensione di vecchiaia può avere diritto, al contempo, ad una parte della pensione di vecchiaia del coniuge deceduto. Ogni soggetto può essere, in linea teorica, titolare di più pensioni appartenenti a diverse funzioni; ciò accade quando uno stesso individuo cumula una prestazione di vecchiaia, con una rendita indennitaria e con una pensione ai superstiti. I dati del CP, opportunamente classificati in funzione delle variabili target Eu-Silc, sono stati utilizzati per individuare i beneficiari dei trattamenti pensionistici che partecipano all'indagine Eu-Silc e per stimare il loro reddito lordo annuo da pensione. Quest'ultimo dato è ottenuto, come prima approssimazione, moltiplicando l'importo mensile al 31 dicembre dell'anno di riferimento dei redditi per il numero massimo di mensilità erogabili nell'anno solare. Chiaramente tale ipotesi non è esente da errori, e conduce inevitabilmente ad una sovrastima del valore dell'importo lordo annuo percepito dalla sottocategoria dei neopensionati. All'interno del CP è però possibile individuare questi ultimi attraverso l'analisi dell'anno di decorrenza della pensione, consentendo così la correzione dei rispettivi importi ad un successivo confronto incrociato con dati di altra fonte. La fonte del CP da sola non permette, dunque, di rilevare il reddito da pensione al netto delle ritenute fiscali, se non nei casi in cui la corrispondente pensione è esente da tassazione, cioè quando l'importo lordo e netto da pensione coincidono. Nell'ambito del nostro sistema di protezione sociale le pensioni esenti da tassazione comprendono : 1) pensioni agli invalidi civili, ai non vedenti e ai non udenti (non reversibili); 2) pensioni o assegni sociali (non reversibili); 3) rendite per infortunio sul lavoro (reversibili); 4) trattamenti diretti per pensione di guerra, per assegno di benemerenzia ai perseguitati politici, per assegno vitalizio agli internati nei campi di sterminio e per assegno annesso alle decorazioni al valore militare (reversibili). Inoltre sono esenti da tassazione tutte le altre pensioni che per importo non superano (in complesso) la "no tax-area".

3.3.3 - Il reddito netto annuo da pensione di fonte Cud/770

L'archivio fiscale dei modelli Cud/770 rileva, al punto 1 della Parte B, i "redditi per i quali è possibile fruire della deduzione di cui all'art. 11 e delle detrazioni di cui all'art. 14 del Tuir". I redditi in esame, come già detto, comprendono diverse poste, tra cui stipendi, pensioni, mensilità aggiuntive, indennità di trasferta, borse di studio, compensi derivanti da rapporti di

¹⁹ Il "Casellario centrale per la raccolta, la conservazione e la gestione dei dati e degli elementi relativi ai titolari di trattamenti pensionistici" è stato istituito in base al D.p.r. n. 1388/71, successivamente modificato dal decreto legge n. 352/78 e dalla legge n. 85/95. Secondo la disciplina vigente tutti gli enti erogatori di prestazioni pensionistiche obbligatorie e integrative e i soggetti erogatori di pensioni e assegni continuativi assistenziali sono tenuti a comunicare periodicamente all'Inps, entro il primo trimestre dell'anno la situazione delle prestazioni pensionistiche vigenti al 31 dicembre dell'anno precedente. Alla fine del 2003 risultavano presenti nel Casellario 659 enti erogatori di prestazioni pensionistiche.

²⁰ Per maggiori dettagli sul contenuto informativo del Casellario centrale dei pensionati si rinvia a Nazzaro O. (1999).

collaborazione coordinata e continuativa, compensi ai soci di cooperative di produzione e lavoro, ecc. L'aggregato al punto 1 fa riferimento al reddito al lordo del prelievo fiscale e al netto dei contributi sociali. Per scorporare i trasferimenti pensionistici dalle restanti componenti economiche sono stati utilizzati gli stessi criteri applicati al caso dei redditi da lavoro, in un'ottica rovesciata. Pertanto, sono state analizzate le stesse variabili, di cui al punto 6 e 7 della Parte B del Cud, indicanti il numero di giorni per i quali spettano le deduzioni, di cui all'art. 11 del Tuir punto, rispettivamente per il lavoro dipendente e la pensione. Parallelamente, in mancanza di simili informazioni, sono stati esaminati i dati retributivi esposti nella Parte C relativi ai lavoratori dipendenti del settore privato o pubblico, ciò al fine di eliminare dal campo di osservazione i presunti percettori di redditi da lavoro. Infine come criterio ultimo, si è considerata la tipologia del sostituto d'imposta: la presenza di un ente di previdenza fa propendere per l'assegnazione del reddito da pensione.

Una volta applicato il procedimento di estrazione della componente di reddito lordo da pensione, il passo successivo consiste nel decurtare l'ammontare del prelievo fiscale. Il criterio adottato considera la differenza tra il reddito al punto 1 (retribuzione lorda al netto dei contributi sociali) e le ritenute Irpef (punto 12), le addizionali regionali (punto 13) e le addizionali comunali (punto 14). Di seguito si illustra la formula adottata nel calcolo del reddito netto da lavoro dipendente sulla base degli elementi rilevati nel modello Cud/770.

$$[2] \quad NY^P = \underbrace{GY^P - SC^P}_{\substack{\text{Punto 1} \\ \text{Parte B} \\ \text{CUD}}} - T^P \quad \leftarrow \substack{\text{Punti 12-13-14} \\ \text{Parte B CUD}}$$

dove:

NY^P : reddito netto da pensione (al netto dei contributi solidarietà e delle ritenute fiscali alla fonte);

GY^P : trattamento pensionistico lordo;

SC^P : contributo di solidarietà (obbligatorio) a carico dei pensionati con pensione superiore ad una soglia definita per legge²¹;

T^P : ritenute fiscali alla fonte applicate al reddito da pensione (Irpef e addizionali regionali e comunali).

La fonte Cud/770, pur rilevando l'importo annuo esatto (lordo/netto) delle pensioni individuali soggette a tassazione ordinaria, non contiene però alcuna informazione atta a distinguere la categoria o funzione di protezione sociale associata. Questo archivio fiscale da solo si dimostra poco utile alla stima delle predette variabili target, mentre acquista notevole rilevanza se affiancato al CP. Per quanto riguarda la stima del numero di mensilità o dei pagamenti delle pensioni individuali si può impiegare la variabile di cui al punto 7 della Parte B del Cud, indicante il numero di giorni per i quali spettano le deduzioni per la pensione.

3.3.4 - Il reddito netto annuo da pensione di fonte 730 e Upf

Il reddito da pensione al netto del prelievo fiscale e contributivo, può essere rilevato presso gli archivi fiscali relativi al modello 730 e al modello Upf, in alternativa²² o congiuntamente alla

²¹ Trattasi di importo lordo annuo pari a venticinque volte quello stabilito dall'articolo 38, comma 1, della legge 28 dicembre 2001, n. 448, rivalutato annualmente nella misura stabilita dal comma 5, lettera d) dello stesso art. 38. Nell'anno 2004 la soglia è pari a 13.398,75 euro lordi mensili. Il contributo in oggetto decorre dal 1° gennaio 2004 e si applica per un periodo di tre anni, pertanto sino a tutto il 31 dicembre 2006.

²² Cfr. precedente nota 11, dove in luogo dei redditi da lavoro dipendente si considerano i redditi da pensione.

fonte Cud/770. Anche per questa tipologia di reddito vale l'equivalenza tra le informazioni riportate nel modello Cud/770 alla Sezione *Dati Fiscali* –Parte B– e i dati alle Sezioni I e III - Quadro C - del modello 730, ovvero dei dati alle Sezioni I e II – Quadro RC – del modello Upf. Il criterio utilizzato per scorporare la componente da reddito da pensione dalle restanti voci si ricava direttamente dalle 6 regole applicate al caso dei redditi da lavoro dipendente e analizzate nel § 3.2.3, rispettivamente per le due fonti di dati. Le mensilità delle pensioni erogate/percepite nell'anno sono stimabili da ambedue le fonti fiscali, utilizzando il dato relativo al periodo (numero di giorni) per il quale spettano le deduzioni di cui art. 11, commi 2 e 3, del Tuir per la pensione". L'impiego fine a stesso di queste basi di dati non è molto utile per la stima delle variabili target, al pari della fonte Cud, mentre diventa essenziale²³ nel caso di integrazione con i dati del CP.

3.4 La costruzione di parte dei trasferimenti non pensionistici

3.4.1 - Premessa

Per allargare il campo di rilevazione delle fonti amministrative su parte dei trasferimenti monetari non pensionistici, inclusi tra le variabili target Eu-Silc “*PY090N, PY100N, PY110N e HY050N*”, si è deciso di ricostruire l'informazione attinente a due tipologie di prestazione, di peculiare interesse per peso economico e rilevanza sociale: gli assegni al nucleo familiare (Anf) e le liquidazioni di fine rapporto (Lfr). Innanzitutto occorre osservare come, in termini monetari, i due trattamenti non pensionistici rappresentino rispettivamente il 3,2 per cento e l'11,6 per cento del totale delle prestazioni sociali in denaro erogate in Italia nell'anno 2004²⁴. Gli assegni al nucleo familiare costituiscono una componente rilevante²⁵ della variabile target Eu-Silc “*HY050N*”, relativa ai trasferimenti monetari alle famiglie che appartengono alla funzione di protezione sociale Esspros: «*Famiglia e cura dei figli*». Le fonti amministrative del Cud/770 costituisce l'unica fonte amministrativa attualmente in grado di rilevare entrambe con un livello di dettaglio *micro* (percettore). In particolare la fonte in questione copre il sotto-universo dei percettori di Anf costituiti dai lavoratori dipendenti del settore privato e l'intero universo dei percettori di liquidazione di fine rapporto. Sono, dunque, esclusi dal campo di indagine i percettori di Anf dipendenti della pubblica amministrazione²⁶, mentre per gli Anf destinati ai pensionati si è fatto riferimento ai dati del CP. Le liquidazioni di fine rapporto, intese come prestazioni sociali erogate a tantum, comprendono una serie di trattamenti che vanno dal trattamento di fine rapporto Tfr, alle indennità premio di fine servizio, alle indennità di buonuscita e alle indennità di anzianità dei parastatali. Tutte queste varie forme di liquidazioni possono coprire tre funzioni di protezione sociale, incluse in Eu-Silc con le seguenti sigle: *PY100N* “trattamenti di vecchiaia” (in caso di quiescenza), *PY090N* “trattamenti di disoccupazione”(a seguito di licenziamento) e *PY110N* “trasferimenti ai superstiti” (a favore degli eredi).

²³ Soprattutto in considerazione del fatto che il modello Cud/770 presenta problemi di mancata copertura dell'universo dei percettori di questi redditi.

²⁴ Cfr. Social protection expenditure and receipts, years 1990-2004, Eurostat.

²⁵ Secondo la fonte citata nella precedente nota, essi rappresentano in Italia il 71 per cento del totale delle prestazioni sociali in denaro afferenti alla funzione Esspros *Famiglia e cura dei figli* (anno 2004). Di contro, nell'indagine Eu-Silc gli Anf rappresentano la quasi totalità delle prestazioni appartenenti a questa funzione, poiché non è stato possibile rilevare le indennità di maternità e i congedi parentali dei lavoratori dipendenti.

²⁶ Secondo una ricostruzione dell'Autore sulle fonti di rilevazione della spesa per gli Anf nell'anno 2002, la quota parte di spesa per i trattamenti familiari ascrivibile agli Anf dei dipendenti pubblici si aggira attorno al 10 per cento. Per un'analisi delle fonti amministrative utilizzabili ai fini della rilevazione delle prestazioni monetarie non pensionistiche si rimanda all'articolo di Consolini P. e De Carli R. (2002).

3.4.2 - Gli assegni al nucleo familiare di fonte Cud/770

La ricostruzione degli importi annui teorici spettanti ai beneficiari degli assegni al nucleo familiare si basa su una formula di calcolo che mette in relazione quattro elementi rilevati nella fonte del Cud/770. Gli elementi in questione riguardano le tre “coordinate Anf” rilevate rispettivamente ai punti 23-25 della Sezione I *Dati previdenziali e assistenziali Inps* – Parte C – del modello Cud/770 (figura 3.9) e il numero di mensilità retribuite stimate in precedenza. In particolare le coordinate Anf si articolano nei seguenti punti: 1) Tabella²⁷; 2) Numero familiari²⁸; 3) Classe (ovvero la fascia di reddito familiare). A ciascuna terna di coordinate è associato un importo mensile dell’assegno, definito ed aggiornato periodicamente sulla base di apposite circolari Inps.

Figura 3.9 – Sezione 1 DATI PREVIDENZIALI ED ASSISTENZIALI INPS - Parte C - Cud/770_2005

The image shows a portion of the Cud/770_2005 form, specifically the 'PARTE C' section for 'DATI PREVIDENZIALI ED ASSISTENZIALI INPS'. The form is divided into several columns and rows. At the top, there are fields for 'Codice fiscale del percipiente' and 'Progressivo Comunicazione'. Below this, the 'PARTE C' section is organized into a grid. The first row contains fields for 'Qualifica' (1, 2, 3), 'Matricola azienda' (4), 'Prov. Lav.' (5), 'IVS' (6), 'DS' (7), 'ALTRE' (8), 'Competenze correnti' (9), and 'Altre competenze' (10). The second row is for 'MESI RETRIBUITI NELL'ANNO', with sub-sections for 'Settimane retrib.' (11), 'Settimane utili' (12), 'Giorni retrib.' (13), and 'Tutti con esclusione di' (14, 15). The 'Mesi retribuiti' are represented by a grid of boxes for each month (T, G, F, M, A, M, G, L, A, S, O, N, D). The 'Mesi con esclusione di' are represented by a grid of boxes for each month (T, G, F, M, A, M, G, L, A, S, O, N, D). The third row contains 'CONTRATTO' (16, 17) and 'Livello inquadramento' (18). The fourth row contains 'Data cessazione' (19), 'Tipo rapporto' (20), 'Trasf. rapporto' (21), and 'Accantonamento T.F.R. spettante' (22). A box highlights the 'COORD. ASSEgni FAMILIARI' section, which includes 'Tabella' (23), 'Num. Fam.' (24), and 'Classe' (25). Arrows point from the 'COORD. ASSEgni FAMILIARI' box to the 'Mesi retribuiti' and 'Mesi con esclusione di' sections.

L’importo annuale dell’assegno si ricava moltiplicando il valore assunto dall’importo mensile per i mesi retribuiti nell’anno di riferimento (formula di calcolo). I problemi incontrati nel processo di ricostruzione degli importi Anf hanno riguardato rispettivamente: *i*) le mancate risposte parziali (*item non response*) nella terna delle coordinate Anf; *ii*) i valori inammissibili in almeno una delle tre coordinate; *iii*) i valori singolarmente ammissibili ma tra loro incoerenti, ovvero tali da non generare valori positivi nei rispettivi importi. Non è stato tuttavia possibile quantificare il problema delle mancate risposte totali (*unit non response*) non avendo a disposizione altre fonti a livello micro che coprono gli stessi percettori. I problemi di mancata rilevazione totale possono essere originati da diversi fattori, quali ad esempio i ritardi nella comunicazione dei dati fiscali del personale da parte di un’azienda, le omissioni in fase di compilazione dei modelli, errori nella trasmissione telematica ecc. Al momento attuale in Istat si sta valutando l’impatto di tale errore, in aggregato per vari settori di attività economica, confrontando la fonte fiscale con la base dati Inps relativa al DM10²⁹. Per quanto concerne le mancate risposte parziali, sono state applicate opportune tecniche di tipo deterministico che permettono il recupero dell’informazione mancante (*missing value*) direttamente dai modelli Cud/770, quali il livello di reddito, il sesso del dichiarante e il numero di componenti il nucleo (DC_24). Purtroppo, al momento della imputazione delle coordinate Anf, non è stato possibile utilizzare variabili quali lo stato civile, il numero dei componenti il nucleo, il numero dei figli

²⁷ Essa si riferisce alla composizione del nucleo familiare ed è utilizzata ai fini della determinazione dell’importo dell’assegno spettante. Con riferimento alla Circolare Inps n. 94 del 15 giugno 2004, vi sono, a decorrere da quella data e per la durata di un anno, un totale di 15 Tabelle così codificate: 11, 12, 13, 14, 15, 16, 17, 18, 19 20A, 20B, 21A, 21B, 21C, 21D.

²⁸ Occorre osservare che la variabile “numero familiari” non rappresenta esattamente il numero dei componenti il nucleo familiare ma è piuttosto riferita all’insieme dei membri della stessa famiglia a cui è riconosciuto il diritto al beneficio. Questi soggetti sono costituiti dal *richiedente, dal coniuge* anche se non convivente purché non separato, *dai figli* fino a 18 anni compiuti, *dai nipoti in linea retta* minori e viventi a carico dell’ascendente, *dai nipoti in linea collaterale* se affidati dai competenti organi, ovvero se minorenni o inabili (purché orfani di entrambi genitori e non titolari di pensioni ai superstiti), *dai fratelli e sorelle* minorenni o inabili (purché orfani di entrambi genitori e non titolari di pensioni ai superstiti). Inoltre, si fa notare che sono considerati *equiparati* ai figli legittimi: i figli legittimati, adottivi, naturali legalmente riconosciuti o giudizialmente dichiarati, nati da precedente matrimonio dell’altro coniuge, e i minori dati in affidamento.

²⁹ Cfr. Peperoni C. e Tuzi F. (2007).

minori, in quanto non aggiornate o indisponibili dall'archivio Sigif o dal registro familiare Eu-Silc. L'uso di queste informazioni avrebbe consentito una maggiore precisione durante la fase di imputazione delle coordinate. Esse sono però state utilizzate *ex-post* per valutare la correttezza delle assunzioni fatte nel processo di *editing* ed imputazione e più generale per analizzare la qualità dei risultati finali. Il problema relativo alle coordinate inammissibili (*unreliable value*) è chiaramente assimilabile al problema delle mancate risposte parziali, una volta individuati ed annullati i rispettivi valori. Più complesso è il problema delle coordinate tra loro incoerenti che, in assenza di altre informazioni, sono stati trattati agendo sulla variabile "numero di familiari" condizionatamente ai valori assunti dalle variabili: "tabella Inps" e "fascia di reddito". Un caso tipico di coordinata con valori inammissibili è quello della variabile "numero di familiari" che assume valori inferiori alla soglia minima consentita per la assegnazione dell'Anf (ad esempio nel caso della «Tabella 11» la soglia minima è 2). In questi casi si è proceduto ad assegnare la modalità più frequente. Per altro verso, il problema delle coordinate incoerenti può essere ricondotto all'incompatibilità tra i valori della variabile "numero di familiari" e quelli della combinazione "Tabella_Inps-Fascia_di_reddito". L'ipotesi sottostante è che la coppia di valori di quest'ultima combinazione siano corretti, mentre si interviene sul numero dei componenti assegnando il valore minimo tale da generare un importo dell'Anf.

Con riferimento alla base dati Cud/770 utilizzata nell'indagine Eu-Silc 2005, si nota la presenza 1.949 coordinate Anf complete (98,3 per cento) e 33 incomplete (1,7 per cento)³⁰. La strategia di imputazione seguita è stata quella di completare l'informazione lungo i vettori delle coordinate incomplete in modo sequenziale, partendo dalla mutabile: *Tabella Inps* (punto 23 – Parte C), per poi procedere all'imputazione della variabile discreta: *Classe reddituale* (punto 25 – parte C) e finire con la variabile conteggio: *Numero familiari* (punto 24 – parte C). Osservando la distribuzione di frequenza della mutabile *Tabella Inps*, si nota che la moda è rappresentata dalla modalità "11" con il 78,7 per cento dei casi, seguono in termini di numerosità la modalità "12" (10,4 per cento) e la modalità "21A" (9,7 per cento). Messe insieme queste tre modalità includono il 98,8 per cento dei casi. Ciascuna modalità, inoltre, corrisponde ad una particolare tipologia familiare. Pertanto, in fase di imputazione si è stabilito di assegnare una delle tre modalità in modo deterministico, sulla base delle distribuzioni di frequenza parziali costruite sulle variabili: "sesso del percettore" (richiedente) e "numero dei familiari" (aventi titolo).

Si osserva così che la categoria "11", corrispondente ai nuclei con entrambi i coniugi ed almeno un figlio minore o disabile, è fortemente connessa alle tipologie familiari con tre o più componenti (DC_24). Infatti la percentuale di casi in cui si presenta la categoria 11, varia dal 92,3 per cento in riferimento alla tipologia con tre membri al 100 per cento per le famiglie con 6 o più componenti.

La categoria "12", riferita alle tipologie delle famiglie monogenitore o conviventi con almeno un figlio minore o disabile, è concentrata sui nuclei familiari con due componenti, dove il titolare dell'assegno è una donna (88,2 per cento). Infine la tipologia "21A", relativa ai coniugi senza figli minori o disabili, è rappresentativa dei nuclei con due componenti il cui titolare è un uomo (78,5 per cento). Di seguito si illustrano i risultati dell'imputazione della mutabile *Tabella Inps* (tavola 3.6).

³⁰ Prendendo in esame il campione teorico Eu-Silc 2005, dove sono incluse anche le famiglie non intervistate, si nota che il numero delle coordinate complete è pari a 2.282 (98,2 per cento) e mentre quelle incomplete ammontano a 42 (1,8 per cento).

Tavola 3.6 – I principali risultati del processo di imputazione sulla variabile “Tabella Inps” – Indagine Eu-Silc 2005

TABELLA INPS (DC_23)	Prima dell'imputazione				Dopo l'imputazione			
	Famiglie del campione teorico		Famiglie intervistate		Famiglie del campione teorico		Famiglie intervistate	
	Frequenza	Percentuale	Frequenza	Percentuale	Frequenza	Percentuale	Frequenza	Percentuale
11	1.814	78,70	1.549	78,75	1.827	78,61	1.561	78,76
12	239	10,37	201	10,22	240	10,33	202	10,19
14	18	0,78	17	0,86	18	0,77	17	0,86
15	4	0,17	3	0,15	4	0,17	3	0,15
17	4	0,17	4	0,20	4	0,17	4	0,20
21A	224	9,72	192	9,76	229	9,85	194	9,79
21B	1	0,04	1	0,05	1	0,04	1	0,05
21C	1	0,04	-	-	1	0,04	-	-
Totale	2.305	100,00	1.967	100,00	2.324	100,00	1.982	100,00

Nell’ambito delle famiglie intervistate, cioè quelle su cui vengono rilevati i dati reddituali, si notano minimi cambiamenti nella distribuzione delle *Table Inps* per effetto dell’imputazione della corrispondente coordinata.

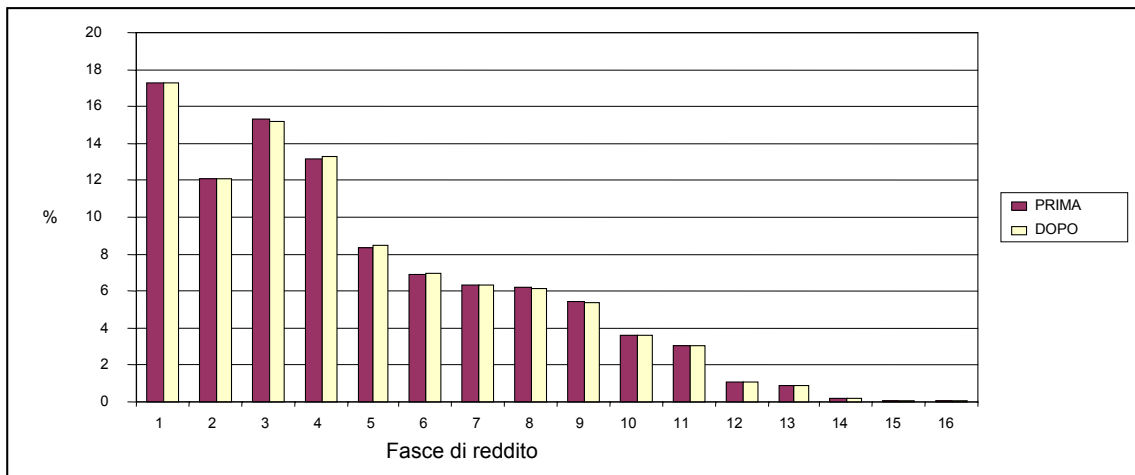
Per quanto concerne la variabile *Classe di reddito* si è proceduto ad imputare i corrispondenti valori mancanti, utilizzando l’informazione relativa al reddito percepito dal lavoratore che richiede l’Anf, così come risulta indicato nei punti 1 e 2 della parte B – Sezione dati fiscali – del modello Cud/770. Questa strategia di imputazione, pur non essendo scevra di difetti, è stata ritenuta la migliore tra le possibili. Le limitazioni sorgono in quanto, ai fini dell’assegnazione dell’Anf, il reddito del nucleo familiare include tanto eventuali componenti aggiuntive di reddito dello stesso percettore (non presenti nel Cud), quanto possibili redditi del coniuge. Mentre è lecito ipotizzare l’assenza di redditi da parte degli altri componenti. Inoltre sussiste il problema del disallineamento del riferimento temporale dei redditi, poiché il reddito familiare, o per meglio dire la classe di reddito esposta nel modello Cud, è riferito all’anno precedente dell’assegnazione dell’Anf³¹, mentre il reddito rilevato nel Cud (del lavoratore richiedente) ed utilizzato in fase di imputazione è riferito all’anno attuale. Per quanto riguarda i dati Cud/770 utilizzati nell’indagine Eu-Silc 2005, si nota come vi siano in tutto 24 record con informazione mancante sulla classe di reddito (1.951 classi valorizzate su 1.982 coordinate complete ed incomplete). La procedura di imputazione assegna valori per 23 record, mentre fallisce in un solo caso. In quest’ultima circostanza, il valore di reddito ricavato dal Cud è talmente elevato da risultare incompatibile con le restanti due coordinate Anf. L’incoerenza riscontrata determina la mancata assegnazione dell’Anf. I restanti valori imputati hanno tuttavia scarso impatto sulla distribuzione finale della classi di reddito (tavola 3.7). Dal grafico 3.1 che mette a confronto le distribuzioni per classi di reddito prima e dopo l’imputazione, si evince una leggera riduzione del peso della classe 3 a cui si contrappone un incremento, in termini percentuali, delle classi 4, 5, e 6.

³¹ Per la precisione, ai fini dell’assegnazione degli Anf per il periodo luglio-dicembre dell’anno *T* si utilizza il reddito familiare dell’anno *T-1*. Mentre per l’assegnazione degli Anf del periodo gennaio-giugno dello stesso anno *T* si considera il reddito familiare all’anno *T-2*.

Tavola 3.7 – I principali risultati del processo di imputazione sulla variabile “Classe di reddito” – Indagine Eu-Silc 2005

FASCIA DI REDDITO (DC_25)	Prima dell'imputazione				Dopo l'imputazione			
	Famiglie del campione teorico		Famiglie intervistate		Famiglie del campione teorico		Famiglie intervistate	
	Frequenza	Percentuale	Frequenza	Percentuale	Frequenza	Percentuale	Frequenza	Percentuale
1	391	17,06	338	17,26	397	17,09	342	17,26
2	291	12,70	237	12,10	294	12,66	240	12,12
3	346	15,10	300	15,32	349	15,02	301	15,19
4	302	13,18	258	13,18	309	13,3	263	13,28
5	188	8,20	164	8,38	193	8,31	168	8,48
6	158	6,89	135	6,89	161	6,93	138	6,97
7	146	6,37	124	6,33	147	6,33	125	6,31
8	131	5,72	121	6,18	133	5,73	122	6,16
9	129	5,63	107	5,46	129	5,55	107	5,40
10	87	3,80	70	3,58	88	3,79	71	3,58
11	74	3,23	60	3,06	74	3,19	60	3,03
12	24	1,05	21	1,07	24	1,03	21	1,06
13	18	0,79	17	0,87	18	0,77	17	0,86
14	5	0,22	4	0,20	5	0,22	4	0,20
15	1	0,04	1	0,05	1	0,04	1	0,05
16	1	0,04	1	0,05	1	0,04	1	0,05
Totale	2.292	100,00	1.958	100,00	2.323	100,00	1.981	100,00

Figura 3.10 – Differenze tra la distribuzioni delle fasce di reddito prima e dopo l'imputazione



La procedura utilizzata per imputare il valori mancanti ed inammissibili sulla variabile *Numero familiari* ha richiesto l'individuazione del valore minimo di questa coordinata utile a generare un importo positivo dell'Anf. Come si può osservare dalla seguente tavola 3.8, nell'indagine Eu-Silc 2005 sono stati imputati 11 valori mancanti ed assegnati 20 nuovi valori (che assumevano in precedenza valori inammissibili e/o incoerenti con le restanti coordinate) della variabile *numero familiari*.

Tavola 3.8 – Matrice delle transizioni del Numero di familiari (aventi titolo) prima e dopo la fase di editing ed imputazione.

N. FAMILIARI (DC_24) PRIMA DELLA IMPUTAZIONE	Numero familiari (DC_24) dopo l'imputazione								Totale
	2	3	4	5	6	7	8	9	
MISSING	5 (45,45)	6 (54,55)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	11 (100,00)
1	0 (0,00)	2 (100,00)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	2 (100,00)
2	331 (95,94)	13 (3,77)	0 (0,00)	1 (0,29)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	345 (100,00)
3	0 (0,00)	811 (99,63)	2 (0,25)	0 (0,00)	1 (0,12)	0 (0,00)	0 (0,00)	0 (0,00)	814 (100,00)
4	0 (0,00)	0 (0,00)	679 (99,85)	0 (0,00)	0 (0,00)	1 (0,15)	0 (0,00)	0 (0,00)	680 (100,00)
5	0 (0,00)	0 (0,00)	0 (0,00)	112 (100,00)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	112 (100,00)
6	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	13 (100,00)	0 (0,00)	0 (0,00)	0 (0,00)	13 (100,00)
7	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	3 (100,00)	0 (0,00)	0 (0,00)	3 (100,00)
8	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	1 (100,00)	0 (0,00)	1 (100,00)
9	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	0 (0,00)	1 (100,00)	1 (100,00)
Totale	336 (16,95)	832 (41,98)	681 (34,36)	113 (5,70)	14 (0,71)	4 (0,20)	1 (0,05)	1 (0,05)	1.982 (100,00)

Degli undici valori mancanti della variabile in esame, cinque assumono valore “2” (45,5 per cento) e i restanti sei prendono valore 3 (55,5 per cento). Di questi casi mostriamo il layout Sas della distribuzione doppia relativa ai caratteri “*Tabella Inps (DC_23)*” e “*Numero di familiari (DC_24)*” successiva all’imputazione:

N.° FAMILIARI DOPO L'IMPUTAZIONE (DC_24)	Tabella Inps (DC_23)		Totale
	11	21A	
2	0 (0,00)	5 (45,45)	5 (45,45)
3	6 (54,55)	0 (0,00)	6 (54,55)
Totale	6 (54,55)	5 (45,45)	11 (100,00)

Da un’ulteriore analisi si osserva come la distribuzione condizionata del *numero di familiari*, data la categoria “21A” della tabella Inps, segnali una concentrazione del 97,3 per cento dei casi sul valore modale “2”. Mentre la stessa distribuzione condizionata, fissata la categoria “11” della tabella Inps, indica come valore modale il termine “3” (48,6 per cento).

Questi risultati empirici confermano la validità della strategia di imputazione messa in atto in relazione all'ultima coordinata esaminata.

In sintesi, al termine della procedura di correzione ed imputazione impiegata nell'indagine Eu-Silc 2005, risultavano 1.981 potenziali percettori di Anf, cioè individui con coordinate complete tali da generare un importo positivo dell'Anf, mentre in un solo caso non è stato possibile procedere all'imputazione. Dei potenziali percettori di Anf, 1.979 possedevano almeno una mensilità retribuita nell'anno di riferimento, per cui venivano assegnato loro un importo dell'Anf. La procedura utilizzata poggia su una serie di ipotesi tra cui l'invarianza delle coordinate durante l'anno solare ed il possesso dei requisiti per l'accesso alla prestazione durante tutto l'anno di riferimento. Per migliorare il processo di imputazione/correzione degli Anf si può fare ricorso a tecniche di imputazione tramite donatore (Scia), oppure effettuare correzioni/imputazioni di tipo longitudinale per le famiglie che vengono reintervistate negli anni successivi di indagine (sino ad un massimo di quattro anni). L'informazione relativa alle coordinate familiari è presente negli archivi fiscali Cud/770 sino all'anno 2005 (periodo d'imposta 2004), negli anni a seguire la stessa informazione viene raccolta dall'Inps attraverso le denunce *E-mens*.

3.4.3 - Le liquidazioni di fine rapporto di fonte Cud/770

L'unica fonte amministrativa in grado di rilevare in maniera esaustiva le liquidazioni di fine rapporto, con un livello di dettaglio micro (percettore), è quella del Cud/770. In particolare nel Quadro B – *Sezione trattamento di fine rapporto* – del modello Cud/770_2005 sono presenti al punto 71 gli importi, soggetti a tassazione separata, relativi a:

- indennità di fine rapporto e indennità equipollenti;
- altre indennità soggette a tassazione separata;
- anticipazioni e acconti sul Tfr e su altre indennità e somme;
- prestazioni pensionistiche complementari erogate in forma di capitale;
- indennità di fine mandato corrisposte a collaboratori coordinati e continuativi ove sussistano le condizioni;
- riscatto anche parziale dal fondo previdenziale, con almeno 8 anni di contributi, per eventuali spese sanitarie, terapie ed interventi straordinari, per acquisto della prima casa per sé o per i figli, per la realizzazione di interventi di manutenzione ordinaria e straordinaria, di restauro o ristrutturazione, dell'abitazione principale;
- riscatto della posizione individuale esercitato per effetto del pensionamento o per la cessazione del rapporto di lavoro, per mobilità o per altre cause indipendenti dalla volontà delle parti;
- anticipazioni di pensioni integrative.

Figura 3.11 – Sezione: Trattamento di fine rapporto, indennità equipollenti ..., campi: 71-83 - Parte B - Cud/770_2005

Codice fiscale del percipiente	Progressivo Comunicazione				
TRATTAMENTO DI FINE RAPPORTO, INDENNITÀ EQUIPOLLENTI, ALTRE INDENNITÀ E PRESTAZIONI IN FORMA DI CAPITALE	TRATTAMENTO DI FINE RAPPORTO, ALTRE INDENNITÀ E SOMME SOGGETTE A TASSAZIONE SEPARATA				
	Indennità, acconti, anticipazioni e somme erogate nell'anno				
	Acconti ed anticipazioni erogati in anni precedenti				
	Ritenute operate nell'anno				
	Ritenute operate in anni precedenti				
	Quota spettante per le indennità erogate ai sensi dell'art. 21.22 c.c.				
71	72	73	74	75	%
DATI RELATIVI AL RAPPORTO DI LAVORO					
Data di inizio rapporto di lavoro		Data di cessazione rapporto di lavoro		Periodo di commisuraz.	
76	giorno	77	giorno	78	anni
	mesi		mese		mesi
	anno		anno	Periodo convenzionale	
Quota eccedente		Giorni di sospensione di lavoro		Tempo deter.	
80		81		82	
				83	
				Codice fiscale del sostituto	

Come si può osservare non tutte le voci incluse al punto 71 fanno parte della nozione di reddito secondo le indicazioni fornite da Eurostat per la definizione delle variabili target di Eu-Silc. Ad esempio, le prestazioni pensionistiche complementari erogate sotto forma di capitale non fanno parte delle definizioni di reddito, come pure il riscatto dal fondo previdenziale per le spese connesse all'abitazione principale e le anticipazioni di pensioni integrative. Per escludere tali componenti economiche si è proceduto a identificare le relative voci nel modello fiscale. In particolare, le liquidazioni in forma di capitale da forme pensionistiche sono indicate in due apposite sezioni del quadro B, distinte secondo il periodo di maturazione: fino al 31 dicembre 2000 e dal 1° gennaio 2001 in poi (figura 3.12).

Figura 3.12 – Sezione: Trattamento di fine rapporto, indennità equipollenti ..., campi: 108-116 - Parte B - Cud/770_2005

PRESTAZIONI IN FORMA DI CAPITALE MATURATE AL 31/12/2000 EROGATE DA FORME PENSIONISTICHE					
Prestazione definitiva 108	Prestazione parziale 109	Erede 110	TFR destinato al fondo pensione 111	Deduzione complessivamente spettante 112	Percentuale 113 %
		<input type="checkbox"/>			
PRESTAZIONI IN FORMA DI CAPITALE MATURATE DAL 1/1/2001 EROGATE DA FORME PENSIONISTICHE					
Prestazione definitiva 114	Prestazione parziale 115	Redditi già assoggettati ad imposta 116			

Nel punto 108 è indicato l'ammontare della prestazione in esame erogata nell'anno 2004 a titolo definitivo, comprese le anticipazioni, riferite alla data di maturazione del 31 dicembre 2000. Nella fattispecie sono comprese le erogazioni in forma di capitale a favore dei soggetti iscritti a forme pensionistiche complementari alla data del 28 aprile 1993, i.c.d. "vecchi iscritti a vecchi fondi". L'ammontare delle erogazioni sono considerate al netto del rendimento finanziario soggetto all'imposta del 12,5 per cento. Il punto 109 riporta la corrispettiva erogazione parziale a titolo di anticipazione della prestazione, relativa ad importi maturati alla data del 31 dicembre 2000. Nel punto 111 si trascrive la quota di Tfr maturato alla data del 31 dicembre 2000 destinato a forme pensionistiche complementari. I punti 112 e 113 includono elementi per il calcolo delle deduzioni spettanti sul Tfr destinato a forme di previdenza complementare. Nei punti da 114 a 116 sono indicati i valori delle prestazioni in forma di capitale erogate nel 2004 dalle forme pensionistiche complementari e individuali, riferibili all'ammontare maturato dal 1° gennaio 2001 e soggette alle disposizioni dell'art. 20 del Tuir. In particolare, il punto 114 va compilato nell'ipotesi di erogazione a titolo definitivo della prestazione in forma di capitale; l'importo deve essere, comprensivo delle anticipazioni e dei riscatti parziali, esercitati ai sensi dell'art. 10, comma 1-bis, del d.lgs. n. 124 del 1993. Il punto 115 va compilato soltanto nelle ipotesi di erogazione parziale della prestazione maturata, a titolo di anticipazione o riscatto parziale imputabile agli importi maturati a decorrere dal 1° gennaio 2001. L'importo deve essere comprensivo del risultato di gestione al netto della relativa imposta sostitutiva dell'11 per cento corrisposta e ridotto dell'importo relativo alle somme non dedotte (contributi e premi). Al punto 116 va indicato l'importo complessivo dei redditi già assoggettati ad imposta riferibile all'importo di punto 114 (risultato di gestione assoggettato ad imposta sostitutiva nella misura dell'11 per cento e contributi e premi non dedotti) ovvero all'importo di punto 115 (contributi e premi non dedotti).

La somma degli importi relativi ai punti 108, 109, 114, 115 e 116 fornisce il dato relativo alle prestazioni in forma di capitale ricevute nel 2004 da forme pensionistiche complementare a titolo parziale e/o definitivo al lordo della tassazione. Una volta eliminati gli elementi che non concorrono alla formazione del reddito, il passo successivo consiste nell'assegnare la funzione di protezione sociale alla corrispondente prestazione o una sua componente (trattamento). Le

liquidazioni di fine rapporto come si è accennato in precedenza possono essere attribuite a tre distinte funzioni di protezione sociale o variabili target: *PY100N*, *PY090N* e *PY130N*. La classificazione della prestazione in una delle tre funzioni dipende da specifiche condizioni o requisiti. Si attribuisce la funzione superstiti quando il percettore della liquidazione è un erede del dante causa titolare della prestazione. La funzione disoccupazione si associa alle indennità di fine rapporto corrisposte a seguito di licenziamento, mentre la funzione vecchiaia concerne le liquidazioni di fine rapporto concesse ai lavoratori dipendenti in quiescenza. La presenza di un'indennità a favore di eredi è segnalata al punto 75, dove si riporta la quota di reddito spettante ai discendenti ai sensi dell'art. 2122 del codice civile. In tale circostanza l'importo netto della liquidazione effettivamente percepito dall'erede è dato dal prodotto dell'aliquota esposta al punto 75 e il saldo tra ammontare lordo riportato al punto 71 e le relative ritenute fiscali indicate al punto 70. In fase di controllo e correzione dell'importo delle liquidazioni si è stabilito di confrontare il totale generale riportato al punto 71 (importo lordo) con la somma delle singole componenti lorde che determinano l'ammontare complessivo delle liquidazioni erogate a favore di un determinato soggetto. In pratica, si è considerato lo scostamento dell'importo iscritto al punto 71 con la somma dei punti: DB_84, DB_86, DB_89, DB_91, DB_99, DB_101, DB_104, DB_108, DB_109, DB_114, DB_115, DB_116 (*SCOMPO_Tfr*). Alla luce di tale confronto si è potuto verificare la presenza di errori di doppio conteggi nella somma delle componenti Tfr. Inoltre si è potuto osservare come in presenza di importi significativi nell'incentivo all'esodo (DB_104) i due aggregati (DB_71 e SCOMPO) in diversi casi non coincidono. Tale differenza è imputabile al criterio di calcolo stabilito dalla normativa fiscale (art. 19, comma 4-bis, Tuir) che prevede l'abbattimento del 50 per cento dell'imponibile relativo a questa posta per i soggetti abbiano superato l'età di 50 anni se donne, ovvero 55 anni se uomini. Pertanto nel punto 71 tale importo compare per intero mentre nel punto DB_104 esso risulta ridotto al 50 per cento.

Figura 3.13 – Sezione: Trattamento di fine rapporto, indennità equipollenti ..., campi: 84-107 – Parte B Cud/770_2005

INDENNITÀ EQUIPOLLENTI ED ALTRE INDENNITÀ CORRISPOSTE NEL 2004				
Indennità equipollenti		Altre indennità		
Ammontare netto 84	Titolo 85	Ammontare netto 86	Titolo 87	
TFR ED ALTRE INDENNITÀ MATURATE AL 31/12/2000				
TFR maturato 88	TFR		Altre indennità	
TFR destinato al fondo pensione 93	Ammontare erogato 89	Titolo 90	Ammontare netto 91	Titolo 92
	Deduzione complessivamente spettante 94	Percentuale 95	Tempo parziale 96 anni mesi 97 %	
TFR ED ALTRE INDENNITÀ MATURATE DAL 1/1/2001				
TFR maturato 98	TFR		Altre indennità	
TFR destinato al fondo pensione 103	Ammontare erogato 99	Titolo 100	Ammontare netto 101	Titolo 102
	Incentivo all'esodo		Tempo parziale	
	Importo 104	Titolo 105	106 anni mesi 107 %	

Al fine di evitare l'inclusione di doppi conteggi nella somma delle componenti del Tfr, si è apportato un correttivo alla posta DB_84 indicante l'ammontare delle indennità equipollenti corrisposte nel 2004. Infatti, in presenza di acconti ed anticipi erogati in anni precedenti (evidenziati nella posta DB_72), la componente di liquidazione indicata in DB_84 può includere eventuali acconti e anticipazioni erogati nel periodo compreso fra il 1974 ed il 2003 (DB_117). A tal fine si è osservato che quando sono soddisfatte entrambe le seguenti condizioni: 1) "DB_72>0"; 2) "DB_84>DB_117>0"; allora alla somma delle componenti del Tfr (SCOMPO_Tfr) deve essere sottratto il valore della variabile DB_117.

Per impostare il valore finale delle liquidazione di fine rapporto si effettua un confronto tra il valore del saldo delle liquidazioni di cui al punto DB_71, in assenza/presenza delle due componenti aggiuntive DB_117 e DB_104, e la sommatoria delle singole voci che compongono la medesima liquidazione (SCOMPO_Tfr). Il risultato finale di questo confronto e delle relative azioni intraprese è illustrato in tavola 3.9. Con riferimento all'ultima casistica è possibile utilizzare una procedura di calcolo alternativa al "massimo dei due valori" scegliendo quel valore che più si approssima al predittore lineare calcolato sulla retta di regressione che lega ciascuna delle due variabili al regressore dell'imposta pagata sul Tfr (cfr. Consolini, 2004).

Tavola 3.9 – Relazioni di compatibilità/incompatibilità tra DB_71 (saldo) e SCOMP_Tfr (somma voci)

RELAZIONI DI COMPABILITÀ/ INCOMPABILITÀ TRA: DB_71 (SALDO FINALE TFR E ASSIMILATI) SCOMPO_TFR (SOMMA DELLE SINGOLE COMPONENTI TFR)	Azioni intraprese	Campione teorico Eu-Silc 2005	
		Frequenza	Percentuale
DB_71-SCOMPO_Tfr =<2	Tfr = DB_71	5.395	88,11
DB_117>0 & DB_71-SCOMPO_Tfr-DB_117 =<2	Tfr = DB_71-DB117	293	4,79
DB_104>0 & DB_71-SCOMPO_Tfr-DB_104 =<2	Tfr = DB_71-DB004	26	0,42
DB_71=0 & SCOMPO_Tfr >2	Tfr =SCOMPO_Tf	22	0,36
DB_71>2 & SCOMPO_Tfr =0	Tfr = DB_71	153	2,5
DB_71-SCOMPO_Tfr >2 & (DB_71-SCOMPO_Tfr /DB_71)=<15%	Tfr= max(DB_71,SCOMPO	43	0,7
DB_71-SCOMPO_Tfr >2 & (DB_71-SCOMPO_Tfr /DB_71)>15% ^(a)	Tfr= max(DB_71,SCOMPO	191	3,12
Totale		6.123	100,00

3.4.4 - I trattamenti non pensionistici ad erogazione diretta Inps di fonte Cud/770

I trattamenti non pensionistici (TMNP) ad erogazione diretta Inps si configurano come un insieme di trasferimenti in denaro, da parte dell'ente erogatore Inps direttamente a favore dei beneficiari (e quindi non tramite il datore di lavoro), comprendente le seguenti fattispecie³²: 1) indennità di malattia; 2) indennità di malattia speciale Tbc; 3) liquidazioni di fine rapporto; 4) indennità di maternità e cura dei figli; 5) assegni al nucleo familiare (vecchio e nuovo tipo); 6) assegni ai nuclei con tre figli minori; 7) indennità di disoccupazione; 8) indennità di mobilità e 9) integrazioni salariali. In particolare, si osserva che la fonte Cud/770 consente di rilevare esclusivamente il dato sul totale degli importi annui erogati per un sottoinsieme di queste tipologie, segnatamente le prestazioni previdenziali³³ a regime di tassazione normale (tipologie 1, 2, 4, 5, 7, 8, e 9). L'identificazione di questa tipologia di redditi avviene tramite opportune variabili spia presenti della stessa fonte (figura 3.14).

³² Per una trattazione più estesa dell'argomento si rinvia a P. Consolini (2000), P. Consolini e R. De Carli (2002).

³³ Il termine "prestazione previdenziale" nel linguaggio parlato è spesso utilizzato impropriamente per indicare i soli trattamenti pensionistici; tuttavia se ci riferisce ai sistemi di classificazioni adottate in questo ambito (SEC) tale termine, in realtà, comprende tutte quelle prestazioni il cui l'accesso è vincolato al versamento di contributi previdenziali da parte del lavoratore o del datore.

Figura 3.14 – Le variabili identificative dei TMNP a tassazione normale del modello Cud/770_2005

Le variabili che all'interno del Cud/770 qualificano un reddito a tassazione normale come *TMNP* (per il sottoinsieme di trattamenti cui sopra) sono così costituite: *a*) codice fiscale del sostituto d'imposta (in questo caso si tratta del codice Inps in qualità di ente erogatore); *b*) numero di giorni per cui spettano le deduzioni di cui all'art. 11 del Tuir (punto 6, Parte B, Cud); *c*) matricola aziendale riportata nel quadro C previdenziale Inps, di cui al punto 4 Parte C (deve essere assente ovvero *missing*) e mesi retribuiti nell'anno per il lavoratore, di cui ai punti 14 e 15 Parte C (assenti o *missing*). Vi è da rilevare che il punto *c*) consente di discriminare i percettori di *TMNP* dai percettori di reddito da lavoro dipendente aventi come datore di lavoro l'Inps. Il dato sull'importo annuo per il complesso dei trattamenti non pensionistici rilevabile da Cud/770 da solo non è molto utile ai fini dell'analisi in quanto a tale valore occorre associare la corrispondente tipologia di prestazione ovvero variabile target Eu-Silc. Dal momento che le fonti amministrative sin qui utilizzate non consentono di attribuire gli importi *TMNP* di fonte Cud alle rispettive variabili target Eu-Silc, si è deciso di utilizzare a tal riguardo le informazioni contenute nel questionario dell'indagine Eu-Silc. In altri termini, la presenza di segnali nel questionario individuale dei redditi che denotano un periodo di malattia o di maternità, di disoccupazione, ecc. del soggetto intervistato forniscono il quadro di riferimento per la corretta attribuzione (si veda successivo capitolo 4).

3.5 La costruzione del reddito da lavoro autonomo

3.5.1 - Premessa

La ricostruzione del reddito da lavoro autonomo a partire dalle fonti fiscali costituisce un processo che presenta elevati livelli di complessità. Il reddito in questione è, infatti, associabile a diverse figure professionali che trovano differente collocazione nell'ambito della disciplina giuridico-fiscale. Così ad esempio, i redditi percepiti dai soci di cooperative di produzione e lavoro, delle cooperative di servizi, delle cooperative agricole e di prima trasformazione, sono classificati come componenti autonome ma vengono assimilati fiscalmente ai redditi dal lavoro dipendente e pertanto

registrati nei modelli Cud/770. Similmente, i redditi derivanti dalle collaborazioni coordinate e continuative pur essendo classificati ai fini statistici come redditi da lavoro autonomo, sono anch'essi assimilati fiscalmente ai redditi da lavoro dipendente e dunque inquadrati nelle dichiarazioni Cud/770. Altre voci economiche, quali i proventi derivanti dall'utilizzazione economica di opere di ingegno, sono poi rilevabili dal modello 730 e dal modello Upf. Infine, tutti i guadagni derivanti da attività professionale e di impresa, in qualità di titolare o partecipante (coadiuvante/socio), sono esclusivamente registrate nell'archivio delle dichiarazioni Upf. Di seguito si illustrano le principali caratteristiche della variabile target sul reddito da lavoro autonomo e le relative fonti del dato.

Tavola 3.10 – Redditi da lavoro autonomo, tipologia di percettori, voci economiche e fonti del dato

Variabile target Eu-Silc	Definizione	Tipologia di percettori	Voci economiche	Fonti del dato (anno 2005)			
				Agenzia delle entrate			ISTAT Indagine Eu-Silc
				Modello Cud/770	Modello 730	Modello Upf	
PY050N Guadagni o perdite in denaro derivanti dal lavoro autonomo	Insieme dei guadagni ricevuti, durante il periodo di riferimento dei redditi, come risultato dell'attività corrente o pregressa di lavoro autonomo	Collaboratori coordinati e continuativi	Remunerazione della attività prestate in collaborazioni a progetto e coordinate e continuative	Parte B e C	Quadro C ^(a)	Quadro RC ^(a)	✓
		Soci di cooperative	Remunerazione dell'attività lavorativa prestata in cooperative di produzione e lavoro, cooperative di servizi, cooperative agricole e di prima trasformazione	Parte B	Quadro C ^(a)	Quadro RC ^(a)	
		Sacerdoti della Chiesa cattolica e ministri del culto	Remunerazione del servizio prestato nelle diocesi e congregate corrisposte sui Fondi del culto, beneficenza e religione	Parte B	Quadro C ^(a)	Quadro RC ^(a)	
		Personale medico ospedaliero, docenti univers. e operat. sanit.	Remunerazione dell'attività medico-sanitaria intramuraria	Parte B	Quadro C ^{(a)(b)}	Quadro RC ^{(a)(b)}	
		Agenti e rappresentanti di commercio che non possiedono partita iva	Provvigioni inerenti a rapporti di commissione, agenzia, mediazione, rappresentanza di commercio e proc. d'affari	Mod.ord. 770 ^(b) : Sezione autonomi			
		Utilizzatori delle proprie opere d'ingegno	Utilizzazione ai fini economici di opere d'ingegno e royalties	Mod.ord. 770: Sezione autonomi	Quadro D	Quadro RL	
		Associati in partecipazione (artt. 2549-2554 c.c.)	Partecipazione agli utili di impresa come corrispettivo dell'apporto esclusivamente di lavoro	(c)	Quadro D	Quadro RL	
		Lavoratori autonomi occasionali e venditori a domicilio (art. 2222 c.c.)	Redditi diversi (art. 67, comma 1, lett I), Tuir)	(c)	Quadro D	Quadro RL	
		Liberi professionisti e artisti	Proventi derivanti dall'esercizio della libera professione e dell'attività artistica			Quadro RE	
		Titolari d'impresa	Remunerazione dell'attività imprenditoriale			Quadro RF/RG	
		Coadiuvati di impresa familiare o coniugale, soci di società di persone,	Remunerazione dell'attività di collaborazione nella conduzione dell'impresa familiare o nella società di persone			Quadro RH	

(a) la voce corrispondente è rilevata solo se il percettore presenta una dichiarazione 730 o Upf.

(b) la rilevazione di questa voce è ancora in fase di progettuale.

(c) se il compenso supera la soglia di 5 mila euro allora sorge l'obbligo da parte del committente di dichiarare i dati retributivi ai fini previdenziali per la Gestione separata Inps, Sezione 4 Parte C del Cud/770.

3.5.2 - Il reddito dei collaboratori coordinati e continuativi

La ricostruzione del reddito dei collaboratori coordinati e continuativi è realizzata a partire dalle informazioni contenute nella Sezione 4 – Parte C – del modello Cud/770 anno 2005. Questa sezione è riservata alla comunicazione, da parte dei committenti, dei dati relativi al periodo di attività e ai redditi e conferiti ai loro collaboratori coordinati e continuativi iscritti alla Gestione separata dell'Inps³⁴. In particolare all'interno della sezione figurano: la data di inizio rapporto di collaborazione (punto 56), la data di fine rapporto (punto 57), i compensi corrisposti ovvero l'imponibile contributivo (punto 58), i contributi dovuti alla suddetta gestione previdenziale Inps da parte del collaboratore in base al disposizioni vigenti nell'anno d'imposta (punto 59), i contributi trattenuti al collaboratore per la quota a suo carico, pari ad un terzo dei contributi dovuti (punto 60) e i contributi effettivamente versati dal committente all'Inps (punto 61). Si dimostra che l'imponibile fiscale dei redditi dei collaboratori si ottiene come differenza tra l'imponibile contributivo e i contributi trattenuti al collaboratore per la quota a suo carico.

Figura 3.15 – Sezione 4: Collaborazioni coordinate e continuative – Parte C B - Cud/770_2005

SEZIONE 4 COLLAB. COORDINATE E CONTINUATIVE		Data inizio		Data fine		Compensi corrisposti	Contributi dovuti	Contributi trattenuti	Contributi versati
56	giorno	56	giorno	57	giorno	58	59	60	61
	mese	mese	mese	mese	mese				
	anno		anno		anno				

La determinazione del *reddito netto* a favore dei collaboratori coordinati e continuativi è relativamente semplice quando sono assenti altre tipologie di reddito, per cui si esso si ottiene come differenza tra l'imponibile fiscale³⁵ e le ritenute fiscali esposte ai punti 12, 13 e 14 della parte B del Cud/770. Viceversa, il calcolo risulta più complesso in presenza di altre voci economiche. In quest'ultima circostanza il reddito netto si determina in due fasi. Nella prima fase si scinde l'imponibile fiscale generale esposto al punto 1 della Parte B nelle seguenti due componenti: *i*) l'imponibile riferito ai collaboratori coordinati continuativi (ricavato dalla differenza tra il punto 58 e 60 della Parte C del Cud) e *ii*) il residuo ascrivibile ad altre forme di reddito. Nella seconda fase, il reddito netto si ottiene come differenza tra l'imponibile riferito al punto *i*) e la corrispondente ritenuta fiscale, ricavata pro-quota in funzione delle due componenti lorde di reddito (rispettivamente dei collaboratori e del residuo). In sintesi, la procedura di calcolo può essere illustrata con le seguenti espressioni:

$$YF_{cococo} = YC_{cococo} - CS_{cococo}$$

$$YF_{RES} = (GY - SC) - YF_{cococo}$$

$$NY_{cococo} = TY \cdot (YF_{cococo} / GY)$$

dove rispettivamente YF denota l'imponibile fiscale, YC l'imponibile contributivo, GY il reddito lordo (al netto dei contributi sociali a carico del datore di lavoro o committente), SC i contributi sociali a carico del lavoratore, TY le ritenute fiscali (Irpef e addizionali regionali e comunali) e NY il reddito netto.

³⁴ Si rammenta che i collaboratori coordinati e continuativi rappresentano un sottogruppo degli iscritti alla gestione separata dell'Inps. Infatti, tra questi ultimi fanno parte anche i liberi professionisti senza una propria Cassa pensioni; gli spedizionieri doganali; i venditori porta a porta (addetti alla vendita a domicilio per importi annui superiori 5 mila euro); i lavoratori autonomi occasionali (importo annuo superiore ai 5 mila euro); gli associati in partecipazione.

³⁵ In tale circostanza l'imponibile fiscale rilevabile al punto 1 della parte B del Cud/770 coincide, in valore, con la differenza tra il punto 58 e il punto 60 della Parte C del Cud.

3.5.3 - *Il reddito dei soci di cooperative e dei ministri del culto*

Per individuare le componenti di reddito autonomo riferibili ai soci di cooperative e ai sacerdoti della Chiesa cattolica, o più in generale ai ministri del culto, che nel Cud compaiono come redditi assimilati al lavoro dipendente, si ricorre all'analisi della variabile "categorie particolari". Tale variabile, presente nella Parte A del modello Cud/770 al punto 9, identifica con un codice l'appartenenza ad una specifica tipologia di lavoratore o percettore. Il codice corrispondente ai soci di cooperativa la variabile è dato dalla lettera "M", mentre quello riferito ai sacerdoti è dato dalla lettera "Q". In presenza di questi due codici, si attua una riclassificazione dei redditi presenti nella sezione dati fiscali della Parte B, assegnati in precedenza alla tipologia dei lavoratori dipendenti, verso la componente autonoma. Il calcolo del reddito netto assegnato a queste due tipologie di percettori segue le stesse regole di calcolo del reddito del lavoro dipendente da fonte Cud/770 esplicitate nel paragrafo 3.2.1.

3.5.4 - *Il reddito derivante dall'attività intramuraria/intramoenia di medici e operatori sanitari*³⁶

I compensi percepiti dai medici del Ssn in relazione all'attività intramoenia costituiscono redditi assimilati ai redditi di lavoro dipendente a norma dell'art. 47 comma 1 lettera e) del D.p.r. 917/86. Su tali redditi non competono pertanto le detrazioni per i redditi di lavoro dipendente, di cui all'art. 14 dello stesso D.p.r., e le detrazioni per i carichi familiari, mentre per converso si applicano le ritenute d'acconto secondo le aliquote Irpef e le deduzione previste all'art. 11 comma 1 del Tuir. Ai fini della determinazione dei redditi assimilati a quelli di lavoro dipendente si applicano le disposizioni dell'art. 48bis del Tuir. Il reddito imponibile è dato da tutti gli emolumenti corrisposti nel periodo di imposta, secondo il principio di cassa, in denaro o in natura, al netto delle trattenute previdenziali. Ai fini della loro rilevazione tramite fonte amministrativa, si osserva come all'interno del Cud essi siano registrati al punto 2 della Parte B, diversamente dai redditi dei soci delle cooperative e dei ministri del culto che sono inclusi invece al punto 1. In analogia con il criterio di identificazione applicato ai redditi esposti nel precedente paragrafo, l'individuazione di questa specifica tipologia di reddito si basa sulla variabile "categorie particolari" che assume in questo caso il valore "R". Come si è accennato al paragrafo 3.2.1, la rilevazione di questa tipologia di redditi è ancora in fase di progettazione, ma verrà implementata ed avviata a regime a partire dall'indagine Eu-Silc 2006.

3.5.5 - *Le provvigioni inerenti a rapporti di commissione, agenzia, mediazione, rappresentanza di commercio e procacciamento d'affari*

In base alla definizione delle variabili target di reddito Eu-Silc, le provvigioni in questione rientrano a pieno titolo nel reddito da lavoro autonomo. Data la peculiarità di queste tipologie di reddito, la normativa fiscale ha disposto la loro collocazione in un apposito modello di dichiarazione: modello semplificato 770 – Sezione Comunicazione Dati Certificazioni Lavoro Autonomo, Provvigioni e Redditi Diversi. Occorre tuttavia precisare che il citato modello include anche altre componenti di reddito da lavoro autonomo che sono comunque codificate

³⁶ L'attività intramoenia è svolta dai soggetti che rispondono a determinati requisiti soggettivi: 1) personale dipendente del servizio sanitario nazionale quale medici chirurghi, odontoiatri, veterinari ed altre professionalità che ricoprono l'incarico di dirigenti del ruolo sanitario quali biologi, farmacisti, fisici, chimici e psicologi; 2) docenti universitari e ricercatori che effettuano attività assistenziale presso cliniche e istituti universitari di ricovero e cura anche se gestiti direttamente dalle università; 3) personale laureato e medico di ruolo in servizio nelle strutture delle facoltà di medicina e chirurgia delle aree tecnico-scientifica e socio-sanitaria; 4) dipendenti degli istituti di ricovero e cura a carattere scientifico con personalità giuridica di diritto privato, dipendenti degli enti ed istituti di cui all'art. 4 comma 12 del d.lgs. 502/1992, nonché dipendenti delle istituzioni pubbliche di assistenza e beneficenza (Ipab) che svolgono attività sanitaria e degli enti pubblici che già applicano al loro personale l'istituto dell'attività libero-professionale intramuraria della dirigenza del servizio sanitario, a condizione che i predetti enti ed istituti abbiano adeguato i propri ordinamenti ai principi di cui all'art. 1 commi da 5 a 19 della L. 662/96 e al D.M. Sanità 31/7/1997; 5) personale autorizzato ad operare presso studi privati o strutture non accreditate, purché l'attività rispetti i criteri fissati dal D.M. Sanità 31/7/1997 e successivamente negli atti adottati dai direttori generali delle strutture sanitarie.

all'interno del modello da un'apposita variabile. Tra queste altre componenti figurano le prestazioni da lavoro rientranti nell'esercizio di arte o professione abituale, remunerazione di brevetti e opere d'ingegno, utili agli associati in partecipazione, prestazioni corrisposte all'attività degli sportivi, eccetera. Per selezionare le voci economiche che corrispondono identicamente alle provvigioni è necessario dunque filtrare le dichiarazioni dei redditi secondo i codici "Q, R, S, T, U"³⁷ associati alla variabile "Causale" (punto 16). Il reddito netto di queste provvigioni si ricava dalla differenza tra l'ammontare lordo corrisposto, di cui al punto 29, e le ritenute a titolo d'acconto (punto 23) o a titolo di imposta (punto 24) nonché le rispettive addizionali regionali a titolo d'acconto (punto 26) o d'imposta (punto 27).

Figura 3.16 – Certificazioni lavoro autonomo, provvigioni e redditi diversi – Modello semplificato 770_2005

70 SEMPLIFICATO 2006
COMUNICAZIONE DATI CERTIFICAZIONI LAVORO AUTONOMO, PROVVIGIONI E REDDITI DIVERSI
EURO

Progressivo Comunicazione
 CODICE FISCALE

agenzia entrate

CODICE FISCALE DEL SOSTITUTO D'IMPOSTA
 (da compilare solo in caso di operazioni straordinarie e successioni)

DATI RELATIVI ALLE SOMME EROGATE

Causale	Anno	Anticipazione	Ammontare lordo corrisposto	Somme non soggette a ritenuta per regime convenzionale	Altre somme non soggette a ritenuta
16	17	18	19	20	21
		<input type="checkbox"/>			
22	Imponibile				
+					
-					
25	Ritenute sospese				
23	Ritenute a titolo d'acconto				
24	Ritenute a titolo d'imposta				
26	Addizionale Regionale a titolo d'acconto				
27	Addizionale Regionale a titolo d'imposta				
28	Addizionale Regionale sospesa				
29	Imponibile anni precedenti				
30	Ritenute operate anni precedenti				
31	Contributi previdenziali a carico del soggetto erogante				
32	Contributi previdenziali a carico del percipiente				
33	Spese rimborsate				
34	Ritenute rimborsate				

3.5.6 - Il reddito derivante dall'utilizzazione economica di opere d'ingegno

Il compenso derivante dall'utilizzo economico delle opere d'ingegno, di invenzioni industriali ecc.³⁸, conseguiti dall'autore o dall'inventore si configurano come reddito da lavoro autonomo tanto per il fisco italiano quanto per Eurostat nella definizione della corrispondente variabile target Eu-Silc. Le fonti fiscali che rilevano il dato sui compensi relativi ai diritti d'autore sono, in alternativa, il modello 730 ed il modello Upf. In relazione al primo modello di dichiarazione, i redditi in esame figurano nel Quadro D al punto D3 "Compensi di lavoro non

³⁷ In particolare, il codice "Q" definisce le provvigioni corrisposte ad agente o rappresentante monomandatario; il codice "R" corrisponde alle provvigioni di agenti o rappresentati plurimandatari; il codice "S" comprende le provvigioni concesse ai commissionari; il codice "T" attiene alle provvigioni versate ai mediatori; il codice "U" include le provvigioni corrisposte ai procuratori d'affari.

³⁸ In pratica si tratta di brevetti, disegni e modelli ornamentali, know-how, libri e articoli per riviste e giornali e così via.

derivanti da attività professionale”, insieme ad altre tipologie di reddito (figura 3.17). Per la selezione della corrispondente tipologia di reddito si deve analizzare il codice numerico riportato in colonna 1 (tipo di reddito): il valore “1” denota per l’appunto i proventi derivanti dall’utilizzazione economica delle opere d’ingegno. Per altro verso, il reddito netto si ottiene come la differenza fra i redditi indicati colonna 2 (al lordo della deduzione forfetaria del 25 per cento) e le rispettive ritenute esposte in colonna 4 (ritenute d’acconto subite).

Figura 3.17 – Quadro D: Altri redditi – Mod. 730_2005

QUADRO D ALTRI REDDITI				
	TIPO DI REDDITO	REDDITI		RITENUTE
D1	1	2		4
D1	Utili e altri proventi equiparati	,00		,00
D2	Altri redditi di capitale	,00		,00
D3	Compensi di lavoro autonomo non derivanti da attività professionale	,00	SPESE	,00
D4	Redditi diversi	,00		,00

Per quanto riguarda il modello Unico persone fisiche il dato sui compensi per i diritti d’autore è registrato, al lordo delle deduzioni forfetarie, al rigo RL24 della Sezione III “Altri redditi da lavoro autonomo” del Quadro RL - Altri redditi. Il reddito netto di questa fonte si ottiene, anche in questo caso, sottraendo all’importo lordo le relative ritenute fiscali indicate nel rigo RL30. A differenza del modello 730, nel modello Upf non consente di attribuire distintamente le ritenute d’acconto ai percettori di questi proventi, ma utilizza lo stesso variabile per includere anche le ritenute di altre tipologie di reddito: compensi dei segretari comunali per attività di levata protesti e redditi degli associati in partecipazione.

Figura 3.18 – Sezione III: Altri redditi da lavoro autonomo – Quadro RL - Mod.Upf_2005

SEZIONE III	RL24	Proventi lordi per l'utilizzazione economica di opere dell'ingegno, di brevetti industriali, ecc. percepiti dall'autore o inventore	,00
Altri redditi da lavoro autonomo	RL25	Compensi derivanti dall'attività di levata dei protesti esercitata dai segretari comunali	,00
	RL26	Redditi derivanti dai contratti di associazione in partecipazione se l'apporto è costituito esclusivamente da prestazioni di lavoro e utili spettanti ai promotori e ai soci fondatori di società per azioni, in accomandita per azioni e a responsabilità limitata	,00
	RL27	Totale compensi, proventi e redditi (sommare gli importi da rigo RL24 a RL26)	,00
	RL28	Deduzioni forfetarie delle spese di produzione dei compensi e dei proventi di cui ai righe RL24 e RL25	,00
	RL29	Totale netto compensi, proventi e redditi (RL27 - RL28; sommare l'importo agli altri redditi Irpef e riportare il totale al rigo RN1 col. 4)	,00
	RL30	Ritenute d'acconto (sommare tale importo alle altre ritenute e riportare il totale al rigo RN23 col. 2)	,00

3.5.7 - Il reddito degli associati in partecipazione

Con il contratto di associazione in partecipazione, l’associante (o committente) conferisce all’associato una partecipazione agli utili della sua impresa, per uno o più affari, a fronte di un determinato apporto dell’associato. Nel contratto di associazione in partecipazione l’apporto dell’associato può essere costituito da un determinato capitale, da capitale e lavoro (c.d. apporto misto) oppure soltanto da lavoro. Il reddito degli associati in partecipazione, quando conferisce esclusivamente un apporto di lavoro, è inquadrato sia dal punto fiscale che previdenziale come redditi da lavoro autonomo. Dal punto di vista delle fonti di rilevazione del dato, si osserva che tanto il modello 730 quanto il modello Unico persone fisiche prevedono un’apposita sezione che accoglie la dichiarazione di questa tipologia di reddito. Nel caso del primo modello di dichiarazione i campi utilizzati per inserire le relative informazioni sono identici a quelli dei compensi per i diritti di autore, esaminati nel precedente sottoparagrafo: Quadro D al punto D3. I redditi degli associati sono tuttavia assumono un differente valore nel codice indicante il tipo di reddito (colonna 1): pari a “3”. Nel caso del secondo modello di dichiarazione (Upf) il dato sui redditi lordi degli associati è registrato nel rigo RL26 del Quadro RL “Altri redditi”. Il reddito

netto si ottiene detraendo le rispettive ritenute esposte al rigo RL30. In presenza di più tipologie di reddito incluse Sezione III del Quadro RL non è possibile attribuire distintamente le ritenute d'acconto alle varie componenti economiche.

3.5.8 - Il reddito dei lavoratori autonomi occasionali e dei venditori a domicilio

Il lavoro autonomo occasionale si caratterizza per una serie di peculiarità che comprendono: l'autonomia del lavoratore circa il tempo e il modo della prestazione, la mancanza di un coordinamento con l'attività del committente, il carattere episodico dell'attività e la mancanza dell'inserimento funzionale del lavoratore nell'organizzazione aziendale. I proventi derivanti dall'attività di lavoro autonomo occasionale sono fiscalmente inquadrati come *redditi diversi* (art. 67 comma 1, lettera l, del Tuir), con base imponibile data dalla differenza fra il reddito percepito e le spese specificamente inerenti alla sua produzione (art. 71, comma 2, Tuir) e con ritenuta d'acconto pari al 20 per cento e successivo conguaglio in sede di dichiarazione dei redditi. Le fonti fiscali deputate alla rilevazione di questi redditi sono alternativamente il modello 730 o il modello Upf. In particolare, nella prima fonte il dato in parola è inserito al rigo D4 “*Redditi diversi*” ed è contraddistinto dal codice “8” sul *tipo reddito* (colonna 1) riferibile ai corrispettivi derivanti da attività di lavoro autonomo non esercitate abitualmente (cfr. box tratteggiato in figura 3.17). Il reddito netto di questa voce si ottiene decurtando dall'imponibile *redditi* (colonna 2) le spese strumentali (colonna 3) e le ritenute d'acconto (colonna 4). Nella fonte Upf il reddito imponibile del lavoro autonomo occasionale è registrato in colonna 1 del rigo RL14 (*Compensi derivanti da attività di lavoro autonomo non esercitate abitualmente*) del Quadro RL “Altri redditi”.

Figura 3.19 – Sezione II-A: Redditi diversi – Quadro RL - Mod.Upf_2005

SEZIONE II-A Redditi diversi		1	2
		Redditi	Spese
RL4	Corrispettivi di cui all'art. 67, lett. a) del Tuir (lottizzazione di terreni, ecc.)	,00	,00
RL5	Corrispettivi di cui all'art. 67, lett. b) del Tuir (rivendita di beni immobili nel quinquennio)	,00	,00
RL6	Corrispettivi di cui all'art. 67, lett. c) del Tuir (cessioni di partecipazioni sociali)	,00	,00
RL7	Proventi derivanti dalla cessione totale o parziale di aziende di cui all'art. 67, lett. h) e h-bis) del Tuir	,00	,00
RL8	Proventi di cui all'art. 67, lett. h) del Tuir, derivanti dall'affitto e dalla concessione in usufrutto di aziende	,00	,00
RL9	Proventi di cui all'art. 67, lett. h) del Tuir, derivanti dall'utilizzazione da parte di terzi di beni mobili ed immobili	,00	,00
RL10	Redditi di natura fondiaria non determinabili catastalmente (censi, decime, livelli, ecc.) e redditi dei terreni dati in affitto per usi non agricoli	,00	
RL11	Redditi di beni immobili situati all'estero	,00	
RL12	Redditi derivanti dalla utilizzazione economica di opere dell'ingegno, di brevetti industriali, ecc. non conseguiti dall'autore o dall'inventore	,00	
RL13	Corrispettivi derivanti da attività commerciali non esercitate abitualmente	,00	,00
RL14	Compensi derivanti da attività di lavoro autonomo non esercitate abitualmente	,00	,00
RL15	Compensi derivanti dall'assunzione di obblighi di fare, non fare o permettere	,00	,00
RL16	Redditi determinati ai sensi dell'art. 71, comma 2-bis del Tuir	,00	,00
RL17	Totale (sommare gli importi da rigo RL4 a RL16)	,00	,00
RL18	Reddito netto (col. 1 rigo RL17 – col. 2 rigo RL17; sommare l'importo di rigo RL18 agli altri redditi Irpef e riportare il totale al rigo RN1 col. 2)	,00	
RL19	Ritenute d'acconto (sommare tale importo alle altre ritenute e riportare il totale al rigo RN21 col. 2)		,00

Il reddito netto in questione si ottiene detraendo alla colonna 1, le spese strumentali in colonna 2 del rigo RL 14 e le ritenute d'acconto al rigo RL19 (cfr. figura 3.19). In presenza di più tipologie di reddito incluse Sezione II-A “*Redditi diversi*” del Quadro RL non è possibile attribuire singolarmente le ritenute d'acconto alle varie voci economiche. Per quanto concerne i venditori a domicilio, si osserva che ai fini Irpef non sono tenuti a dichiarare le somme percepite a titolo di provvigione, dato che, in base all'art. 25-bis D.p.r. n. 600/73, la ritenuta operata dalla società committente è applicata a titolo d'imposta definitiva. L'esclusione dei loro redditi dalla base imponibile si estende anche ai contribuenti titolari di partita Iva.

3.5.9 - *Il reddito dei liberi professionisti e artisti*


Fanno parte di questa categoria i redditi da lavoro autonomo che derivano dall'esercizio delle arti e delle professioni come da comma 1 dell'art. 53 (ex art. 49) del Tuir. Il reddito in questione può essere soggetto a tre regimi fiscali: *a*) il regime analitico, *b*) il regime forfetario³⁹ (c.d. “contribuenti minimi”) di cui ai commi 171-173 dell'art. 3, della legge n. 662/96; *c*) regime agevolato, di cui agli artt. 13 e 14 della legge n. 388/2000, per i quali è previsto il pagamento di una imposta sostitutiva. Rientrano in questo ambito anche i proventi percepiti per prestazioni di volontariato o cooperazione rese ad organizzazioni non governative riconosciute idonee ai sensi dell'art. 28 della legge n. 49/87, qualora dette prestazioni discendano dall'assunzione di obblighi riconducibili ad un rapporto di lavoro autonomo. L'attività che genera questa fattispecie di reddito da lavoro autonomo deve avere le seguenti caratteristiche: *i*) essere professionale, nel senso che il soggetto deve porre in essere una molteplicità di atti coordinati e finalizzati verso un identico scopo con regolarità, stabilità e sistematicità; *ii*) essere abituale, nel senso che deve caratterizzarsi per lo svolgimento di atti in maniera non episodica, né saltuaria o comunque non programmata; *iii*) avere per oggetto attività diversa da quella imprenditoriale. La giurisprudenza e il Ministero dell'Economia e delle Finanze sono dovuti intervenire più volte per distinguere tra reddito d'impresa e reddito da lavoro autonomo derivante dall'esercizio di arti e professioni. Si ricorda che rientrano fra le attività imprenditoriali quelle svolte da professionisti che esercitano la professione all'interno di un'attività organizzata in forma d'impresa (Cass., Sez. I, 11 febbraio 1988 n. 1468). Mentre sono riconducibili all'attività derivante dall'esercizio di arti e professioni le attività svolta da: agronomi (C.T.C., Sez. XXII, 13 luglio 1982, n. 4998); pony express (Cass., Sez. lavoro, 25 gennaio 1993 n. 811); guide turistiche (R.M. n. 10/148 del 27 marzo 1974); spedizionieri doganali (C.M. n. 22 del 7 settembre 1988).

Il reddito di questi soggetti è rilevato nel Quadro RE dei modelli di dichiarazione Upf. Per la determinazione del reddito netto si procede in due passi. Nel primo passo si calcola il valore massimo tra gli imponibili esposti rispettivamente nel rigo RE20 colonna 2 (regime analitico: reddito o perdita derivante dall'esercizio delle attività professionali) e nel rigo RE23 colonna 2 (regime sostitutivo: reddito o perdita derivante dall'esercizio delle attività professionali al netto dell'imposta sostitutiva). Nel secondo, si somma il valore massimo ottenuto al passo 1 con l'imponibile riferito al regime fiscale forfetario, esposto al rigo RE24 colonna 2, e si decurta la relativa ritenuta d'acconto (rigo RE25). Vi è da osservare che in applicazione del regime fiscale forfetario o agevolato, il reddito imponibile dell'artista e del libero professionista sottostima di gran lunga il valore dell'effettivo guadagno dello stesso lavoratore al netto delle spese sostenute per l'attività (elusione fiscale). Viceversa, quando si calcola il reddito imponibile secondo il regime analitico, il dato fiscale si accosta di più alla definizione Eu-Silc di reddito lordo autonomo, in quanto ottenuto come differenza tra:

³⁹ Questo regime fiscale riguarda le persone fisiche e le imprese familiari che iniziano un'attività artistica o professionale ovvero d'impresa possono assoggettare il reddito di lavoro autonomo o d'impresa dell'esercizio in cui iniziano l'attività e nei due successivi ad un'imposta sostitutiva dell'imposta sul reddito delle persone fisiche nella misura del 10 per cento.

- ammontare dei compensi, in denaro o in natura, percepiti nel periodo d'imposta (anche sotto forma di partecipazione agli utili) al netto dell'Iva e dei contributi previdenziali e assistenziali posti dalla legge a carico del soggetto che li corrisponde;
- ammontare delle spese sostenute nel periodo stesso nell'esercizio dell'arte o professione ammesse in deduzione (art. 54, D.p.r. n. 917/1986).

Figura 3.20 – Quadro RE: Reddito di lavoro autonomo: esercizio arti e professioni - Mod.Upf_05

		PERIODO D'IMPOSTA 2004				
		EURO €				
		CODICE FISCALE <input type="text"/>				
		REDDITI QUADRO RE Reddito di lavoro autonomo derivante dall'esercizio di arti e professioni				
Determinazione analitica	RE1	Codice attività ¹	parametri e studi di settore: cause di esclusione ²			
	RE2	Compensi derivanti dall'attività professionale o artistica	¹	²	³	⁴
	RE3	Altri proventi lordi	¹	²	³	⁴
	RE4	Compensi lordi non annotati nelle scritture contabili	Parametri e studi di settore ¹	Maggiorazione ²	Concordato ⁴	
	RE5	Totale compensi (RE2 colonna 2 + RE3 colonna 2 + RE4 colonna 4)				
	RE6	Quote di ammortamento e spese per l'acquisto di beni di costo unitario non superiore a euro 516,46				
	RE7	Canoni di locazione finanziaria relativi ai beni mobili				
	RE8	Canoni di locazione non finanziaria e/o di noleggio				
	RE9	Spese relative agli immobili				
	RE10	Spese per prestazioni di lavoro dipendente e assimilato				
	RE11	Compensi corrisposti a terzi per prestazioni direttamente afferenti l'attività professionale o artistica				
	RE12	Interessi passivi				
	RE13	Consumi				
	RE14	Spese per prestazioni alberghiere e per somministrazione di alimenti e bevande in pubblici esercizi				²
		(Ammontare sostenuto ¹ ,00)				Ammontare deducibile ² ,00
	RE15	Spese di rappresentanza				²
		(Ammontare sostenuto ¹ ,00)				Ammontare deducibile ² ,00
	RE16	50% delle spese di partecipazione a convegni, congressi e simili o a corsi di aggiornamento professionale				
	RE17	Reddito detassato				
	RE18	Altre spese documentate				
RE19	Totale spese (sommare gli importi da rigo RE6 a RE18)					
RE20	Reddito (o perdita) delle attività professionali e artistiche				²	
	Incentivi di portata temporale limitata ¹ ,00				² ,00	
RE21	Reddito adeguato					
RE22	Reddito soggetto ad imposta sostitutiva (Vedere istruzioni) ¹		Imposta sostitutiva ²	³	⁴	
RE23	Reddito (o perdita) al netto dell'imposta sostitutiva (sommare l'importo agli altri redditi e riportare il totale nel quadro RN) (di cui a determinazione separata ¹ ,00)				² ,00	
Determinazione forfetaria	RE24	Reddito (o perdita) delle attività professionali e artistiche in regime forfetario (sommare l'importo agli altri redditi e riportare il totale nel quadro RN)				²
		Volume d'affari e altri compensi ¹ ,00			² ,00	
	RE25	Ritenute d'acconto				² ,00

3.5.10 - Il reddito dei titolari d'impresa

La nozione di reddito d'impresa trae origine dalla disciplina civilistica. In particolare, all'articolo 2082 del Codice civile, si definisce imprenditore colui il quale esercita professionalmente un'attività economica organizzata al fine della produzione o dello scambio di beni o di servizi. In realtà, la nozione fiscale di esercizio dell'attività d'impresa risulta più estesa, in quanto, oltre a comprendere quella civilistica, include anche l'esercizio di attività di natura commerciale non organizzate in forma di impresa. La tassazione delle imprese, trova nel titolo II del D.p.r. 917/86 la disciplina sul reddito d'impresa e solo in parte e in modo del tutto residuale,

il rimando al titolo relativo all'Irpef. Nell'ambito dei redditi d'impresa la disciplina fiscale distingue quelli che derivano dall'esercizio di imprese commerciali (art. 55, D.p.r. n. 917/1986) da altri redditi d'impresa (categoria residuale).

Per impresa commerciale si intende l'esercizio per professione abituale, ancorché non esclusiva, di:

- un'attività industriale diretta alla produzione di beni o servizi;
- un'attività intermediaria alla circolazione dei beni;
- un'attività di trasporto per terra, per acqua o per aria;
- un'attività bancaria o assicurativa;
- altre attività ausiliarie alle precedenti.

Costituiscono, inoltre, altri redditi d'impresa:

- i redditi derivanti dall'esercizio di attività organizzate in forma d'impresa dirette alla prestazione di servizi che non rientrano nell'articolo 2195 del codice civile;
- i redditi derivanti dallo sfruttamento di miniere, cave, torbiere, saline, laghi, stagni e altre acque interne;
- le attività agricole e zootecniche che eccedono i limiti di reddito agrario;
- l'allevamento di animali con mangimi ottenibili per meno di ¼ dal terreno;
- l'attività di trasformazione e commercializzazione dei prodotti agricoli o zootecnici che non rientrino nell'esercizio normale dell'agricoltura o che abbiano per oggetto prodotti ottenuti per meno del 50 per cento dal terreno o dagli animali su di esso allevati;
- la coltivazione di vegetali realizzati con strutture fisse o mobili, anche provvisorie, quando la superficie adibita alla produzione eccede il doppio di quella del terreno su cui la produzione stessa insiste;
- le attività indicate nell'art. 32 del D.p.r. n. 917/1986 e, in ogni caso, se tali attività sono svolte da società ed enti commerciali soggetti ad IRES, ovvero se gestite da S.n.c. o da S.a.s.

Non sono invece assimilati fiscalmente ai redditi di impresa ma fanno parte dei redditi agricoli i redditi derivanti dall'allevamento di animali con mangimi ottenibili per almeno un quarto dal terreno e i redditi derivanti dalle attività dirette alla produzione di vegetali tramite l'utilizzo di strutture fisse o mobili, anche provvisorie, se la superficie adibita alla produzione non eccede il doppio di quella del terreno su cui la produzione stessa insiste. Nel definire l'esercizio delle attività agricole, la norma di cui sopra pone alcuni limiti, il superamento dei quali influenza la natura del reddito prodotto, che, limitatamente alla quota imputabile all'attività eccedente, assume la qualifica di reddito d'impresa.

Per la rilevazione del reddito del titolare d'impresa occorre analizzare distintamente i seguenti tre Quadri del modello di dichiarazione Upf: *RF*, *RG* e *RD*. I quadri *RF* e *RG* devono essere compilati dalle persone fisiche titolari di reddito di impresa, ai sensi dell'art. 55 del Tuir. Il reddito derivante dall'esercizio delle attività commerciali (o altre attività d'impresa) va indicato nel quadro *RF* o nel quadro *RG* a seconda che sussista l'obbligo di tenuta della contabilità ordinaria o della contabilità semplificata (nel qual caso il reddito è determinato in base all'art. 66 del Tuir). In particolare per quanto riguarda i redditi dei titolari d'impresa che si avvalgono della *contabilità ordinaria* il reddito netto si ricava dalla differenza il dato rigo *RF60* colonna 2 (reddito imponibile di spettanza del titolare d'impresa) e la quota di ritenute spettanti al titolare ottenute dalla seguente formula:

$$RF61_4 \cdot (RF60_2 / (RF57_2 + RF60_2));$$

dove *RF61_4* indica le ritenute complessive, *RF57_2* denota i redditi d'impresa imputabili ai collaboratori e *RF60_2* rappresenta, come già accennato, il reddito imponibile di spettanza del titolare.

Figura 3.21 – Sezione del Quadro RF – Redditi d’impresa in contabilità ordinaria - Mod.Upf_05

RF57	Quote imputate ai collaboratori dell’impresa familiare o al coniuge di azienda coniugale non gestita in forma societaria (di cui a determinazione separata	1		2	
			,00		,00
RF58	Reddito d’impresa (o perdita) di spettanza dell’imprenditore (di cui a determinazione separata	1		2	
			,00		,00
RF59	Perdite d’impresa in contabilità ordinaria portate in diminuzione del reddito (di cui degli anni precedenti	1		2	
			,00		,00
RF60	Reddito d’impresa di spettanza dell’imprenditore al netto delle perdite d’impresa in contabilità ordinaria (sommare tale importo agli altri redditi e riportare nel quadro RN) (di cui a determinazione separata	1		2	
			,00		,00
			Crediti di imposta limitati sui fondi comuni di investimento	Crediti di imposta ordinari sui fondi comuni di investimento	Altri crediti
RF61	Dati da riportare nel quadro RN	1		2	3
			,00	,00	,00
			Ritenute	Eccedenze di imposta di società trasparenti	Acconti versati da società trasparenti
		4		5	6
			,00	,00	,00

Per quanto concerne i redditi dei titolari d’impresa che si avvalgono del regime di contabilità semplificata, il reddito netto si ricava anch’esso dalla differenza tra l’imponibile di spettanza dell’imprenditore (rigo RG34 colonna 2) e quota di ritenute a lui imputabili:

$$RG35_4 \cdot (RG34_2 / (RG31_2 + RG34_2));$$

dove RG35_4 indica le ritenute complessive, RG31_2 denota i redditi d’impresa imputabili ai collaboratori e RG34_2 rappresenta il reddito imponibile di spettanza del titolare.

Figura 3.22 – Sezione del Quadro RG – Redditi d’impresa in contabilità semplificata - Mod.Upf_05

RG31	Quote imputate ai collaboratori dell’impresa familiare o al coniuge di azienda coniugale non gestita in forma societaria (di cui a determinazione separata	1		2	
			,00		,00
RG32	Reddito d’impresa (o perdita) di spettanza dell’imprenditore al lordo delle perdite d’impresa in contabilità ordinaria (di cui a determinazione separata	1		2	
			,00		,00
RG33	Perdite d’impresa in contabilità ordinaria portate in diminuzione del reddito (di cui degli anni precedenti	1		2	
			,00		,00
RG34	Reddito d’impresa (o perdita) di spettanza dell’imprenditore al netto delle perdite d’impresa in contabilità ordinaria (sommare tale importo agli altri redditi e riportare nel quadro RN) (di cui a determinazione separata	1		2	
			,00		,00
			Crediti di imposta limitati sui fondi comuni di investimento	Crediti di imposta ordinari sui fondi comuni di investimento	Altri crediti
RG35	Dati da riportare nel quadro RN	1		2	3
			,00	,00	,00
			Ritenute	Eccedenze di imposta di società trasparenti	Acconti versati da società trasparenti
		4		5	6
			,00	,00	,00

Infine, i redditi dei titolari di imprese agricole e di allevamento la cui attività è classificata a titolo di *reddito agricolo* sono rilevate nel Quadro *RD*. In questo caso, il loro reddito netto si ottiene dalla differenza tra l’imponibile associato al titolare dell’impresa (rigo RD17) e le ritenute da lui subite (rigo RD18).

Figura 3.23 – Sezione del Quadro RD – Redditi di attività di allevamento agricolo - Mod.Upf_05

RD14	Quote imputate ai collaboratori dell'impresa familiare o al coniuge di azienda coniugale non gestita in forma societaria		,00
RD15	Reddito (o perdita) dell'impresa di spettanza dell'imprenditore		,00
RD16	Perdite d'impresa in contabilità ordinaria portate in diminuzione dal reddito (di cui degli anni precedenti ¹ ,00)	²	,00
RD17	Reddito d'impresa di spettanza dell'imprenditore al netto delle perdite d'impresa in contabilità ordinaria		,00
RD18	Ritenute d'acconto subite di spettanza dell'imprenditore		,00

3.5.11 - I proventi dei compartecipanti al reddito di imprese e società di persone

Fanno parte di questa categoria le quote di reddito, derivanti da partecipazioni in società ed associazioni o da collaborazioni in imprese familiari o coniugali, attribuite ai seguenti soggetti:

- soci delle società di persone ed equiparate di cui all'art. 5 del Tuir;
- collaboratori di imprese familiari;
- coniuge che partecipa all'azienda coniugale non gestita in forma societaria;
- persone fisiche, membri di Gruppi Europei di Interesse Economico (GEIE) residenti nel territorio dello Stato o, se non residenti, con stabile organizzazione nel territorio dello Stato;
- soci di società che hanno optato per il regime della trasparenza ai sensi dell'art. 116 del Tuir.

Secondo l'ordinamento italiano si può parlare di *impresa familiare* quando la gestione di un'azienda è affidata oltre che all'imprenditore titolare, anche al coniuge, ai parenti e agli affini. Sono considerati parenti del titolare e quindi potenziali compartecipanti:

- i discendenti entro il terzo grado come il figlio o il nipote; gli ascendenti (es. il padre), i collaterali (fratelli);
- gli affini, entro il secondo grado i parenti del coniuge;
- gli affini, entro il secondo grado i coniugi del figlio, del genitore e del fratello del titolare.

Il reddito prodotto dall'impresa familiare è diviso tra i *familiari*, per un limite che non può eccedere il 49 per cento e posto che questi abbiano prestato la propria attività in modo continuativo e prevalente, e l'*imprenditore* la cui quota deve essere non inferiore al 51 per cento. Nel caso in cui si verificano delle perdite durante l'esercizio di impresa, queste devono essere imputate unicamente al titolare e non al collaboratore familiare.

L'*azienda coniugale* si riferisce, da canto suo, ad aziende costituite dai coniugi, prima o dopo il matrimonio, per il compimento di attività in comune. Per l'azienda coniugale non gestita in forma di società, ai fini dell'imputazione del reddito, si applicano le regole dell'impresa familiare. In questo caso però il reddito da imputare al coniuge è pari al 50 per cento dell'ammontare risultante dalla dichiarazione del titolare. L'imprenditore deve, inoltre, rilasciare al coniuge dell'azienda in comunione un prospetto analogo a quello che il titolare dell'impresa familiare rilascia ai collaboratori familiari. Se l'azienda viene gestita in forma societaria, si applica la disciplina prevista per le società di persone ed il reddito o le perdite sono imputati a ciascun coniuge secondo le regole per le società di persone.

Il *GEIE* è un'associazione tra imprese e professionisti appartenenti a Stati diversi della CEE che ha una propria soggettività tributaria. È disciplinato dall'articolo 11 del d.lgs. n. 240/1991. Il *GEIE* residente nel territorio dello Stato, sebbene non sia soggetto a nessuna

imposta sul reddito, deve presentare la dichiarazione per determinare l'imponibile da attribuire ai membri del gruppo.

Le informazioni sui redditi dei compartecipanti sono contenute nel Quadro *RH* del modello di dichiarazione dei redditi Upf. In particolare ai fini della nostra indagine si è voluto distinguere i compartecipanti che svolgono un'attività professionale ed artistica (c.d. "lavoro autonomo" in materia fiscale) da quelli che collaborano ad un'attività d'impresa. I primi sono stati infatti classificati alla stregua dei liberi professionisti e artisti, tenuti alla compilazione del quadro fiscale RE dell'Upf (si veda § 3.5.7), mentre i secondo sono stati inquadrati come coadiuvanti, collaboratori o soci di imprese familiari, coniugali e società che svolgono attività diversa da professionale ed artistica. Il *reddito netto* della prima categoria di lavoratori è data dalla differenza tra l'imponibile indicato al rigo RH14 colonna 2 (RH14_2) e le ritenute d'acconto (rigo RH17) a loro imputabili:

$$RH17 \cdot (RH14_2 / (RH14_2 + RH15_2));$$

dove RH17 indica le ritenute complessive, RH14_2 denota il reddito imponibile dei collaboratori che svolgono attività professionale ed artistica e RH15_2 rappresenta il reddito imponibile di delle altre forme di collaboratori. Chiaramente il reddito netto di questi ultimi si ottiene decurtando dal loro imponibile la quota restante di ritenute d'acconto.

Figura 3.24 – Sezione del Quadro RH – Compartecipanti al reddito d'impresa- Mod.Upf_05

Sezione III Determinazione del reddito Dati comuni alla sez. I ed alla sez. II	RH7	Redditi di partecipazione in società esercenti attività d'impresa (di cui a determinazione separata	1	,00	2	,00	
	RH8	Perdite di partecipazione in società esercenti attività d'impresa in contabilità ordinaria				,00	
	RH9	Differenza tra rigo RH7 e RH8 (se negativo indicare zero) (di cui a determinazione separata	1	,00	2	,00	
	RH10	Perdite d'impresa in contabilità ordinaria				,00	
	RH11	Differenza tra rigo RH9 e RH10 (di cui a determinazione separata	1	,00	2	,00	
	RH12	Perdite d'impresa di esercizi precedenti				,00	
	RH13	Perdite di partecipazione in impresa in contabilità semplificata				,00	
	RH14	Totale reddito (o perdita) di partecipazione in società esercenti attività d'impresa (di cui a determinazione separata	1	,00	2	,00	
	RH15	Totale reddito (o perdita) di partecipazione in associazioni tra artisti e professionisti o in società semplici (di cui a determinazione separata	1	,00	2	,00	
	RH16	Totale reddito (o perdita) di partecipazione (sommare tale importo agli altri redditi e riportare il totale al rigo RN1 col. 4)	di cui a determinazione separata		3		
			Impresa	Lavoro autonomo			
			1	,00	2	,00	,00
	Sezione IV Riepilogo	RH17	Totale ritenute d'acconto				,00
RH18		Totale crediti d'imposta	Crediti d'imposta limitati sui fondi comuni di investimento	1	,00	2	,00
			Crediti d'imposta ordinari sui fondi comuni di investimento				,00
RH19		Reddito sul quale spetta la deduzione				,00	
RH20		Totale credito per imposte estere ante opzione				,00	
RH21		Totale oneri detraibili				,00	
RH22		Totale eccedenza				,00	
RH23		Totale acconti				,00	

Capitolo 4 – Strategie per l'integrazione dei dati campionari e di fonte amministrativa: soluzioni al problema dell'attribuzione dei redditi

4.1 - Introduzione

Una volta completata la ricostruzione delle diverse componenti di reddito (variabili target dell'indagine di Eu-Silc) a partire dai giacimenti informativi di natura fiscale, o più in generale di natura amministrativa, il passo seguente consiste nel *riconciare*, dapprima, il contenuto delle variabili tra le diverse fonti amministrative (ottenendo così una sintesi dei dati degli archivi amministrativi) e successivamente nell'individuare un sistema di regole che permettano la soluzione ai problemi di incoerenza e di divergenza (*inconsistency*) tra le informazioni rilevate su una stessa unità da fonte campionaria e amministrativa (integrata). Le strategie adottate variano a seconda della variabile target considerata, ovvero della corrispondente componente di reddito. Nel presente capitolo verranno analizzate tre distinte strategie, impiegate rispettivamente nell'attribuzione finale dei redditi da pensione, dei redditi da lavoro dipendente e da lavoro autonomo.

Per le altre componenti di reddito, quali ad esempio le liquidazioni di fine rapporto, non è stata applicata alcuna tecnica di risoluzione dei problemi di incoerenza, poiché l'unica fonte fiscale a disposizione "Cud/770" forniva maggiori garanzie in termini di qualità rispetto al dato campionario. Si è stabilito, pertanto, di correggere il valore riportato sul questionario dell'indagine con il dato registrato negli archivi del Cud; ciò per tutti i casi in cui sussisteva una corrispondenza tra le due fonti (*units matched*). Mentre, per i casi in cui non vi era alcuna corrispondenza (*units not matched*) si è deciso di mantenere il valore dichiarato nell'indagine Eu-Silc. Riguardo, poi, agli assegni familiari (Anf) ci si è limitati a donare il valore registrato nella fonte del Cud/770 alla corrispondente unità campionaria, mentre non si è intrapresa alcuna azione di correzione nei casi in cui il dato era esclusivamente di natura campionario. Ciò in ragione del fatto che la fonte Cud/770 è tipicamente affetta da errori di sottocopertura, cioè rileva solamente un segmento della popolazione dei beneficiari di Anf: i lavoratori del settore privato che percepiscono gli assegni familiari dell'Inps in forma indiretta, ovvero tramite il datore di lavoro¹.

Nell'ambito dei redditi da pensione si è realizzata una prima fase di armonizzazione del contenuto delle variabili di reddito rispettivamente rilevate nel Casellario centrale dei pensionati e nelle fonti fiscali. In pratica, dal primo archivio si è calcolata la variabile *derivata*: "importo lordo annuale per tipologia di pensione", inteso come prodotto tra l'importo mensile rilevato al 31 dicembre ed il numero di mensilità corrispondenti, cioè al fine di omologarlo al dato di fonte fiscale².

Successivamente si sono classificate le pensioni da CP sulla base del regime fiscale in: esenti o meno da imposizione. In ultimo si sono confrontati gli importi annui delle pensioni da

¹ Il datore anticipa le somme degli assegni per conto dell'Inps ed effettua il conguaglio in sede di versamento dei contributi tramite il modello DM10. Per approfondimenti si rinvia a P. Consolini e R. De Carli (2002).

² In particolare, per i trattamenti del Casellario con decorrenza antecedente l'anno di riferimento dei redditi si è proceduto, per ciascuna tipologia di pensione, a moltiplicare l'importo mensile per il numero massimo di mensilità erogabili in un anno. Mentre, per le pensioni decorrenti nello stesso anno di riferimento dei redditi si è proceduto a calcolare l'importo annuo in due fasi: dapprima moltiplicando l'importo mensile per il numero di mensilità intercorrenti tra il mese di decorrenza e la fine dell'anno (31 dicembre) e, successivamente, applicando i ratei di tredicesima (se dovuti).

CP soggette a tassazione con gli importi registrati presso gli archivi fiscali. Il valore finale dei trasferimenti pensionistici si ricava in modo semplice ed immediato quando il valore degli importi annui lordi delle pensioni tassabili da CP è coerente con il dato fiscale, ovvero quando le pensioni sono rilevate esclusivamente dal Casellario centrale dei pensionati (mancati agganci con i dati fiscali) o alternativamente da fonte fiscale (mancati agganci con i dati Inps). In presenza di valori difforni tra le due fonti si sono dovute formulare opportune ipotesi per realizzare la loro riconciliazione (ad es. correggendo il numero di mensilità da CP).

Infine, nei casi in cui non era possibile stabilire alcun abbinamento tra unità campionarie e fonti amministrative (*units not matched*), il valore delle pensioni (importi netti) è stato fissato in base ai dati riportati nelle interviste (5,3 per cento dei casi in Eu-Silc 2005).

L'identificazione della componente di reddito da pensione tramite fonte amministrativa (integrata) precede logicamente la fase di assegnazione del reddito da lavoro dipendente alla persona fisica per mezzo della stessa fonte. In altri termini, la prima operazione è propedeutica alla seconda e ne condiziona i risultati. Una volta identificati i redditi da lavoro dipendente di fonte amministrativa si procede ad effettuare un'analisi della coerenza dei valori ivi riportati. I risultati sperimentali ottenuti in questa fase indicano l'elevato livello di qualità delle informazioni sui redditi dei dipendenti e sulle mensilità retribuite presenti nella fonte Cud/770. Essa, come vedremo più avanti, costituisce il perno centrale per l'attribuzione dei livelli di reddito da lavoro dipendente di fonte amministrativa.

Dopo avere assegnato il valore del reddito dei dipendenti da fonte amministrativa, il passo successivo consiste nel confrontare il dato amministrativo con quello campionario. Allo scopo di armonizzare il contenuto delle due fonti, il reddito individuale mensile da lavoro dipendente di fonte Eu-Silc viene, dapprima, scisso nelle componenti esenti o meno da tassazione e successivamente, per la sola componente tassabile, moltiplicato per il numero di mensilità retribuite nell'anno. La variabile derivata così ottenuta rappresenta l'importo annuo della componente tassabile da lavoro dipendente che è l'aggregato direttamente osservabile dalla fonte fiscale. Definito il termine di confronto sui redditi da lavoro dipendente l'analisi procede con l'identificazione dei possibili casi di incoerenza che emergono dal confronto delle due tipologie di fonti.

Tavola 4.1 – Analisi delle incoerenze tra dati di fonte campionaria e amministrativa.

FONTE AMMINISTRATIVA INTEGRATA	Fonte campionaria (indagine Eu-Silc)	
	Perceptor di reddito da lavoro dipendente	Non perceptor di reddito da lavoro dipendente
Perceptor di reddito da lavoro dipendente	<i>Coerenza (1)</i>	<i>Incoerenza (2)</i>
Non perceptor di reddito da lavoro dipendente	<i>Incoerenza (1)</i>	<i>Coerenza (2)</i>

Come si osserva dalla tavola 4.1 i nodi critici sono rappresentati dalle incoerenze del primo e del secondo tipo, ovvero dalle unità che secondo la fonte amministrativa percepiscono redditi da lavoro dipendente mentre non li percepiscono in base alla fonte campionaria e viceversa. La soluzione ai problemi di incoerenza, come vedremo in seguito, viene effettuata tenendo conto della struttura dei redditi rispettivamente di fonte campionaria e amministrativa. La stessa analisi delle incoerenze è applicata al caso dei redditi da lavoro autonomo. Nei casi di coerenza del primo tipo, cioè quando ambedue le fonti segnalano la presenza di redditi da lavoro dipendente per la stessa unità statistica, si procede ad analizzare i livelli di reddito (variabile derivata) rilevati rispettivamente dalle due fonti e ad esaminare eventuali divergenze. Le ipotesi che vengono utilizzate per dirimere i casi di divergenza nei livelli di reddito rilevati nelle fonti citate differiscono a seconda della tipologia di reddito analizzata.

In relazione ai redditi da lavoro dipendente si è ipotizzato essere prossimo al “vero” il dato fiscale quando prevale sul valore campionario ovvero quando è di poco inferiore (soglia del 10 per cento in termini di differenze relative). Viceversa, si è considerato come prossimo al “vero” il dato campionario quando esso risulta essere almeno doppio rispetto al dato fiscale, sottintendendo la presenza di somme pagate in nero o di compensi di fuori busta ai lavoratori dipendenti. Negli altri casi intermedi si è considerato come prossimo al vero l’importo che, tra fonti fiscali e indagine campionaria, più si accosta al valore imputato dal software *IVAware*.

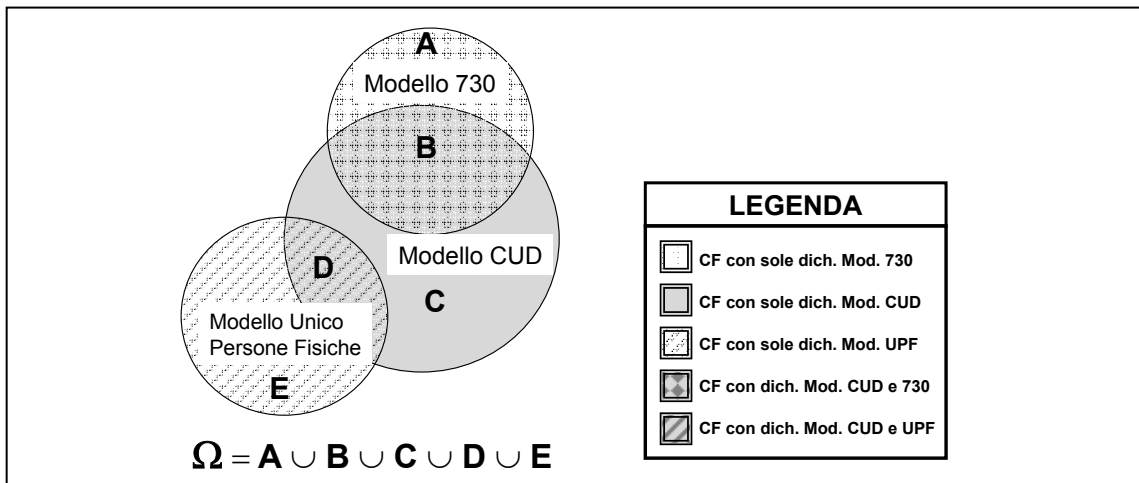
Riguardo ai redditi da lavoro autonomo si è invece fissata la regola del valore massimo, secondo la quale si attribuisce il massimo tra i valori rilevati nell’indagine campionaria e nelle fonti fiscali; ciò distintamente per i redditi dei professionisti e per i titolari/compartecipanti di imprese.

4.2 - Il disegno per l’integrazione dei microdati di fonte amministrativa

Condizione essenziale affinché si possa stabilire una connessione fra più archivi è la presenza di una comune chiave identificativa del record, che nel nostro caso è costituita dal “codice fiscale individuale (*CF*)”. L’operazione di *merge* tra le diverse fonti fiscali deputate alla rilevazione delle principali componenti reddito (*Cud/770*, *730* e *Upf*) consente di ottenere un archivio integrato contenente 5 distinti segmenti o sottoinsiemi di riferimento per l’analisi³ (figura 4.1). In questa rappresentazione grafica si ipotizza, per semplicità, l’inesistenza di problemi di copertura tra i vari archivi fiscali. Pertanto l’archivio dei modelli *Cud* è esaustivo delle dichiarazioni riguardanti i percettori dei redditi da lavoro dipendente, dei redditi da pensione (tassabili) e delle liquidazione di fine rapporto. Similmente, l’archivio dei modelli *730* comprende il complesso delle dichiarazioni dei titolari di redditi, aggiuntivi o sostitutivi a quelli da lavoro dipendente e assimilati (da fabbricati, terreni, compensi di lavoro autonomo occasionali, ecc.), o di coloro che richiedono rimborsi d’imposta per detrazioni/deduzioni. Come si evince dalla figura 1, uno stesso individuo può essere contestualmente presente nell’archivio dei modelli *Cud* e dei modelli *730* (insieme **B**). Tale situazione si verifica quando i titolari di redditi da lavoro dipendente e assimilati dichiarano di possedere altri redditi oppure semplicemente richiedono il rimborso d’imposta. Le dichiarazioni *730* non si sovrappongono a quelle del *Cud* quando i titolari possiedono esclusivamente redditi diversi da quelli da lavoro dipendente e assimilati (insieme **A**). Lo stesso ragionamento si applica, pur con le opportune differenze in termini di componenti di reddito coinvolte, nei confronti delle dichiarazioni dei modelli *Cud* e *Unico persone fisiche*. Viceversa le dichiarazioni *730* e *Upf* sono alternative e per definizione non si sovrappongono.

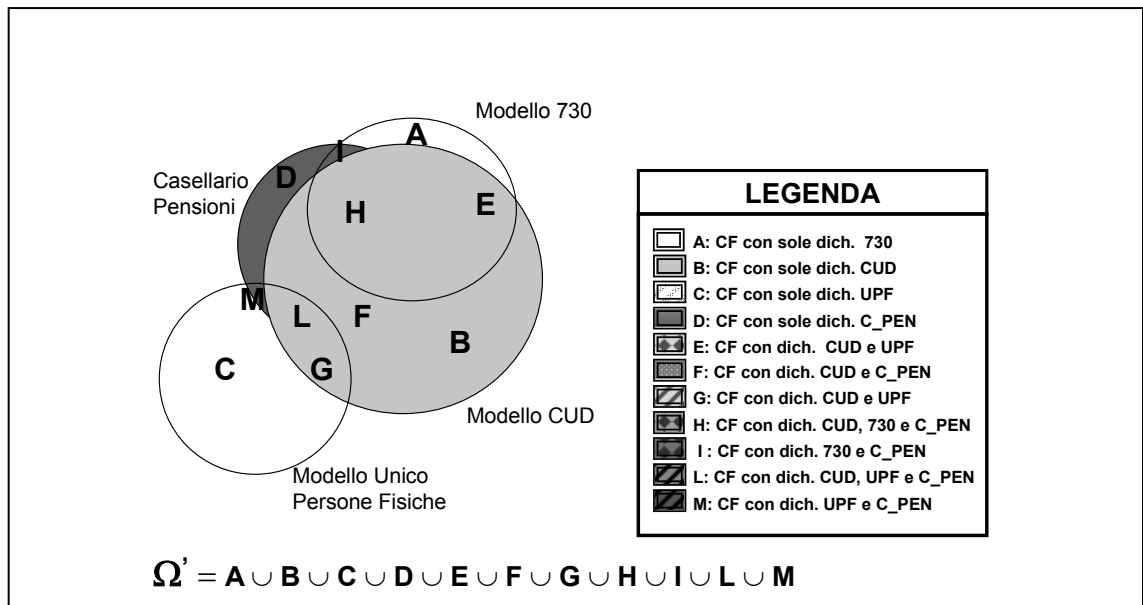
³ I cinque segmenti dell’archivio integrato sono rappresentati da: A) archivio delle persone fisiche con sole dichiarazioni *730*; B) archivio delle persone fisiche con sole dichiarazioni *Cud*; C) archivio delle persone fisiche con sole dichiarazioni *Unico persone fisiche (Upf)*; D) archivio delle persone fisiche con dichiarazioni *Cud* e modello *730*; E) archivio delle persone fisiche con dichiarazioni *Cud* e *Upf*.

Figura 4.1 – Struttura dell’archivio integrato dei dati fiscali (Ω)



Se alla precedente struttura dell’archivio integrato dei dati fiscali aggiungiamo le informazioni del Casellario centrale dei pensionati (abbreviato: Casellario pensionati), si perviene alla definizione all’archivio amministrativo integrato che costituisce la base per l’assegnazione delle componenti individuali di reddito. Una volta realizzato il *merge* tra i quattro archivi fiscali e amministrativi (Cud, 730, Upf e Casellario pensionati) e dunque generato l’archivio integrato dei redditi di fonte amministrativa, il disegno per l’integrazione procede dapprima con la fase di individuazione della prima componente di reddito, cioè i “trasferimenti pensionistici”, e successivamente con l’individuazione delle restanti componenti di reddito da lavoro, in funzione anche degli esiti della prima fase. Tale fase preliminare fornisce lo spettro delle possibili combinazioni con cui i percettori di reddito, ed in particolare di trasferimenti pensionistici, figurano tra le varie fonti. Come è possibile osservare dal grafico in figura 4.2 vi sono in totale 11 combinazioni di tipologie di fonti del dato in grado di rilevare i dati reddituali sulle persone fisiche e, segnatamente, sui titolari di pensione.

Figura 4.2 – Struttura dell’archivio amministrativo integrato (Ω')

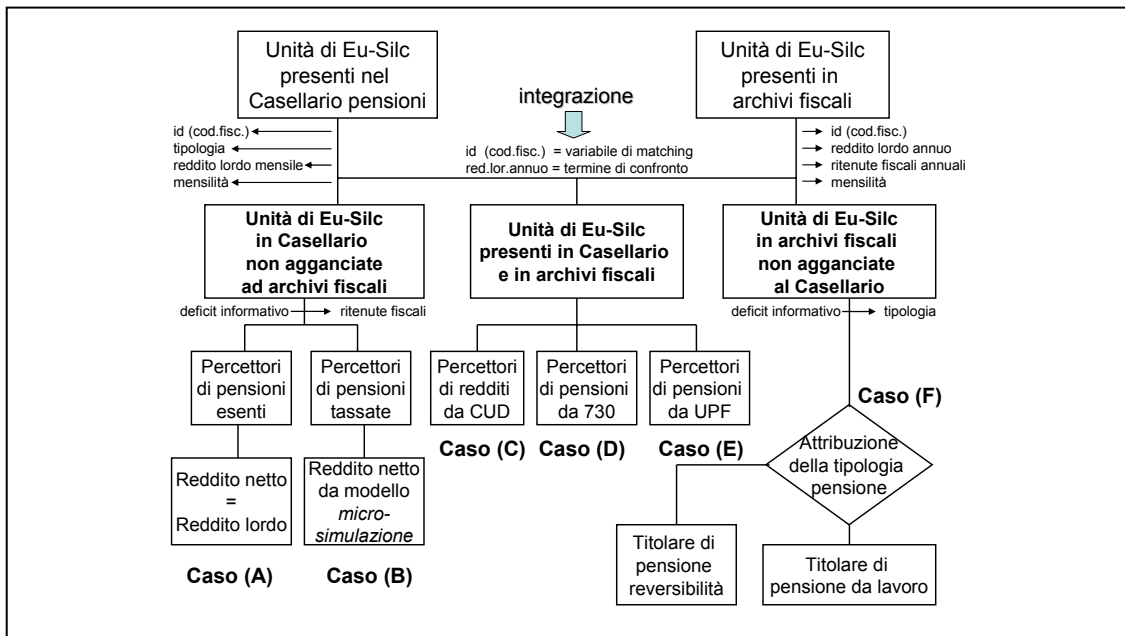


4.3 - L'identificazione dei redditi da pensione da fonte amministrativa

Nel processo di identificazione dei percettori di reddito da pensione in base alle fonti del dato, vi è un primo gruppo di analisi costituito dai pensionati rilevati esclusivamente dal CP, rappresentati dal sottoinsieme **D** di figura 4.2 (7,47 per cento del totale dei pensionati 2005 rilevati in almeno una fonte). In tale segmento figurano quasi esclusivamente i beneficiari la cui pensione è esente dall'imposizione fiscale, ma in via residuale possono entrare a farvi parte anche i titolari di pensioni tassabili la cui chiave individuale non si aggancia (*units not linked*) con gli archivi fiscali, per motivi di sottocoperta delle fonti fiscali o di errori nella chiave di abbinamento del CP (0,74 per cento). Un secondo gruppo riguarda i percettori di reddito rilevati esclusivamente dal Cud/770 (sottoinsieme **B** in figura 4.2), al cui interno figurano marginalmente i pensionati: 0,62 per cento del totale dei pensionati rilevati in almeno una fonte. In quest'ultimo caso, è possibile rilevare esclusivamente il livello del reddito da trasferimento (netto e lordo) ma non la tipologia di pensione corrispondente, informazione che è donata dall'archivio del CP (cfr. figura 4.3). Tra le due estremità rappresentate da gruppi citati, figura una serie di possibilità di combinazioni per tipologia di fonte ugualmente interessanti. La combinazione più rilevante è costituita dai record in cui le informazioni sui pensionati sono presenti contestualmente nel CP e nell'archivio del Cud/770: sottoinsiemi **F**, **L** e **H** in figura 4.2 (90,19 per cento). Se si calcolano le “differenze relative in valore assoluto” tra l'importo lordo delle pensioni tassabili da CP e quello delle pensioni di fonte Cud/770 per ciascun titolare di pensione, si osserva che nell'84,14 per cento dei casi esse risultano inferiori alla soglia del 5 per cento. Questo risultato fornisce una prima dimostrazione dell'elevato grado di coerenza delle informazioni sui livelli delle pensioni tra le due fonti amministrative. Dal punto di vista operativo si può osservare che le unità statistiche dell'indagine Eu-Silc che si agganciano al Casellario pensioni ma non agli archivi fiscali (sottoinsieme **D** di figura 4.2) devono essere analizzate secondo le tipologie di pensioni da esse percepite (cfr. caso A e B di figura 4.3). La presenza di componenti pensionistiche esenti da tassazione non pone alcun problema

computazionale nel calcolo del reddito netto in quanto la componente lorda (la sola rilevata nel CP) coincide con quella netta (caso A). Viceversa in presenza di trattamenti pensionistici soggetti ad imposizione, si pone il problema di calcolare il reddito netto a partire dalla componente lorda rilevata nel CP (caso B).

Figura 4.3 – Processo di integrazione sui redditi da pensione tra fonti amministrative e fiscali



4.3.1 - Il calcolo del reddito netto delle pensioni tassabili di sola fonte CP

La procedura che calcola il reddito netto a partire dalle pensioni rilevate esclusivamente dal Casellario centrale dei pensionati⁴ (casi A e B di figura 4.3) utilizza la tipologia pensionistica come fattore discriminante nell'applicazione dei vari passi di elaborazione. In particolare, quando l'unità statistica risulta titolare di sole tipologie di pensioni totalmente esenti (ad es. pensioni sociali) da imposizione il programma assegna banalmente un identico valore all'importo annuo lordo e netto della corrispondente pensione. Viceversa, nel caso pensioni parzialmente tassate come le pensioni di invalidità, non avendo scorporato le componenti esenti da quelle tassabili, si è utilizzato quale discriminare l'importo di queste ultime. Per semplicità sono state considerate come tassabili le pensioni di importo annuo superiore ai 7 mila euro, mentre sono state classificate come esenti le restanti pensioni con valori al sotto della suddetta soglia. Si rammenta che il valore di 7 mila euro corrisponde alla "no-tax area" per la tassazione dei redditi delle persone fisiche riferite all'anno 2005. Il reddito netto delle pensioni di invalidità si ottiene in due passi, dapprima si ricava l'imposta che grava sui questi redditi e successivamente si decurta dal reddito lordo.

In presenza di pensioni da lavoro e/o di reversibilità (che si ricorda essere gravate da imposizione fiscale per la parte eccedente i 7 mila euro) si utilizza la medesima soglia (no tax area) delle pensioni di invalidità per separare il caso banale dei redditi netti coincidenti con i redditi lordi, da quello in cui si deve procedere al calcolo della relativa imposta. In questo caso le

⁴ Con riferimento alle basi dati dell'anno 2005 questa casistica comprende il 7,46 per cento delle unità statistiche a cui si associa almeno una pensione rilevata in una o più fonti amministrative e fiscali.

istruzioni per il calcolo dell'imposta e del reddito netto sono analoghe a quelle sopra riportate e riferite alle pensioni di invalidità. Se dovessero essere simultaneamente presenti pensioni di reversibilità e da lavoro si attribuisce l'imposta pro-quota rispetto al valore dei rispetti importi lordi.

4.3.2 - Riconciliazione del reddito da pensione tra fonte CP e fonti fiscali

Lo strumento delle *differenze relative* in termini di valori di reddito rilevati sulle stesse unità statistiche da diverse fonti del dato (amministrative e fiscali) rappresenta il perno della struttura decisionale utilizzata in fase di definizione dei livelli delle pensioni e, più in generale, in fase di attribuzione delle componenti di reddito in presenza di informazioni da Casellario pensioni e fonti fiscali (cfr. casi C⁵, D ed E di figura 4.3).

A rigore, per effettuare correttamente un confronto al livello di singola variabile target tra gli importi lordi annui delle pensioni del Casellario centrale dei pensionati e quelli degli archivi fiscali occorre fissare a priori quali siano le tipologie pensionistiche da CP soggette o meno a tassazione (armonizzazione del contenuto). Ciò in quanto le componenti pensionistiche esenti da tassazione, seppure rilevate nel CP, per definizione esulano dal contesto fiscale. Non costituiscono un problema per l'analisi comparativa le pensioni da lavoro (vecchiaia e anzianità) e quelle di reversibilità (con l'eccezione delle rendite Inail erogate ai superstiti), essendo teoricamente⁶ sempre gravate da imposta. Del pari semplice è il caso delle pensioni sociali e delle indennità di accompagnamento rappresentano, dove però non si applica alcuna tassazione (esenti). I trattamenti pensionistici inclusi nella variabile target "trasferimenti monetari per la funzione invalidità (PY130N/G)" sono invece parzialmente⁷ tassati e rappresentano una casistica di difficile trattazione.

La procedura di integrazione tra i valori delle pensioni delle diverse fonti è strutturata in diverse fasi che prevedono l'utilizzo di una batteria di differenze relative tra gli importi pensionistici rilevati presso le varie fonti.

La prima tipologia di differenze relative *DFRELI_PEN* fornisce informazioni sullo scostamento relativo degli importi annui per i trattamenti pensionistici teoricamente soggetti a tassazione (pensioni da lavoro e reversibilità) rilevati rispettivamente da fonte CP e fonte Cud (sottoinsiemi **F**, **L**, **H** di figura 4.2):

$$DFRELI_PEN = |CPLAV_E_13 + CPREV_T_13 - DB_PEN_RLOR| / DB_PEN_RLOR$$

dove *CPLAV_E_13* indica l'importo annuale lordo della pensione da lavoro ottenuta come prodotto dell'importo mensile rilevato al 31 dicembre dell'anno di riferimento per il numero di mensilità corrispondenti⁸, *CPREV_T_13* denota a sua volta l'importo annuale lordo della prestazione individuale "pensione di reversibilità", infine *DB_PEN_RLOR* rappresenta l'importo delle pensioni tassabili di fonte Cud.

⁵ Con riguardo alla base dati anno di riferimento 2005, la percentuale di casi in cui il reddito da pensione è contestualmente rilevato da fonte Casellario pensioni e da fonte Cud è pari al 90,19 per cento del complesso dei titolari di pensione rilevati in una delle predette fonti (Caso C).

⁶ In realtà queste due tipologie di pensioni sono *effettivamente* gravate da imposta quando il loro importo lordo supera la soglia della cosiddetta "no-tax area" che il legislatore ha fissato nella misura di 7 mila euro per i redditi da pensione riferiti all'anno 2005.

⁷ In particolare, l'esenzione si applica alle pensioni erogate agli invalidi civili, ai non vedenti e ai non udenti, le rendite (dirette) per infortunio sul lavoro e per le malattie professionali, le pensioni (dirette) di guerra, per assegno di benemeranza ai perseguitati politici. Sono, al contrario, soggette a tassazione l'assegno di invalidità o le pensioni di inabilità di inabilità pagate dall'Inps in favore dei lavoratori dipendenti ed autonomi.

⁸ Le mensilità pensionistiche si deducono a loro volta dalla data di decorrenza della pensione (mese e anno).

Il secondo gruppo di differenze relative $DFREL2_PEN$ fornisce le stesse informazioni del caso precedente, quando però nell'argomento da confrontare⁹ si includono anche le pensioni di invalidità di fonte CP $CPINV_E_13$ (parzialmente soggette ad imposizione):

$$DFREL2_PEN = |CPLAV_E_13 + CPREV_T_13 + CPINV_E_13 - DB_PEN_RLOR| / DB_PEN_RLOR$$

La terza categoria di differenze relative $DFREL3_PEN$ confronta l'importo annuo da fonte CP delle pensioni di invalidità: $CPINV_E_MR = (CPINV_E \times MESI_RAT)$, delle pensioni da lavoro $CPLAV_E_MR = (CPLAV_E \times MESI_RAT)$ e quella di reversibilità $CPREV_T_MR = (CPREV_T \times MESI_RAT)$, calcolate secondo un diverso criterio di *annualizzazione* (importi mensili del Casellario pensioni per il numero di mensilità dedotte dalle sole informazioni presenti nelle fonti fiscali), con l'importo lordo delle pensioni di fonte Cud (DB_PEN_RLOR):

$$DFREL3_PEN = |CPLAV_E_MR + CPREV_T_MR + CPINV_E_MR - DB_PEN_RLOR| / DB_PEN_RLOR$$

La quarta categoria di differenze relative $DFREL4A_PEN$ paragona, invece, l'importo annuo da fonte CP delle pensioni di invalidità $CPINV_E_13$ delle pensioni da lavoro $CPLAV_E_13$ e di reversibilità $CPREV_T_13$ con la somma degli importi lordi annui dei redditi da pensione DB_PEN_RLOR e da lavoro dipendente DB_LAV_RLOR da fonte Cud. In questa circostanza si ipotizza la presenza di errori nella procedura di scomposizione ed attribuzione delle componenti reddituali (reddito da lavoro dipendente e da pensione) di fonte Cud; per cui risultano impropriamente escluse dai redditi da pensione componenti altrimenti attribuite ai redditi da lavoro dipendente.

La quinta tipologia di differenze relative $DFREL4B_PEN$ considera nell'argomento da confrontare l'importo *annualizzato* (secondo il numero di mensilità dedotte dalle sole informazioni presenti nelle fonti fiscali) delle pensioni da fonte CP di invalidità ($CPINV_E \times MESI_RAT$) delle pensioni da lavoro ($CPLAV_E \times MESI_RAT$) e di reversibilità ($CPREV_E \times MESI_RAT$), e l'importo lordo dei redditi da pensione (DB_PEN_RLOR) e da lavoro dipendente (DB_LAV_RLOR) da fonte Cud. Nell'applicazione di questa differenza relativa vale lo ragionamento esposto per il caso precedente.

La sesta categoria di differenze relative $DFREL4C_PEN$ mette a confronto l'importo annuo da fonte CP delle pensioni di invalidità $CPINV_E_13$ delle pensioni da lavoro $CPLAV_E_13$ e di reversibilità $CPREV_T_13$ con l'importo dei redditi da lavoro dipendente di fonte Cud DB_LAV_RLOR . Anche in relazione a quest'ultimo caso si ipotizza la presenza di errori nella procedura di scomposizione ed attribuzione delle componenti reddituali (reddito da lavoro dipendente e da pensione) di fonte Cud; per cui le componenti economiche attribuite, in un primo momento, ai redditi da lavoro dipendente fanno in realtà parte dei redditi da pensione.

Le differenze relative $DFREL4D_PEN$ della settima tipologia comparano l'importo *annualizzato* (importi mensili di fonte CP per il numero di mensilità dedotte dalle sole informazioni presenti nelle fonti fiscali) delle pensioni di fonte CP di invalidità ($CPINV_E \times MESI_RAT$), da lavoro ($CPLAV_E \times MESI_RAT$) e di reversibilità ($CPREV_E \times MESI_RAT$) con l'importo dei redditi da lavoro dipendente di fonte Cud (DB_LAV_RLOR).

I ragionamenti fatti in considerazione della sesta tipologia di differenze relative si applicano, *mutatis mutandis*, al caso della settima tipologia.

L'ottava tipologia di differenze relative $DFREL5A_PEN$ prende in esame lo scostamento relativo degli importi annuali da fonte CP delle pensioni di invalidità ($CPINV_E_13$), lavoro ($CPLAV_E_13$) e reversibilità ($CPREV_E_13$) rispetto al dato annuo delle pensioni di fonte 730 $RLOR_PEN_M730$ (per le unità statistiche appartenenti ai sottoinsiemi **H** e **I** di figura 4.2). Gli

⁹ Relativamente alle unità statistiche dell'intersezione tra Casellario e Cud per i sottoinsiemi **F**, **L**, **H** di figura 4.2

esiti di questo tipo di confronto assumono rilevanza soprattutto quando la fonte Cud non rileva il dato reddituale (mancata copertura – sottoinsieme **I**) o lo rileva scostandosi significativamente dall'informazione del Casellario pensioni (segnalando possibili errori di registrazione nel Cud).

La nona tipologia di differenze relative *DFREL5B_PEN* analizza lo scostamento relativo degli importi *annualizzati* (secondo il numero di mensilità dedotte dalle sole informazioni presenti nelle fonti fiscali) da fonte CP delle pensioni di invalidità (*CPINV_E×MESI_RAT*), lavoro (*CPLAV_E×MESI_RAT*) e reversibilità (*CPREV_T×MESI_RAT*) rispetto al dato annuo delle pensioni di fonte 730: *RLOR_PEN_M730* (per le unità statistiche appartenenti ai sottoinsiemi **H** e **I** di figura 4.2).

La decima tipologia di differenze relative *DFREL6A_PEN* mostra lo scostamento relativo degli importi annuali da fonte CP delle pensioni di invalidità (*CPINV_E_13*), lavoro (*CPLAV_E_13*) e reversibilità (*CPREV_E_13*) rispetto al dato annuo delle pensioni di fonte Upf (Unico persone fisiche) *RLOR_PEN_Upf* (applicata alle unità statistiche appartenenti ai sottoinsiemi **L** e **M** di figura 4.2).

L'undicesima tipologia di differenze relative *DFREL6A_PEN* mostra lo scostamento relativo degli importi annualizzati (secondo il numero di mensilità dedotte dalle sole informazioni presenti nelle fonti fiscali) da fonte CP delle pensioni di invalidità (*CPINV_E×MESI_RAT*), lavoro (*CPLAV_E×MESI_RAT*) e reversibilità (*CPREV_E×MESI_RAT*) rispetto al dato annuo delle pensioni di fonte Unico persone fisiche: *RLOR_PEN_Upf* (per le unità statistiche appartenenti ai sottoinsiemi **L** e **M** di figura 4.2).

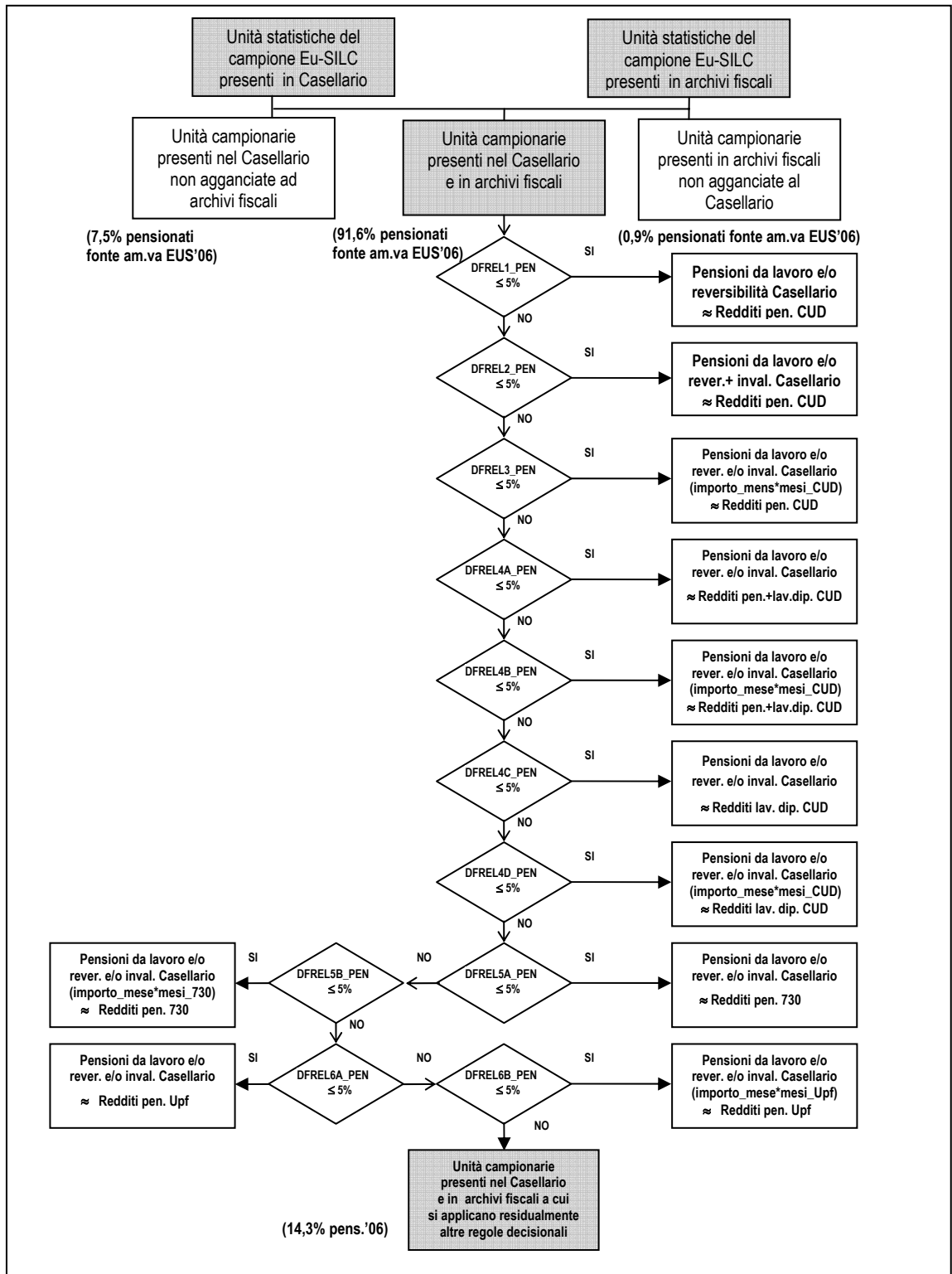
Una volta definita la batteria delle principali differenze relative il passo successivo consiste nel fissare una soglia di tolleranza sugli scostamenti relativi tra i dati di fonte amministrativa (Casellario pensioni) e le restanti fonti fiscali che qualifichi come accettabili i valori di reddito di una delle due tipologie di fonti. Ai fini l'integrazione dei dati pensionistici di diversa fonte si è stabilito di assegnare una soglia pari al 5 per cento; di conseguenza si assegna un valore alla variabile *target* per tutti quei record che riportano differenze relative in valore assoluto, tra i importi pensionistici delle due tipologie di fonti, contenuti nell'intervallo 0-5 per cento. In particolare, si assume come valore finale il dato sul reddito lordo/netto pensionistico rilevato presso la fonte fiscale. Con riguardo al complesso delle unità statistiche che nell'anno 2005 presentano dati pensionistici contestualmente nel Casellario pensioni e nella fonte Cud/770, si osserva che ben l'84,1 per cento di esse possiede un valore di soglia inferiore al 5 per cento in almeno una delle prime nove batterie di differenze relative considerate¹⁰. Mentre se si considera la batteria completa delle 11 differenze relative si nota che il vincolo del 5 per cento viene soddisfatto dall'85,7 per cento dei soggetti aventi almeno una pensione rilevata in ambedue le fonti. Per il restante 14,3 per cento non coperto si devono applicare nuove regole per riconciliare i valori del Casellario pensioni a quelli delle restanti fonti fiscali.

¹⁰ Se lo stesso numero di soggetti lo si raffronta rispetto al complesso delle unità statistiche aventi almeno un reddito pensionistico in una delle fonti citate si perviene ad una percentuale del 76 per cento circa. In altri termini, più di tre quarti delle unità campionarie di Eu-Sile'06, a cui è associata una pensione da fonte amministrativa e/o fiscale, possiede valori di reddito lordo annuo pensionistico convergenti tra il Casellario pensioni ed il Cud.

Le ipotesi di lavoro seguite per la riconciliazione dei casi residuali prendono in esame altri fattori, segnatamente la natura delle pensioni e il sistema fiscale. Al riguardo rilievo l'ipotesi sull'esenzione dal prelievo fiscale delle pensioni di invalidità e quella sull'applicazione della *no-tax area*, anche in relazione al cumulo di più pensioni. In particolare, si osserva come una gran parte dei casi residuali¹¹ si spiegano con la presenza di importi pensionistici (per trattamenti di invalidità, reversibilità e vecchiaia/anzianità, distinti o cumulati) inferiori alla soglia della *no-tax area* (7 mila euro). In questi casi è lecito assumere che vi siano scarti anche significativi tra gli importi lordi da archivi fiscali e da CP e che i dati reddituali di quest'ultima fonte (sistematicamente più alti dei primi) si accostino di più al concetto di reddito lordo. Ricade in quest'ultima casistica il 5,7 per cento delle unità campionarie Eu-Silc'06 a cui è associato un trasferimento pensionistico di fonte amministrativa e/o fiscale. Per altro verso, le differenze tra gli importi lordi annui delle due fonti del dato sono riconducibili al cumulo di pensioni tutte di invalidità, in parte esenti (3,1 per cento). Anche in questo caso è lecito assumere come prossimo al dato vero (lordo) il valore degli importi lordi di fonte CP. Ugualmente rilevante è l'assunzione che la *no-tax area* trovi applicazione anche nei confronti di quei redditi pensionistici (pensioni da lavoro e/o reversibilità) di poco superiori alla soglia dei 7 mila euro e comunque inferiori agli 8 mila euro (3,3 per cento). Infine, per riconciliare parte delle osservazioni contenenti dati pensionistici delle due fonti, si è deciso di rilassare la condizione relativa alla soglia del 5 per cento sulla batteria delle 11 differenze relative, per includere anche i casi compresi nell'intervallo 5-10 per cento (0,9 per cento). Il restanti casi in cui le fonti di discostano sottintendono errori legati all'attribuzione dei redditi nelle fonti fiscali e alla presenza di segnali di errore di duplicazione nel Cud non individuabili tramite altre fonti fiscali (1,3 per cento).

¹¹ Relativi, cioè, al caso in cui tutte le 11 tipologie di differenze relative non soddisfano la soglia del 5 per cento.

Figura 4.4 – Albero decisionale per l'integrazione dei dati pensionistici da fonte Casellario pensioni e fonti fiscali: differenze relative tra importi annui delle due tipologie di fonti Eu-Silc'06



4.3.3 - *Attribuzione della tipologia di pensione ai redditi di sola fonte fiscale*

L'attribuzione della tipologia di pensione in assenza di informazioni da fonte CP è un'operazione indispensabile ai fini dell'assegnazione alla corrispondente variabile target del reddito pensionistico rilevato dalla sola fonte fiscale. Trattandosi di pensioni tassabili (in quanto presenti nelle dichiarazioni Cud, ovvero 730/Upf) il campo di osservazione si restringe alle sole pensioni da lavoro, di reversibilità e di invalidità e infortunio sul lavoro.

Le informazioni utilizzate allo scopo dell'assegnazione comprendono l'età, il sesso, lo stato civile del titolare del reddito in questione, così come sono rilevate nel registro familiare (Modello SILC/06/REG) allegato al questionario individuale dell'indagine Eu-Silc. I criteri di attribuzione delle pensioni la lavoro variano a seconda del sesso del titolare in quanto nella legislazione previdenziale italiana sono previste soglie di pensionamento differenziate per genere. Ai titolari di reddito da pensione si assegna una pensione da lavoro (vecchiaia) nei casi in cui possiedono i requisiti di legge per ritirarsi dal lavoro, mentre si attribuisce loro una pensione di reversibilità quando risultano essere vedovi e in età non pensionabile, ovvero una pensione di invalidità/inabilità in assenza di requisiti per ottenere una delle precedenti pensioni (età al di sotto della soglia pensionabile e stato civile diverso da vedovo). Dal momento che la procedura di attribuzione delle tipologie di pensione si applica ad un numero esiguo di casi (0,91 per cento del totale dei pensionati) non si è ritenuto conveniente includere altri elementi decisionali, quali gli anni di contributi versati dal lavoratore (pensioni anzianità) oppure le tipologie pensionistiche eventualmente dichiarate dagli intervistati in Eu-Silc¹².

4.4 - **L'identificazione dei redditi da lavoro dipendente da fonte amministrativa**

L'identificazione dei percettori di reddito da lavoro dipendente, tramite l'archivio amministrativo integrato, si realizza quasi interamente sulla base dei dati riportati nella fonte fiscale Cud/770 (99,15 per cento) e solo marginalmente tramite le altre fonti 730/Upf (0,85 per cento). Si deve precisare che non sempre la fonte Cud/770 contiene informazioni corrette, complete e coerenti ai fini dell'attribuzione dei livelli di reddito da lavoro percepiti. Infatti, se si considerasse esclusivamente il contenuto informativo del Cud/770 si rischierebbe di attribuire impropriamente alla componente "reddito da lavoro dipendente" elementi reddituali facenti in realtà parte dei trasferimenti pensionistici; per i quali si è trovata traccia nel Casellario pensioni (1,07 per cento dei pensionati rilevati nella precedente fase). Per altro verso, il processo di identificazione delle varie componenti di reddito da fonte Cud/770 potrebbe risultare teoricamente indeterminato qualora venissero utilizzati differenti criteri di assegnazione con esiti tra loro incoerenti. Può accadere così che le informazioni contenute nella Parte B - dati fiscali - del Cud (ad es. il numero di giorni per cui spettano per le deduzioni rispettivamente da lavoro dipendente e da pensione) portino ad un'assegnazione diversa rispetto a quelle di altre Sezioni dello stesso modello di dichiarazione (si pensi ad esempio alla presenza/assenza di valori reddituali nella parte dati previdenziali). In questa circostanza si è stabilito di dare priorità al criterio che classifica le fonti di reddito secondo le informazioni della parte B, salvo poi rivedere l'attribuzione quando vengono incrociati i dati di fonte CP. Con riferimento all'indagine Eu-Silc 2006 (anno di riferimento 2005) si osserva che soltanto lo 0,74 per cento dei percettori finali di reddito da lavoro dipendente era in realtà erroneamente classificato in prima istanza (in base ai dati della parte B del Cud/770) come titolare di pensione (casistica *h* di tavola 4.2). Una volta

¹² Quest'ultima scelta scaturisce dall'osservazione empirica in quanto, confrontando la distribuzione delle tipologie pensionistiche sui dati campionari con quella dell'universo (Casellario), si rileva una distorsione nella struttura pensionistica dichiarata dagli intervistati; con sovra-rappresentazione delle pensioni da lavoro a scapito delle rimanenti tipologie.

scorporati i trasferimenti pensionistici dai redditi assimilati da lavoro dipendente il passo successivo consiste nel detrarre da quest'ultima voce i compensi derivanti da collaborazioni coordinate e continuative. Ciò che si ottiene a residuo da questa seconda operazione sono i redditi effettivi da lavoro dipendente. Tale procedura investe rispettivamente l'1,39 per cento dei percettori di redditi assimilati da Cud (non pensionistici) che detengono contestualmente redditi da lavoro dipendente e collaborazioni coordinate e continuative ed il 5,56 per cento dei percettori di redditi assimilati che possiedono esclusivamente compensi da co.co.co. (casistiche *f* e *g* tavola 4.2). Un elenco completo delle casistiche e problematiche incontrate durante la fase per il consolidamento e la riconciliazione delle informazioni sui redditi da lavoro dipendente (importo annuale e mensilità retribuite) e del relativo peso in termini percentuali rispetto al totale dei casi considerati (percettori di uno più redditi assimilati al lavoro dipendente rilevati in almeno un fonte fiscale) è riportato nella seguente tavola 4.2.

Tavola 4.2 – Analisi della copertura e degli errori presenti nelle fonti del dato sul reddito da lavoro dipendente

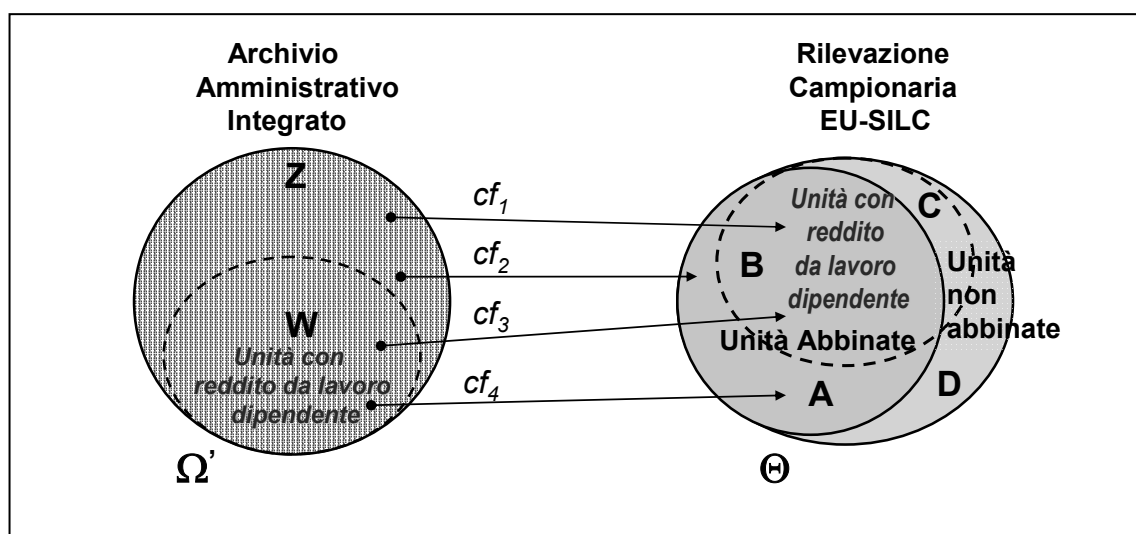
CASISTICA	Problematica	Azione intrapresa	Per cento sui percettori di redditi assimilati al lavoro dipendente (varie fonti fiscali)
a	Redditi da lavoro dipendente rilevati esclusivamente dal 730/Upf	Inserimento del dato retributivo di fonte 730/Upf	0,85
b	Redditi assimilati esclusivamente composti da lavoro dipendente rilevati nel Cud: mensilità errate nel quadro previdenziale (12 flag_mensili)	Correzione delle mensilità in base ai dati della Parte B	3,18
c	Redditi assimilati esclusivamente composti da lavoro dipendente rilevati nel Cud: mensilità mancanti nel quadro previdenziale (12 flag_mensili)	Imputazione delle mensilità in base ai dati della Parte B	4,33
d	Redditi assimilati esclusivamente composti da lavoro dipendente rilevati nel Cud: importi Cud inferiori ai valori rilevati nel 730/Upf	Correzione di redditi e mensilità in base al 730/Upf	1,33
e	Redditi assimilati esclusivamente composti da lavoro dipendente rilevati nel Cud: importi Cud doppi rispetto ai dati 730/Upf (errori di duplicazione)	Correzione dei redditi e delle mensilità in base al 730/Upf	0,07
f	Redditi assimilati misti da lavoro dipendente e da collab. coordinate e continuative rilevati nel Cud	Selezione della componente reddito da lavoro dipendente	1,39
g	Redditi assimilati esclusivamente composti da collaborazioni coordinate continuative rilevati nel Cud	Riclassificazione del reddito assimilato Cud come reddito da collaboraz. coord. contin.	5,56
h	Redditi da lavoro dipendente classificati in prima istanza come pensioni rilevate nel Cud	Attribuzione dei redditi a componente lavoro dipendente	0,74
i	Redditi assimilati misti pensione e lavoro dipendente rilevati nel Cud e ottenuti come differenza dai valori del Casellario pensioni	Attribuzione valori di reddito lavoro dip. in funzione del Casellario pensioni	0,31
j	Redditi assimilati misti lavoro dipendente e pensione Cud: mensilità non rispettano vincolo sugli importi mensili: "reddito da lavoro" > "pensione"	Correzione delle mensilità per rispettare il vincolo sugli importi mensili lav.dipendente > pensione	0,25
k	Redditi assimilati da lavoro dipendente desunti dal solo quadro previdenziale INPDAP	Attribuzione valori di reddito lavoro da quadro previdenziale	0,24
l	Redditi assimilati esclusivamente composti lavoro dipendente rilevati nel Cud: importi e mensilità valide	Nessuna	81,75
Totale			100,00

Dalla tavola in questione si evince che nella maggioranza dei casi (81,75 per cento) la parte B dati fiscali del Cud fornisce informazioni complete e corrette e coerenti (importi e mensilità retribuite) rispetto ad altre sezioni dello stesso archivio sui percettori di reddito da lavoro dipendente (casistica *l*). Oltre a ciò per un'ulteriore quota, pari al 7,5 per cento, la stessa Parte B del Cud rileva informazioni sulle mensilità retribuite formalmente corrette¹³ (e contestualmente degli importi), seppure incoerenti con i dati riportati in altre sezioni del medesimo archivio Cud (casistiche *b e c*). Di contro, la parte B del Cud riporta informazioni inesatte¹⁴ ed incoerenti sugli importi rispetto al dato di altre fonti amministrative in una percentuale esigua di casi: 2,2 per cento (casistiche *d, e, i, j, k*).

4.5 - Soluzioni alle incoerenze tra dati campionari e fiscali

Prima di esaminare i casi di incoerenza tra i dati sui redditi di natura campionaria e amministrativa occorre innanzitutto fare chiarezza su un aspetto importante che concerne la possibilità o meno di abbinare le due fonti in questione. Infatti, generalizzando il discorso, si può parlare di incoerenza quando due o più fonti del dato presentano valori difformi su una stessa unità di analisi. Resta implicita in questa definizione che l'unità sia osservabile su più fonti del dato (almeno due) e pertanto il campo di osservazione sulle incoerenze si restringe ai soli casi in cui vi è una corrispondenza tra unità statistiche delle varie fonti (almeno due). Se si prende a titolo illustrativo la seguente figura 4.5, si nota che il nostro campo di osservazione è ristretto alle sole unità campionarie che trovano un abbinamento con i dati dell'archivio amministrativo integrato (Ω'): sottoinsiemi **A** (unità abbinate con redditi da lavoro dipendente) e **B** (unità abbinate con redditi da lavoro dipendente).

Figura 4.5 – Tipologie di incoerenza sui dati del reddito da lavoro dipendente rilevati da più fonti.



¹³ Nel senso che esse assumono valori tali che l'importo mensile (variabile derivata) ricade nell'intervallo dei valori ammissibili della distribuzione dei redditi da lavoro dipendente.

¹⁴ Con il termine "inesatto" si intende che l'importo rilevato nella fonte Cud contiene errori di sottostima (superiori al 10 per cento in termini di differenze relative rispetto ad altra fonte), errori di duplicazione (importi pressoché doppi rispetto ad altra fonte), errori di incompatibilità (rispetto al dato di altra fonte: Casellario), errori di mancata registrazione.

Vi è da osservare, che le unità dell'archivio amministrativo integrato (a loro volta ripartibili, a seconda che possiedono o meno una dichiarazione di redditi da lavoro dipendente, negli **W** e **Z**) costituiscono un sottoinsieme delle unità del campione teorico di partenza, segnatamente quelle a cui è stato abbinato con successo un codice fiscale¹⁵ e dove, al tempo stesso, è associata una dichiarazione/documento dei redditi (Cud, 730, Upf, CP). Il confronto dei dati dell'archivio amministrativo, in realtà, è fatto rispetto alle unità del campione effettivo (rispondenti) Θ e non a quelle del campione teorico. Questi ultimi due insiemi differiscono per numero di unità sia a causa delle mancate risposte totali (cadute) che a causa dell'aggiunta di nuovi membri nella famiglia intervistata (componenti di fatto) non rilevati in precedenza durante la fase di creazione del campione. Pertanto si deduce che le cause di mancato abbinamento tra le fonti in esame sono imputabili tanto alla mancata individuazione del codice fiscale o di una dichiarazione di redditi, quanto alle mancate risposte totali e alla presenza di nuovi componenti familiari non rilevabili in precedenza¹⁶. Limitandoci ad esaminare i casi in cui è possibile stabilire un abbinamento tra le fonti ($\Theta' = \{\omega \in \Theta, \vartheta \in \Omega' : \omega = \vartheta\}$, dove $\Theta' = \mathbf{A} \cup \mathbf{B}$), emergono due tipologie di incoerenze: indicate a titolo illustrativo con i codici **cf₁** e **cf₄** in figura 4.1. Formalizzando il discorso, dato che ogni unità è posta in corrispondenza biunivoca con il suo codice fiscale, la prima tipologia di incoerenza è associata alle unità del sottoinsieme $\mathbf{B}' = \{\{b \in \mathbf{B}, z \in \mathbf{Z} : b = z\}\}$, formato dalle unità del campione che dichiarano nell'intervista di aver percepito un reddito da lavoro dipendente e che sono abbinate ad unità dell'archivio amministrativo integrato non aventi questa tipologia di reddito nelle rispettive dichiarazioni fiscali; la seconda tipologia di incoerenza è a sua volta legata alle unità del sottoinsieme $\mathbf{A}' = \{a \in \mathbf{A}, w \in \mathbf{W} : a = w\}$, rappresentate dalle unità campionarie che dichiarano di non possedere redditi da lavoro dipendente e che tuttavia sono abbinate a unità dell'archivio amministrativo con redditi di questa tipologia nelle rispettive dichiarazioni fiscali. Non si può invece parlare di incoerenza in relazione alle unità campionarie titolari di reddito da lavoro dipendente Θ_L ($\Theta_L = \mathbf{B} \cup \mathbf{C}$) dove non sia possibile stabilire alcun abbinamento con le unità di fonte amministrativa: sottoinsieme $\mathbf{C} = \{c \in \Theta_L, \omega \in \Omega' : c \neq \omega\}$. Come vedremo in seguito, la procedura utilizzata dirime prima i casi di incoerenza tra i percettori di reddito da lavoro dipendente e subordinatamente quelli tra i percettori di reddito autonomo. Le assunzioni fatte per risolvere le incoerenze prendono in esame anche le dichiarazioni degli intervistati circa la loro condizione lavorativa attuale e pregressa e il calendario lavorativo retrospettivo (anno di riferimento dei redditi).

Infine per quanto concerne i casi in cui vi è coerenza tra le fonti e compresenza di redditi (come l'esempio del codice **cf₃** in figura 4.1) il problema si sposta nel confronto tra i livelli di reddito assunti nelle rispettive fonti e nella costruzione di ipotesi che stabiliscano quale delle due informazioni sia più attendibile o prossima al vero. Il ragionamento sin qui esposto si estende, *ceteris paribus*, al caso dei percettori di reddito dal lavoro autonomo e co.co.co.

4.5.1 - I redditi da lavoro dipendente¹⁷

Il problema di incoerenza del *primo tipo* è caratterizzato dalla presenza nel questionario Eu-Silc di informazioni sul reddito da lavoro dipendente e dall'assenza di dati sulla stessa

¹⁵ Nel senso che il soggetto con quel codice è stato identificato dall'anagrafe tributaria.

¹⁶ Nell'indagine Eu-Silc 2006 il problema dei codici fiscali non agganciati all'anagrafe tributaria ha riguardato il 2,58 per cento delle unità del campione teorico, mentre il problema delle mancate risposte totali (cadute) ha riguardato il 15,4 per cento delle famiglie eleggibili (16,5 per cento degli individui eleggibili), infine i componenti di fatto (ovvero i membri della famiglia presenti al momento dell'intervista ma non rilevati in fase di formazione del campione) ammontano al 2,4 per cento degli individui eleggibili (cioè i soggetti di età ≥ 15 anni del campione effettivo).

¹⁷ L'impianto metodologico e le procedure di calcolo finalizzate alla soluzione del problema del primo e del secondo tipo di incoerenza sui redditi da lavoro dipendente sono state sviluppate rispettivamente dai colleghi del Servizio CEF/D: Ricci Roberta e Di Laurea Davide. Eventuali errori o omissioni contenuti nel sottoparagrafo sono da imputare esclusivamente all'Autore.

tipologia di reddito nelle dichiarazioni fiscali. L'ipotesi di lavoro utilizzata per la soluzione di questa incoerenza è che il reddito da lavoro dipendente di fonte campionaria vada riclassificato in altre componenti economiche certificate dalla fonte fiscale, oppure ricalcolato o ancora integrato con altre fonti di guadagno rilevabili esclusivamente dal fisco, in funzione della condizione professionale svolta nei 12 mesi dell'anno dei redditi come da questionario (calendario retrospettivo) e dei livelli e della composizione del reddito presente nelle rispettive fonti del dato. In altri termini, le soluzioni adottate hanno visto la trasformazione dello status originario di percettore di reddito da lavoro dipendente in qualche altra figura, ovvero la variazione della relativa struttura dei redditi. In merito all'indagine Eu-Silc 2006 si dimostra che le strutture di reddito rilevate presso le fonti campionaria e amministrativa assumono, nel caso di incoerenza del primo tipo, la seguente configurazione:

Tavola 4.3 – Differenze nella struttura di reddito tra le due tipologie di fonti del dato (incoerenza 1)

REDDITO DI FONTE FISCALE	Reddito di fonte campionaria	Variabile di classificazione GRUPPO_R	Numero osservazioni campionarie
Nessun tipo di reddito	dipendente	10	163
co.co.co.	dipendente	11	115
co.co.co.	dipendente e autonomo	12	6
co.co.co.	dipendente e co.co.co.	12.1	9
co.co.co. e autonomo	dipendente	13	31
co.co.co. e autonomo	dipendente e autonomo	14	10
co.co.co. e autonomo	dipendente e co.co.co.	15	1
co.co.co. e autonomo	dipendente, autonomo e altra tipologia di reddito	15.1	1
Autonomo	dipendente	16	319
Autonomo	dipendente e autonomo	17	39
Autonomo	dipendente e co.co.co.	18	2
Autonomo	dipendente e altra tipologia di reddito	19	6
Autonomo	dipendente, autonomo e altra tipologia di reddito	20	1
Autonomo	dipendente, co.co.co. e altra tipologia di reddito	21	1
TMNP ¹⁸	dipendente	30	59
TMNP	dipendente e indennità disoccupazione	31	37
TMNP	dipendente e Cig	32	4
TMNP e autonomo	Dipendente	33	6
TMNP e autonomo	dipendente e indennità disoccupazione	34	3
TMNP e autonomo	dipendente, autonomo e altra tipologia di reddito	35	1
Borsa lavoro	Dipendente	40	5
Borsa lavoro	dipendente, e altra tipologia di reddito	41	1
Borsa lavoro e autonomo	Dipendente	42	1
Borsa lavoro e co.co.co.	dipendente e co.co.co.	43	1
TOTALE			822

Le incoerenze affiorate in questa edizione coinvolgono un totale di 822 individui. Le azioni intraprese nei riguardi delle suddette incoerenze sono così articolate:

1. eliminazione della componente di reddito da lavoro dipendente e sua sostituzione con il reddito da co.co.co. di fonte amministrativa. I gruppi interessati riguardano le categorie del Gruppo_R: 11-12-12.1-13-14-15-15.1 del tavolo 4.3. Tale decisione coinvolge quei

¹⁸ Il termine TMNP sta ad indicare i trattamenti monetari non pensionistici, comprendenti le indennità di disoccupazione, la C.I.G., la malattia e la maternità del lavoratore dipendente.

soggetti che nel calendario retrospettivo (12 mesi dell'anno di riferimento dei redditi) del questionario Eu-Silc dichiarano di non aver mai assunto la condizione di lavoratore dipendente, ovvero dichiarano all'indagine un reddito netto complessivo paragonabile¹⁹ al reddito da collaborazione coordinata della fonte fiscale (112 casi);

2. rimozione del reddito da lavoro dipendente e inserimento del reddito da lavoro autonomo di provenienza fiscale. L'operazione coinvolge gli individui della variabile GRUPPO_R : 16-17-18-19-20-21. Anche in questa circostanza per intraprendere l'azione si considera l'assenza di esperienze lavorative alle dipendenze nel calendario retrospettivo del soggetto oppure l'accostamento tra il suo reddito complessivo campionario e il reddito autonomo di fonte fiscale (130 casi);
3. trasformazione dello status originario di percettore di reddito da lavoro dipendente in percettore di TMNP (Cig / indennità di disoccupazione) in coerenza con il dato fiscale (14 casi). I soggetti interessati alla trasformazione dello status afferiscono ai GRUPPO_R: 31-32-34. Il campo di applicazione di questa correzione riguarda i soggetti che hanno dichiarato nell'intervista di aver ricevuto in aggiunta indennità di disoccupazione e/o Cig e che risultano al fisco possessori di TMNP. Inoltre si richiede che la differenza relativa tra l'importo dichiarato come dipendente nel campione e l'importo TMNP da fonte fiscale sia compreso tra più o meno il 25 per cento;
4. rettifica della struttura dei redditi con integrazione della componente di reddito da collaborazione coordinata (di fonte fiscale) alla componente reddito da lavoro dipendente rilevata nell'indagine campionaria. Le categorie coinvolte riguardano i GRUPPO_R: 11-12.1-15-15.1. In questa operazione si assume come corretto il livello di reddito da co.co.co. rilevato al fisco e si ricava il valore del reddito da lavoro dipendente come differenza tra il dato campionario (reddito dipendente) e l'importo fiscale rilevato in altra tipologia (43 casi);
5. rettifica della struttura dei redditi con inserimento del reddito da lavoro autonomo da fonte fiscale al reddito da lavoro dipendente rilevato nell'indagine campionaria. Questa operazione investe i soggetti appartenenti al GRUPPO_R=16. Analogamente al caso precedente, si considera come valido l'importo del reddito autonomo di provenienza fiscale e si fissa il reddito lavoro dipendente come differenza tra il valore campionario e l'importo fiscale di altra tipologia (113 casi);
6. trasformazione dello status di percettore di reddito dipendente campionario in favore di quello di percettore di borsa lavoro di natura campionaria (2 casi). I soggetti potenzialmente interessati a questa operazione riguardano il GRUPPO_R: 40,41,42,43;
7. nessuna azione con relativo mantenimento dello status originario di percettore di reddito da lavoro dipendente e/o della relativa struttura di reddito (408 casi). In questa situazione le informazioni fiscali non sono in grado di supportare alcuna modifica circa lo status di percettore o della relativa struttura di redditi.

Le incoerenze del *secondo tipo*, contrassegnate dall'assenza di dati individuali sul reddito da lavoro dipendente nel questionario Eu-Silc e dalla contestuale presenza di questo tipo di informazioni nelle dichiarazioni fiscali, hanno interessato nell'indagine Eu-Silc 2006 un totale di 1.841 soggetti. L'ipotesi di lavoro adottata per la soluzione di questa incoerenza è che il reddito da lavoro dipendente di fonte fiscale debba integrare e completare le informazioni campionarie,

¹⁹ Nel senso che rientra nell'intervallo compreso tra il +/- 20 per cento della differenza relativa tra il reddito campionario e fiscale.

e dunque sia da inserire nella struttura dei redditi complessivi, a meno di segnali di misclassificazione della tipologia di reddito fra le due fonti. In presenza di elementi probatori, colti in sede di intervista, tali da far reputare come estremamente inverosimile la presenza di un reddito dipendente si procede alla relativa riallocazione.

Dal punto di vista operativo si è proceduto, dapprima, ad individuare e correggere possibili errori generati dalla procedura implementata in una fase anteriore che attribuisce una tipologia di reddito in base alle sole informazioni di fonte amministrativa²⁰ e, successivamente, a verificare la presenza di una serie di segnali del questionario discordanti rispetto all'assegnazione di parte amministrativa. Tra i segnali del questionario che depongono per la riallocazione del reddito da lavoro dipendente verso altre forme sono stati considerati: 1) assenza di una condizione professionale da lavoratore dipendente in tutti i 12 mesi dell'anno di riferimento dei redditi (calendario retrospettivo); 2) posizione professionale da lavoratore autonomo e nessun cambiamento di attività nei 12 mesi precedenti; 3) posizione professionale dell'attività svolta in passato diversa da lavoro dipendente. In altri termini, in assenza di segnali di lavoro dipendente ed in presenza di segnali di lavoro di tipo autonomo nel periodo temporale di riferimento dei redditi si reputa che il reddito da lavoro di fonte fiscale sia più verosimilmente riferibile al lavoro autonomo di cui il questionario porta evidenza. Di conseguenza, il reddito di fonte fiscale viene concordemente riclassificato. Volendo esemplificare i passi della procedura, si può immaginare di suddividere il problema della riallocazione del reddito esaminato in due punti: a seconda della presenza o meno di altre componenti individuali di reddito rilevate nelle fonti amministrativa e campionaria. Infatti, in quest'ultima situazione, cioè in assenza di altre forme di reddito tra le due fonti, la componente di reddito da lavoro unicamente rilevata nella fonte fiscale²¹ entra direttamente (1.154 casi). Nell'altra circostanza, cioè in presenza di altre tipologie di reddito (687 casi), si vanno a verificare i segnali provenienti dalla fonte campionaria. Al termine dell'intera procedura applicata all'indagine Eu-Silc 2006 vengono riclassificati i redditi da lavoro dipendente (di provenienza fiscale) di 217 soggetti verso altre tipologie in coerenza con l'informazione campionaria.

Quello che si intende evitare, attraverso la rimozione delle incoerenze del primo e del secondo tipo, è l'inclusione di tipologie di reddito che, rispetto alle finalità dell'indagine, andrebbero a ragione classificate diversamente; ovvero l'insorgere di errori di duplicazione nei redditi, o relative sottovoci, in quanto già dichiarati in sede d'indagine ma classificati in maniera differente.

4.5.2 - I redditi da lavoro autonomo e parasubordinato

La soluzione al problema delle incoerenze del primo e secondo tipo sui redditi da lavoro autonomo e parasubordinato (co.co.co.) è stata elaborata in modo tale da non contraddire i risultati della precedente riattribuzione del reddito da lavoro dipendente. In tale fase quindi sono state considerate esclusivamente le possibili incoerenze derivanti dalla diversa allocazione tra fonte campionaria e fonte fiscale di queste due sole componenti economiche.

Indirizzando la questione sull'incoerenza del primo tipo relativa alla componente da lavoro autonomo²² (in senso stretto), si possono configurare quattro diverse combinazioni nella struttura dei dati (tavola 4.4): 1) presenza di reddito da lavoro autonomo e da collaborazione coordinata di fonte campionaria ed esistenza della sola componente co.co.co di fonte

²⁰ Si pensi ad esempio alla valorizzazione della variabile DA_9 riferita alle "categorie particolari" secondo le modalità "M" (soci di cooperative) o "P" (titolari di borse di studio) presenti nella Parte A del modello Cud/770.

²¹ Vi è osservare che tale regola non è stata applicata in soli 9 casi in quanto le altre informazioni di natura amministrativa portavano a riclassificare questi redditi verso altre tipologie: compensi a soci di cooperativa, borse di studio ecc.

²² Essa attiene alle unità del campione che dichiarano nell'intervista di aver percepito un reddito da lavoro autonomo e che sono abbinate ad unità dell'archivio amministrativo integrato non aventi questa tipologia.

amministrativa; 2) presenza del solo reddito autonomo da questionario Eu-Silc ed esistenza della sola componente co.co.co nel dato amministrativo; 3) presenza di ambedue le tipologie di reddito da fonte campionaria e mancanza di redditi di questa natura dal dato amministrativo; 4) presenza del solo reddito da lavoro autonomo e assenza di redditi autonomi e parasubordinanti da fonte amministrativa. In relazione a ciascuna di queste quattro combinazioni si procede ad individuare e correggere la possibile causa di incoerenza.

In merito al punto 1, si tenta di risolvere l'incoerenza facendo ricorso a due distinte ipotesi di lavoro basate ambedue sul confronto dei livelli di reddito campionari. Da una parte si assume, la presenza di errori di duplicazione nelle due componenti di reddito (lavoro autonomo e co.co.co.) campionarie e dall'altra si ipotizza l'esistenza di evasione totale sul reddito da lavoro autonomo altrimenti dichiarato nel questionario.

La soluzione all'incoerenza di cui al punto 2 è stata elaborata tenendo conto di tre distinte evenienze: A) attività lavorativa principale autodichiarata nell'intervista in qualità di: "collaborazione coordinata e continuativa" oppure "prestazione d'opera occasionale"; B) la differenza relativa tra il reddito da lavoro autonomo campionario e il reddito da co.co.co. di fonte fiscale è inferiore o uguale al 20 per cento; C) attività lavorativa principale diversa da quella del punto A e differenza relativa tra le componenti di cui al punto B maggiore del 20 per cento (domina il livello di reddito campionario). Nelle situazioni riferite ai punti A e B si ipotizza la presenza di un errore di misclassificazione tra le due componenti da parte degli intervistati e si interviene di conseguenza attribuendo la tipologia di co.co.co. desunta dal fisco al reddito autonomo campionario. Nel caso C si assume, invece, la compresenza dei due redditi e si deriva il livello di reddito da lavoro autonomo come differenza tra il dato campionario (autonomo) ed il dato fiscale (co.co.co.).

Infine per le due restanti combinazioni di reddito, di cui a punti 3 e 4, si ipotizza che gli archivi fiscali siano affetti da deficit informativi e si assume come valida la struttura di reddito di provenienza campionaria.

Tavola 4.4 – Analisi delle incoerenze tra i redditi autonomi di fonte campionaria e amministrativa

FONTE AMMINISTRATIVA INTEGRATA		Fonte campionaria (indagine Eu-Silc)	
		Presenza di reddito da lavoro autonomo (in senso stretto)	
		Presenza di reddito da CO.CO.CO.	Assenza di reddito da CO.CO.CO.
Assenza di reddito da lavoro autonomo (in senso stretto)	Presenza di reddito da CO.CO.CO.	Esito 1 (3 casi)	Esito 2 (187 casi)
	Assenza di reddito da CO.CO.CO.	Esito 3 (13 casi)	Esito 4 (485 casi)

Se analizza l'incoerenza del primo tipo, relativamente alla componente da collaborazione coordinata e continuativa²³ (in senso stretto), si perviene anche in questa circostanza a quattro diverse combinazioni nella struttura dei dati. Vi è da osservare che una di queste combinazioni coincide con il caso 4 analizzato in precedenza (esito 4 di tavola 4.4), mentre un'altra, parimenti al caso 3 sopra menzionato, non modifica la struttura del reddito campionario (presenza dell'unica componente di reddito da co.co.co. di provenienza campionaria e assenza di redditi da fonti fiscali). Le situazioni più complesse da affrontare in questo ambito riguardano: 1) la contestuale presenza di reddito da collaborazione coordinata e da lavoro autonomo di fonte

²³ Essa riguarda le unità del campione che dichiarano nell'intervista di aver percepito un reddito da co.co.co. e che sono abbinate ad unità dell'archivio amministrativo integrato non aventi questa tipologia.

campionaria a fronte dell'esistenza della sola componente autonoma di fonte amministrativa (15 casi); 2) la discordanza nella classificazione dei redditi tra le fonti, per cui nel dato campionario si rileva la sola componente di reddito co.co.co. mentre nel dato fiscale si registra la sola componente di reddito autonomo (46 casi).

Nella prima circostanza si ipotizza che la struttura del reddito di natura fiscale sia corretta e pertanto si impone il contributo della sola componente di reddito da lavoro autonomo. Inoltre, si provvede a controllare e correggere eventuali errori di duplicazione, segnalati da importi simili²⁴ nelle due componenti di reddito campionarie.

Nella seconda fattispecie si suppone che l'incoerenza sia dovuta ad un errore di misclassificazione, per cui l'intervistato è portato ad interpretare la propria consulenza ad un committente come collaborazione coordinata continuativa (occasionale) anziché come prestazione professionale (e quindi non occasionale), in accordo con la relativa documentazione fiscale. In questo caso si interviene riclassificando il reddito da co.co.co. in reddito da lavoro autonomo.

Se si passa ad esaminare l'incoerenza del secondo tipo sui redditi da lavoro autonomo²⁵, si giunge identicamente a quattro distinte combinazioni nella struttura del reddito. Tra queste ultime, una coincide con la seconda fattispecie riferita all'incoerenza del primo tipo sui redditi co.co.co., mentre altre due attengono alla presenza di redditi esclusivamente da fonte amministrativa. Pertanto, l'unica combinazione rilevante ai fini dell'analisi è quella che vede il concorso di ambedue le tipologie di reddito da fonte amministrativa e la presenza della sola componente di reddito da co.co.co. campionaria (47 casi). In tale circostanza si ipotizza che la struttura dei redditi desunta dal fisco sia corretta e si provvede ad aggiungere un'ulteriore voce di entrata: reddito autonomo. L'importo di quest'ultima componente è fissato secondo il valore rilevato nella fonte fiscale, mentre l'importo del reddito da co.co.co. è dato da: i) il suo valore campionario, quando esso risulta inferiore alla somma degli importi delle due componenti rilevate dal fisco; ii) la differenza tra il suo valore campionario e il livello di reddito autonomo di fonte fiscale, nel caso opposto quando il primo risulta superiore alla somma degli importi fiscali. Infine, per quanto concerne l'incoerenza del secondo tipo sui redditi da co.co.co.²⁶ si perviene, *mutatis mutandis*, ad un risultato analogo a quello applicato alla stessa incoerenza sui redditi autonomi. In questo caso l'unica combinazione rilevante è quella che vede il concorso di entrambe le componenti di reddito fiscali e la sola rilevazione dei redditi da lavoro autonomo di natura campionaria (144 casi). Similmente al caso precedente si impone la struttura del dato fiscale e si aggiunge però come componente di reddito quella da co.co.co.. Per la definizione degli importi relativi alle due componenti vale lo stesso criterio applicato al passo precedente con l'inversione delle componenti di reddito implicate.

4.5.3 - I redditi da pensione

Anche per quanto concerne i redditi da pensione sussistono le stesse incoerenze tra fonti amministrative e fonti fiscali, contrassegnate dall'assenza di dati individuali sui trattamenti pensionistici nel questionario Eu-Silc e dalla contestuale presenza di questo tipo di informazioni nelle dichiarazioni fiscali e viceversa. Come è stato già detto al paragrafo introduttivo del capitolo 4 l'incoerenza tra le fonti è stata risolta inserendo sia titolari di trasferimenti pensionistici di sola provenienza fiscale che i pensionati rilevati esclusivamente in Eu-Silc. In

²⁴ Nel senso che la differenza relativa tra i valori delle rispettive componenti campionarie ricade nell'intervallo compreso tra il +/- 20 per cento.

²⁵ Essa coinvolge le unità del campione che dichiarano nell'intervista di non aver percepito alcun reddito da lavoro autonomo e che tuttavia sono abbinate ad unità dell'archivio amministrativo integrato percettrici di questa forma di reddito.

²⁶ Essa coinvolge le unità del campione che dichiarano nell'intervista di non aver percepito alcun reddito da lavoro autonomo e che tuttavia sono abbinate ad unità dell'archivio amministrativo integrato percettrici di questa forma di reddito.

particolare, in quest'ultima circostanza sono state effettuate le opportune procedure di imputazione, controllo correzione sugli importi mancanti o incorretti²⁷. Allo stesso tempo il problema della riconciliazione tra i valori delle pensioni per le due fonti è stata affrontata imponendo il dato amministrativo, ritenuto più affidabile, al dato campionario.

4.6 - Riconciliazione del reddito campionario e fiscale

4.6.1 - I redditi da lavoro dipendente

L'ipotesi sottostante la riconciliazione dei livelli retributivi individuali tra le due fonti del dato è che il valore *vero* sia contenuto nella fonte fiscale, sempre che il reddito rilevato nel questionario "risulti non superiore" a quello di origine fiscale. In quest'ultima circostanza si è tentato di riconciliare le discrepanze attraverso una serie di regole applicate selettivamente in funzione della distanza dei livelli di reddito tra le due fonti.

In riferimento a differenze relative tra il dato campionario e amministrativo comprese nell'intervallo (0,1), cioè quando il valore campionario è al più due volte il valore fiscale, sono stati applicati alcuni criteri decisionali legati a: 1) esistenza di elementi retributivi da questionario totalmente o parzialmente esenti da tassazione; 2) attendibilità delle risposte fornite dall'intervistato. La presenza nel questionario di taluni compensi parzialmente o totalmente esenti da tassazione ha fatto propendere nella correttezza dell'informazione natura campionaria²⁸. Viceversa in assenza di questi compensi si è privilegiato il dato amministrativo a patto che la differenza relativa tra valori delle due fonti risultasse comunque contenuta in una soglia del 10 per cento. In mancanza di queste due condizioni, si è proceduto a valutare l'attendibilità della risposta dell'intervistato relativamente al reddito dichiarato, ricorrendo al giudizio del rilevatore o all'identificazione del rispondente: persona interessata/altro familiare (proxy). Il dato campionario viene convalidato solo se il giudizio del rilevatore è positivo (risposte abbastanza/molto verosimili) e se, al tempo stesso, la persona rispondente coincide con la persona interessata. L'ipotesi sottostante la convalida del dato campionario è che il soggetto abbia percepito compensi fuori busta non risultanti dalle dichiarazioni dei redditi. Al contrario, se vengono i requisiti di affidabilità, si considera corretta l'informazione amministrativa.

In corrispondenza, poi, di forti discrepanze nei valori delle retribuzioni tra le fonti, rappresentate da differenze relative superiori al 100 per cento tra il dato campionario rispetto all'amministrativo, si è stabilito di considerare la distanza tra i valori *osservati* nelle due fonti ed il valore *imputato* dal package statistico *IVEware*²⁹ sulla stessa variabile reddito dipendente una volta eliminato il valore assunto dalla corrispondente unità campionaria. Tra le variabili utilizzate nel modello di regressione multipla incluso nelle procedure di *IVEware* sono stati considerati diversi elementi correlati con il tenore di vita del soggetto (possesso di beni durevoli,

²⁷ La fase di controllo e correzione si è limitata a verificare l'eventuale *incompatibilità* degli importi delle pensioni di natura campionaria rispetto ai valori massimi e minimi esibiti dalle fonti amministrative (in caso positivo si è proceduto a cancellare ed imputare un nuovo valore compatibile) e la *coerenza* dell'età pensionabile rispetto ai limiti vigenti fissati dalla legge.

²⁸ Di conseguenza si è assunto che la differenza nei livelli delle due fonti è ascrivibile a tali elementi retributivi.

²⁹ Sviluppato dal *Survey Research Center Institute for Social Research dell'University of Michigan*, il package si compone di diverse applicativi in linguaggio Sas e Fortran che realizzano l'imputazione multivariata dei valori mancanti tramite una sequenza di modelli di regressione multipla univariata (approccio "*variable by variable*"), partendo dalla variabile con minore tasso di non risposta ed utilizzando come predittori le variabili senza valori mancanti. Ciascun modello mette in relazione tra la variabile da imputare con tutte le altre variabili e utilizza sia i valori osservati sia i valori imputati nei passi precedenti. Il singolo modello è selezionato separatamente senza bisogno di definire un modello multivariato per l'intero insieme dei dati di cui, però, si assume l'esistenza. A seconda della natura della variabile da imputare è selezionato un differente modello di regressione tra quelli disponibili. Così, ad esempio per le variabili continue fa riferimento al modello di regressione lineare normale (con eventuale trasformazione della variabile se non è normale nella scala originaria), viceversa per le variabili binarie utilizza il modello di regressione logistico, infine per le variabili categoriche considera i modelli di tipo log-lineare. Per approfondimenti si rinvia a Manzari A. (2004), Raughunatan E. et al. (2001) e Vitaletti S. (2008).

titolo di godimento e caratteristiche dell'abitazione principale, indicatori di deprivazione, ecc.) e quindi indirettamente con il suo livello di reddito.

4.6.2 - I redditi da lavoro autonomo e parasubordinato

L'ipotesi sottostante la riconciliazione degli importi del reddito da lavoro autonomo nelle due fonti del dato è che il valore *vero* sia potenzialmente sottostimato in entrambi i casi; da una parte per motivi di incompletezza dell'informazione fiscale (componenti deducibili o evase al fisco) e dall'altra per la reticenza a dichiarare il vero dell'intervistato (*under-reporting*). L'attribuzione del livello di reddito da lavoro autonomo finale, inteso come migliore approssimazione, è stata realizzata distintamente per i guadagni da "libera professione e collaborazioni coordinate e continuative" e da attività imprenditoriali.

In merito alla prima tipologia di reddito (libera professione e collaborazioni coordinate continuative) si è stabilito di considerare come "prossimo al vero" il valore massimo tra il dato individuale dichiarato al fisco e il dato rilevabile in Eu-Silc, sia esso il valore osservato, calcolato in fase di soluzione delle incoerenze, ovvero imputato tramite *IVEware*³⁰. Si precisa che sino all'anno di indagine 2005 il questionario Eu-Silc non prevedeva la rilevazione separata delle due componenti di reddito da lavoro autonomo e da co.co.co. . Pertanto, nei casi in cui la fonte fiscale contemplava la presenza di ambedue le voci economiche, si è proceduto a stimare l'importo da co.co.co. campionario come quota del reddito "autonomo e parasubordinato" dichiarato congiuntamente nell'intervista, coerentemente con la composizione dei redditi rilevata nei documenti fiscali.

Riguardo ai "redditi di impresa" si è utilizzato un diverso approccio nella riconciliazione tra i dati amministrativi e campionari rispetto alla voce precedente. Dapprima si sono identificati i "lavoratori autonomi fittizi"; ovvero quei soggetti che in Eu-Silc dichiaravano di "*non aver mai lavorato in passato e di essere attualmente in condizione non professionale*" e che, tuttavia, al fisco risultavano percettori di redditi di partecipazioni in società. Quando questi soggetti vivevano presso famiglie senza altri lavoratori autonomi³¹, si è ipotizzato realisticamente che essi percepivano in realtà redditi da capitale, ovvero una remunerazione del capitale da loro prestato in una o più attività senza impegno lavorativo (nell'indagine Eu-Silc 2004 vi sono 42 casi). Viceversa, quando essi vivevano presso famiglie in cui vi era almeno un lavoratore autonomo (che dichiara in Eu-Silc di essere autonomo o di aver lavorato nella settimana di riferimento o in precedenza), si è stabilito di attribuire proporzionalmente il reddito dei primi (lavoratori autonomi fittizi) agli effettivi lavoratori autonomi. La peculiarità della procedura applicata ai percettori di reddito da impresa è che si effettua un confronto a livello familiare tra i redditi di impresa riportati nella fonte fiscale e campionaria, al netto dell'eventuale aggregato di reddito da libera professione. Rispetto al caso dei liberi professionisti, si assume come "prossimo al vero", il valore massimo dell'aggregato familiare dei redditi di impresa tra le due fonti del dato e non quello dei singoli individui. Tuttavia, anche quando su base familiare emerge una dominanza dei dati reddituali di fonte fiscale rispetto ai corrispettivi valori campionari, può accadere che a livello individuale tale ordinamento sui redditi non sia rispettato. In altri termini, può accadere che taluni redditi autonomi individuali dichiarati in Eu-Silc risultino superiori rispetto ai dati della controparte fiscale, nonostante a livello familiare sia vero il contrario. Tale incongruenza rappresenta un spia dell'eventuale presenza di lavoratori autonomi fittizi. In queste

³⁰ Nel caso dei redditi da lavoro autonomo si è fatta precedere la fase di imputazione sui valori mancanti della fonte campionaria, tramite il pacchetto *IVE-ware*, alla fase riconciliazione dei valori delle due fonti. Pertanto, a differenza della riconciliazione dei dati sui redditi dei dipendenti, una parte dei confronti sui redditi degli autonomi non prendono in esame i valori campionari effettivamente osservati, quanto piuttosto quelli imputati.

³¹ Nel senso che nessun componente della famiglia dichiara di essere lavoratore autonomo o di aver lavorato nella settimana di riferimento o in precedenza come autonomo.

circostanze si assume che la vera struttura dei redditi da impresa riferita a ciascun membro sia quella osservata in Eu-Silc, mentre l'ammontare corretto dei redditi familiari di impresa sia rappresentato dal valore più elevato, cioè quello fiscale. Viceversa, quando il valore aggregato dei redditi da impresa di fonte campionaria domina il corrispettivo dato di fonte fiscale, allora si ritiene che sia sempre corretto il dato individuale Eu-Silc. In analogia a quanto esposto in precedenza, è possibile che qualche valore individuale del reddito di impresa di fonte fiscale sia più elevato del corrispondente dato campionario, a dispetto dell'ordinamento osservato sull'aggregato famiglia. Questa circostanza fornisce anch'essa un segnale della presenza di percettori di redditi di impresa fittizi, e suggerisce pertanto di utilizzare come vera la struttura dei redditi individuali di fonte campionaria. Nel campione Eu-Silc 2004 sono stati individuati 52 lavoratori autonomi fittizi appartenenti a famiglie con almeno un lavoratore autonomo. Se si confrontano gli importi del reddito da lavoro autonomo rilevati contestualmente nell'indagine campionaria e nelle fonti fiscali (5.120 osservazioni), si osserva come in 2.956 casi queste ultime riportino valori più elevati rispetto alla prima.

Conclusioni e prospettive future

Il lavoro illustrato costituisce il primo esempio in Italia di applicazione di procedure di abbinamento esatto tra più fonti del dato finalizzate alla produzione di statistiche sociali su scala nazionale. Lo scopo dell'attuale ricerca è di illustrare metodi e soluzioni - dapprima sperimentate e, in seguito, applicate con successo - nel processo di integrazione di basi di dati campionarie ed amministrative relativamente all'indagine sui redditi Eu-Silc. La sperimentazione e la realizzazione di metodi finalizzati all'arricchimento e al completamento dell'informazione da più fonti, seppure indirizzate al contesto della produzione dei dati sui redditi, rappresentano utili linee guida applicabili anche in altri settori delle statistiche sociali. L'articolazione del processo di integrazione in fasi sequenziali - che inizia dall'analisi del fenomeno e dalla ricognizione delle fonti, si sviluppa attraverso la scelta della chiave per l'abbinamento, l'armonizzazione delle unità e delle variabili, l'analisi delle incoerenze e della completezza del dato, e termina con la riconciliazione dei valori tra le diverse fonti - rappresenta un modello universale che può, dunque, essere esteso a qualsiasi altra indagine in campo sociale. I vantaggi relativi all'impiego di più fonti integrate, amministrative e campionarie, su un determinato settore di analisi sono dati dal miglioramento delle qualità della informazione, per effetto sia della riduzione dell'errore di misurazione che del contenimento della distorsione per mancata risposta. Un esempio, a tal proposito, è fornito dai risultati dell'applicazione del processo di integrazione per l'anno di indagine 2004, dove è stata eliminata del tutto la sottostima sul numero stimato dei percettori da fonte campionaria (-7 per cento) rispetto al dato Casellario pensioni (rilevazione totale e nostro *benchmark*). I limiti insiti nel procedimento che combina le informazioni da più fonti sono legati essenzialmente alla natura dei dati amministrativi, il cui contenuto informativo non sempre è armonizzabile e riconducibile alle definizioni e alle variabili che costituiscono l'oggetto di interesse della ricerca. Occorre infatti ricordare che il dato di natura amministrativa, per sua natura, risponde ad esigenze di natura contabile, organizzativa e funzionale riconducibili ad attività e a compiti propri di una struttura pubblica/privata e, pertanto, non sono sempre utilizzabili per finalità di produzione statistica. Inoltre, problemi di tempestività nel rilascio degli archivi amministrativi e nell'aggiornamento delle relative variabili di studio, di copertura rispetto alla popolazione target, di mutamenti nel tempo del contenuto delle variabili, di accessibilità al dato per ragioni di protezione della privacy, possono compromettere seriamente il loro impiego ai fini statistici. Fortunatamente, non è questo il caso dei giacimenti informativi usati nella presente indagine.

Di recente è stato istituito un apposito Gruppo di lavoro interistituzionale Istat-Mef (Dipartimento delle Politiche Fiscali), avente lo scopo di vagliare le possibilità di utilizzo a fini statistici dell'archivio dei modelli fiscali Cud/770, che rappresenta una base dati fondamentale durante il processo di integrazione illustrato nei precedenti capitoli. I primi risultati di questo progetto dimostrano la buona qualità dei dati in esso presenti, in termini di copertura rispetto al dominio dei sostituiti d'imposta e dei percettori di reddito. Il lavoro conferma, d'altra parte, l'elevata coerenza tra i dati amministrativi e fiscali (e quest'ultimi nel loro interno), seppure su un sottoinsieme del collettivo dei percettori, ma al tempo stesso evidenzia l'esigenza di corredare l'informazione amministrativa con quella campionaria. Infatti, almeno nel settore delle indagini sui redditi, non è possibile perseguire l'obiettivo della riduzione del carico di risposta tramite l'utilizzo esclusivo di dati amministrativi senza perdere informazione sui percettori e sui rispettivi redditi.

Vincoli connessi alla disponibilità di risorse umane non hanno finora consentito di estendere il processo di integrazione in direzione di altre componenti di reddito quali: i trasferimenti monetari non pensionistici, i redditi da attività finanziarie e da capitale reale, e parte dei redditi autonomi (agenti e rappresentanti di commercio). Tuttavia, forti dei risultati positivi conseguiti nell'applicazione dell'integrazione delle fonti rispetto ai tre principali voci di reddito (reddito da lavoro dipendente, autonomo e trasferimenti pensionistici), resta come nostro obiettivo operativo primario la copertura di restanti componenti economiche.

Bibliografia

- Artina V. 2005. *Dichiarazione 730/2005 semplificato e ordinario: guida alla compilazione quadro per quadro*. Milano: Ipsosa. (Collana Dichiarazioni fiscali 2005)
- Artina V., G. Cremona, P. Monarca e N. Tarantino. 2005. *Unico persone fisiche 2005: Guida alla compilazione quadro per quadro*, Milano: Ipsosa. (Collana Dichiarazioni fiscali 2005)
- Atkinson A.B., L. Rainwater e T.M. Smeeding. 1995. *Income distribution in Oecd countries*. Paris: Oecd.(Social policy studies, n.18)
- Azteni et al. 1999. *Basi di dati*. Milano: McGraw-Hill.
- Biemer P. e L. Lyberg. 2003. *Introduction to survey quality*. John Wiley & Sons.
- Blum O. 2005. *Evaluation of editing and imputation supported by administrative records*. Unece work session on statistical data editing WP7, 16-18 May 2005, Ottawa (Canada),.
- Broccoli M. et al. 2006. *Il fattore territorio nel controllo di processo nelle statistiche socio-demografiche*. Sessione parallela “Il fattore territorio nella qualità dei dati”, Sesta conferenza nazionale di statistica, 6-8 novembre 2006, Roma.
- Ceccato F. 1998. *Il Sistema di classificazione delle prestazioni pensionistiche (Scpp)*. Roma, Istat. (Documenti Istat, n. 19).
- Chapman A.D. 2004. *Principles of data quality*. Report for the Global biodiversity information facility (Gbif), Copenhagen.
- Cody R. 1999. *Cody's data cleaning techniques*. Cary NC, USA: Sas Institute Inc.
- Consolini P. 2000. *Le prestazioni sociali monetarie non pensionistiche: aspetti istituzionali e classificazioni statistiche*. Roma: Istat. (Documenti Istat, n. 2)
- Consolini P. e R. De Carli. 2002. *Le prestazioni sociali monetarie non pensionistiche: unità di analisi, fonti e rappresentazione statistica dei dati*. Roma: Istat (Documenti Istat, n. 1).
- Consolini P. 2003. *Administrative data based statistics: the case of non-pension cash benefit*. Proceeding of the 17th roundtable on business survey frames, volume II, pp. 423-430, Roma, 26-31 October, 2003.
- Consolini P. 2004. *L'indagine sperimentale sull'archivio fiscale modd. 770 anno 1999: analisi della qualità del dato e stime campionarie*. Roma: Istat. (Contributi Istat, n. 29)
- Consolini P., M. Di Marco, R. Ricci e S. Vitaletti. 2006. *Administrative and survey microdata on self-employment: the Italian experience with the Eu-Silc project*. Iariw 29th general conference, Joensuu (Finland), 20-26 August, 2006.
- Di Marco M. 2006. *International comparability of microdata on incomes: lessons from the Eu-Silc project*. VIII International meeting on quantitative methods for applied sciences, Certosa di Pontignano (Siena), 11-13 settembre.

- Epland J. 2006. *Challenges in income comparability: experiences from the use of the register data in Norwegian Eu-Silc*. VIII International meeting on quantitative methods for applied sciences, Certosa di Pontignano (Siena), 11-13 settembre.
- Eurostat. 1996. *Esspros manual, population and social conditions, methods*. Luxembourg.
- Eurostat. 2004. *Description of target variables*, doc. Eu-Silc 065/04, Luxembourg.
- Fellegi I. P. e A. B. Sunter. 1969. "A theory for record linkage". *Journal of the American statistical association*. vol. 64, pp. 1183-1210.
- Guerrucci D. 1996. *Confronto tra i sistemi di classificazione dell'assistenza sociale*, contributo al gruppo di lavoro Istat-Usci, Roma.
- Herzog T. N., Scheuren F. J. e W. E. Winkler. 2007. *Data quality and record linkage techniques*. New York: Springer ed.
- Maletic, J. I. e A. Marcus. 2000. *Data cleansing: beyond integrity analysis*. Proceedings of the conference on information quality (IQ2000), Boston, pp. 200-209.
- Manzari A. 2004. *Valutazione comparativa di alcuni metodi di imputazione singola delle mancate risposte parziali per dati quantitativi*. Roma: Istat. (Contributi Istat, vol. 18)
- Nuccitelli A., F. Bosio e L. Fioriti. 2004. *L'applicazione Reclink per il record linkage: metodologia implementata e linee guida per la sua utilizzazione*. Roma: Istat. (Documenti Istat, n. 10)
- Nazzaro O. 1999. *Statistiche sulle pensioni e sui percettori di prestazioni pensionistiche: progetto di realizzazione delle pubblicazioni per la collana Informazioni*. Roma: Istat. (Documenti Istat, n. 16)
- Newcombe H. B. 1988. *Handbook of record linkage: methods for health and statistical studies, administration, and business*. Oxford (UK): Oxford University Press.
- Peperoni C. e F. Tuzi. 2007. *Trattamenti monetari non pensionistici: metodologia sperimentale per la stima degli assegni al nucleo familiare*. Roma: Istat. (Contributi Istat, n. 9)
- Rahm E. e H.H. Do. 2000. "Data cleaning: problems and current approaches". *Ieee bulletin of the technical committee on data engineering* vol. 23, n. 4, December 2000, pp. 1-13
- Raughunathan T.E., J.M. Lepkowski, J. Van Hoewyk e P. Solenberger. 2001. "A multivariate technique for multiply imputing missing values using a sequence of regression models". *Survey Methodology* vol. 27, n. 1, June 2001, pp. 85-95.
- Rendtel U. et al. 2004. *Report on quality of income data*. WP5 Chitex project, final version, January 2004.
- Scanu M. 2003. *Metodi statistici per il record linkage*, Roma: Istat. (Metodi e Norme, n. 16)
- Spinelli V. 2007. *Processo di acquisizione e trattamento informatico degli archivi relativi al modello di dichiarazione 770*. Roma: Istat. (Documenti Istat, n. 4)
- Studio Borgini Peverelli e Associati. 2005. *Dichiarazione 770 semplificato e ordinario: guida alla compilazione quadro per quadro*. Milano: Ipsoa. (Collana Dichiarazioni fiscali 2005)
- Van der Laan P. 2000. "Integrating administrative registers and household surveys". *Netherlands Official Statistics* vol. 15, Summer 2000.

- Verma Vijay. 2006. *Eu statistics on income and living conditions (Eu-Silc): issues in comparability and data quality*. VIII International meeting on quantitative methods for applied sciences, Certosa di Pontignano (Siena), 11-13 settembre, 2006.
- Ver Ploeg M., R.A. Moffitt e C.F. Citro. 2002. *Studies of welfare populations: data collection and research*. Washington: National Academies Press.
- Vitaletti S. 2008. “Correzioni e imputazioni delle informazioni sui redditi”. In *L’indagine europea sui redditi e le condizioni di vita delle famiglie (Eu-Silc)*, par. 4.4. Roma: Istat. (Metodi e Norme, n. 37)
- Wallgren A. e B. 2007. *Register-based statistics: Administrative data for statistical purposes*. John Wiley & Sons.

Appendice

Criteri per la generazione automatica del codice fiscale

Caratteri indicativi del cognome

Ai fini del calcolo dei tre primi caratteri distintivi del codice fiscale si utilizzano le lettere in maiuscolo del COGNOME e si tralasciano gli eventuali spazi (*blank*), gli apici, gli accenti e i trattini. Così, ad esempio, il cognome «De Mauro» è da considerarsi come «DEMAURO», del pari il cognome «D'Alessio» viene in realtà tradotto in «DALESSIO», e così via. Per i soggetti di sesso femminile si considera il cognome da nubile. In sintesi, le regole per il calcolo dei tre caratteri in esame sono riconducibili ad una delle nove condizioni, legate al numero di consonanti e vocali presenti nel cognome, di seguito specificate:

1.	[Nr. consonanti ≥ 3]	→ 1 ^a , 2 ^a , 3 ^a consonante
2.	[Nr. consonanti = 2] [Nr. vocali ≥ 1]	→ 1 ^a , 2 ^a consonante, 1 ^a vocale
3.	[Nr. consonanti = 2] [Nr. vocali = 0]	→ 1 ^a , 2 ^a consonante, "X"
4.	[Nr. consonanti = 1] [Nr. vocali ≥ 2]	→ 1 ^a consonante, 1 ^a vocale, 2 ^a vocale
5.	[Nr. consonanti = 1] [Nr. vocali = 1]	→ 1 ^a consonante, 1 ^a vocale, "X"
6.	[Nr. consonanti = 1] [Nr. vocali = 0]	→ 1 ^a consonante, "X", "X"
7.	[Nr. consonanti = 0] [Nr. vocali ≥ 3]	→ 1 ^a , 2 ^a , 3 ^a vocale
8.	[Nr. consonanti = 0] [Nr. vocali = 2]	→ 1 ^a , 2 ^a vocale, "X"
9.	[Nr. consonanti = 0] [Nr. vocali = 1]	→ 1 ^a vocale, "X", "X"

Per meglio comprendere le modalità con cui si ottengono le prime tre lettere del codice fiscale dal cognome si fornisce qualche esempio:

COGNOME	LETTERE PER IL CODICE FISCALE	METODO UTILIZZATO (CONDIZIONE)
Rossetti	RSS	1 ^a , 2 ^a , 3 ^a consonante (1.)
Rossi	RSS	1 ^a , 2 ^a , 3 ^a consonante (1.)
Delia	DLE	1 ^a , 2 ^a consonante, 1 ^a vocale (2.)
Delia Rossi	DLR	1 ^a , 2 ^a , 3 ^a consonante (1.)
Alo'	LAO	1 ^a consonante, 1 ^a vocale, 2 ^a vocale (4.)
Re	REX	1 ^a consonante, 1 ^a vocale, "X" (5.)

Caratteri indicativi del nome

I successivi tre caratteri che compongono il codice fiscale (posizione 4,5,6) sono ricavati dal NOME considerato in maiuscolo e senza eventuali spazi, apostrofi e accenti e trattini. Le regole che presidono al calcolo dei caratteri in questione possono essere sinteticamente espresse nelle seguenti 10 condizioni:

1.	[Nr. consonanti ≥ 4]	→ 1 ^a , 3 ^a , 4 ^a consonante
2.	[Nr. consonanti = 3]	→ 1 ^a , 2 ^a , 3 ^a consonante
3.	[Nr. consonanti = 2] [Nr. vocali ≥ 1]	→ 1 ^a , 2 ^a consonante, 1 ^a vocale
4.	[Nr. consonanti = 2] [Nr. vocali = 0]	→ 1 ^a , 2 ^a consonante, "X"
5.	[Nr. consonanti = 1] [Nr. vocali ≥ 2]	→ 1 ^a consonante, 1 ^a vocale, 2 ^a vocale
6.	[Nr. consonanti = 1] [Nr. vocali = 1]	→ 1 ^a consonante, 1 ^a vocale, "X"
7.	[Nr. consonanti = 1] [Nr. vocali = 0]	→ 1 ^a consonante, "X", "X"
8.	[Nr. consonanti = 0] [Nr. vocali ≥ 3]	→ 1 ^a , 2 ^a , 3 ^a vocale
9.	[Nr. consonanti = 0] [Nr. vocali = 2]	→ 1 ^a , 2 ^a vocale, "X"
10.	[Nr. consonanti = 0] [Nr. vocali = 1]	→ 1 ^a vocale, "X", "X"

Anche in questo caso l'applicazione delle regole è molto semplice, di seguito si illustrano alcuni esempi:

NOME	LETTERE PER IL CODICE FISCALE	METODO UTILIZZATO (CONDIZIONE)
Mariangela	MNG	1 ^a , 3 ^a , 4 ^a consonante (1.)
Carlo	CRL	1 ^a , 2 ^a , 3 ^a consonante (2.)
Carlo Maria	CLM	1 ^a , 3 ^a , 4 ^a consonante (1.)
Luca	LCU	1 ^a , 2 ^a consonante, 1 ^a vocale (3.)
Zoe	ZOE	1 ^a consonante, 1 ^a vocale, 2 ^a vocale (5.)
Li	LIX	1 ^a consonante, 1 ^a vocale, "X" (6.)

Per dare una maggiore estensione al sistema di calcolo dei codici fiscali si è provveduto a ricodificare alcune lettere straniere che compongono il nominativo della persona e che tuttavia non fanno parte dell'alfabeto italiano. A tale scopo ci si è attenuti alle regole fissate dal Ministero delle finanze in materia di nominativi stranieri. Pertanto, durante il processo di calcolo dei caratteri distintivi, rispettivamente del cognome e del nome, sono stati adottate le seguenti ricodifiche:

- La lettera "Š" dell'alfabeto sloveno è sostituita dalla lettera "S";
- La lettera "Ñ" dell'alfabeto spagnolo è sostituita dalla lettera "N";
- La lettera "Ç" dell'alfabeto francese è sostituita dalla lettera "C";
- La lettera "Ä" dell'alfabeto tedesco è sostituita dalle lettere "AE";
- La lettera "Ö" dell'alfabeto tedesco è sostituita dalle lettere "OE";
- La lettera "Ü" dell'alfabeto tedesco è sostituita dalle lettere "UE";
- La lettera "ß" dell'alfabeto tedesco è sostituita dalle lettere "SS";
- La lettera "Â" è sostituita dalla lettera "C";
- La lettera "Å" è sostituita dalla lettera "S";
- La lettera "Ă" è sostituita dalla lettera "Z";

Caratteri indicativi delle data di nascita e del sesso

I caratteri indicativi della data di nascita e del sesso delle persona fisica compongono i successivi 5 *digit* del codice fiscale (posizione 7, 8, 9, 10, 11). In particolare i primi due caratteri corrispondono alle ultime due cifre dell'anno di nascita, ad esempio l'anno 1977 è riportato come 77, mentre i nati del 1902, al pari dei nati dell'anno 2002, sono indicati con 02, ecc. Il terzo carattere che forma la cinquina in esame è rappresentato da un lettera che tiene conto del mese di nascita secondo la seguente tavola di raccordo:

MESE DI NASCITA	LETTERA PER LA 9 ^a POSIZIONE
Gennaio	A
Febbraio	B
Marzo	C
Aprile	D
Maggio	E
Giugno	H
Luglio	L
Agosto	M
Settembre	P
Ottobre	R
Novembre	S
Dicembre	T

Infine gli ultimi due caratteri dei 5 *digit* analizzati corrispondono esattamente al giorno di nascita della persona, qualora di sesso maschile, oppure al giorno di nascita incrementato di 40 unità

quando ci si riferisce a persona di genere femminile. In particolare, per gli uomini le due cifre saranno comprese tra 01 e 31, con l'avvertenza di aggiungere uno zero non significativo per i numeri inferiori a dieci. Mentre per le donne le due cifre in questione saranno comprese tra 41 e 71.

Caratteri indicativi del comune italiano o stato estero di nascita (codici territoriali)

I caratteri indicativi del luogo di nascita, sia esso comune italiano o stato estero, formano una quaterna (posizione 12, 13, 14, 15) e sono costituiti da un carattere alfabetico seguito da tre caratteri numerici. I valori di tale quaterna variano, dunque, in funzione del comune italiano o stato estero di nascita e sono tabulati nel volume “Codice dei Comuni d’Italia” o nel volume “Codice degli Stati esteri”, redatti a cura della Direzione generale del catasto e dei servizi tecnici erariali. In particolare per i comuni italiani è stato realizzato un archivio storico integrato, comprendente gli 8.101 comuni italiani presenti alla data del 1° gennaio 2006¹ e i 2.112 comuni soppressi², che combina le informazioni dell’elenco dei comuni Istat (denominazione e codice Istat) con le informazioni della banca dati dell’Agenzia del territorio – Comuni d’Italia (denominazione e codice territoriale)³. In definitiva l’archivio finale riporta le informazioni sui 10.207 comuni (esistenti / soppressi) per i quali è stato possibile stabilire una corrispondenza tra i due archivi. Per sei comuni, inclusi nell’elenco dei comuni soppressi, non è stato possibile stabilire alcuna corrispondenza e sono stati pertanto scartati. La predisposizione di tale archivio integrato rappresenta un passo indispensabile per l’aggancio delle informazioni di Sigif sul luogo di nascita (denominazione e codice Istat relativo ai comuni italiani) con le informazioni dell’Agenzia del territorio (codice territoriale comuni italiani). Per gli stati esteri è stata utilizzata una tabella di corrispondenza (<codici_Istat> e <codici_territoriali> Agenzia del territorio) elaborata dalla Ragioneria Generale dello Stato.

Carattere di controllo

Il carattere di controllo rappresenta il sedicesimo ed ultimo carattere identificativo del codice fiscale ed è costituito da una lettera. Esso è in funzione dei primi 15 caratteri ed è determinato nel modo seguente:

- Si assegna ai caratteri in posizione dispari i valori corrispondenti contenuti nella seguente tabella:

CARATTERE	VALORE	CARATTERE	VALORE
A o 0	1	N	20
B o 1	0	O	11
C o 2	5	P	3
D o 3	7	Q	6
E o 4	9	R	8
F o 5	13	S	12
G o 6	15	T	14
H o 7	17	U	16
I o 8	19	V	10
J o 9	21	W	22
K	2	X	25
L	4	Y	24
M	18	Z	23

¹ Essa fa riferimento all’elenco dei comuni italiani al 1° gennaio 2006. L’Istat attribuisce ad ogni comune esistente un codice composto da sei cifre. Le prime tre cifre del codice identificano la provincia di appartenenza del comune.

² Nel 1999 è stato attribuito un codice nell’ambito della provincia di ultima appartenenza anche ai comuni soppressi prima dell’anno 1966 e ai comuni ceduti ad altra nazione documentati negli archivi dell’Istat.

³ La banca dati fa riferimento a due file DBF rispettivamente per i comuni d’Italia (aggiornato al 25 novembre 2005) e per gli Stati esteri (aggiornato allo stesso anno).

- Si assegna ai caratteri in posizione pari i valori corrispondenti contenuti nella seguente tabella:

CARATTERE	VALORE	CARATTERE	VALORE
A o 0	0	N	13
B o 1	1	O	14
C o 2	2	P	15
D o 3	3	Q	16
E o 4	4	R	17
F o 5	5	S	18
G o 6	6	T	19
H o 7	7	U	20
I o 8	8	V	21
J o 9	9	W	22
K	10	X	23
L	11	Y	24
M	12	Z	25

- Si sommano tutti i valori corrispondenti ai 16 caratteri (distinti per posizione pari o dispari);
- Si divide la somma di tutti valori per 26 e si considera il resto di tale divisione;
- Il resto della divisione viene convertito in “lettera” secondo la seguente tabella:



RESTO DELLA DIVISIONE	LETTERA	RESTO DELLA DIVISIONE	LETTERA
0	A	13	N
1	B	14	O
2	C	15	P
3	D	16	Q
4	E	17	R
5	F	18	S
6	G	19	T
7	H	20	U
8	I	21	V
9	J	22	W
10	K	23	X
11	L	24	Y
12	M	25	Z

Metodi e Norme - Nuova serie - Volumi pubblicati

Anno 2000

6. *L'indice del costo della vita valevole ai fini dell'applicazione della scala mobile delle retribuzioni. Dalle origini alla cessazione (1945-97)*
7. *Le nuove stime dei consumi finali delle famiglie secondo il Sistema Europeo dei Conti SEC95*

Anno 2001

8. *La nuova indagine sulle cause di morte. La codifica automatica, il bridge coding e altri elementi innovativi*
9. *Il settore delle costruzioni in contabilità nazionale. I nuovi standard europei dettati dal SEC95*
10. *Indagini sociali telefoniche. Metodologia ed esperienze della statistica ufficiale*
11. *Elenco dei comuni al 31 maggio 2001* 
12. *Classificazione delle professioni* 

Anno 2002

13. *Le statistiche culturali in Europa*
14. *Gli investimenti lordi di contabilità nazionale dopo la revisione: nota metodologica*
15. *Panel Europeo sulle famiglie*


Anno 2003

16. *Metodi statistici per il record linkage*
17. *Metodologia e organizzazione dell'indagine multiscopo sulla domanda turistica "Viaggi e vacanze"*
18. *Classificazione delle attività economiche. Ateco 2002*


Anno 2004

19. *Inventario sulle fonti e metodi di calcolo per le valutazioni a prezzi costanti - Italia*
20. *Metodologia e tecniche di tutela della riservatezza nel rilascio di informazione statistica*
21. *Metodologia di stima degli aggregati di contabilità nazionale a prezzi correnti*
22. *Numeri indici dei prezzi alla produzione dei prodotti industriali venduti sul mercato interno - Base 2000=100*



Anno 2005

23. *I conti economici nazionali per settore istituzionale: le nuove stime secondo il Sec 95* 
24. *La rete di intervistatori Capi dell'Istat per la conduzione dell'indagine continua sulle Forze di Lavoro*
25. *Il monitoraggio del processo e la stima dell'errore nelle indagini telefoniche*
26. *Classificazione delle forme giuridiche delle unità legali*

Anno 2006

27. *Gli stranieri nella rilevazione continua sulle forze di lavoro*
28. *L'indagine campionaria sulle nascite: obiettivi, metodologia e organizzazione*
29. *Rilevazione mensile sull'occupazione, gli orari di lavoro e le retribuzioni nelle grandi imprese*
30. *La classificazione Istat dei titoli di studio italiani. Anno 2003* 
31. *Il sistema di indagini sociali multiscopo. Contenuti e metodologia delle indagini*
32. *La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione*
33. *Il calcolo della spesa pubblica per la protezione dell'ambiente - Linee guida per riclassificare i rendiconti delle amministrazioni pubbliche*

Anno 2007


34. *Come si progetta il monitoraggio del lavoro sul campo di un'indagine sulle famiglie* 
35. *Istruzioni integrative per l'applicazione della Icd-10 nella codifica delle cause di morte* 


Anno 2008

36. *La progettazione e lo sviluppo informatico del sistema CAPI sulle forze di lavoro*
37. *L'indagine europea sui redditi e le condizioni di vita delle famiglie (Eu-Silc)*

Anno 2009

38. *Integrazione di dati campionari Eu-Silc con dati di fonte amministrativa*

 dati forniti su floppy disk

 dati forniti su cd-rom



***Produzione editoriale
&
Altri servizi***

La produzione editoriale

LE PUBBLICAZIONI A CARATTERE GENERALE

Annuario statistico italiano 2008
pp. XXIV+840+1 cd-rom; € 50,00
ISBN 978-88-458-1595-9

Atti - Ottava Conferenza Nazionale di Statistica
Roma, 28 e 29 novembre 2006
pp. 160+1 cd-rom; € 22,00
ISBN 978-88-458-1569-0

Bollettino mensile di statistica
pp. 168 circa; € 15,00
ISSN 0021-3136

Compendio statistico italiano 2007
(con appendice in inglese)
pp. 344; € 15,00
ISBN 978-88-458-1568-3

Italian Statistical Abstract 2006
pp. 380; € 15,00
ISBN 978-88-458-1541-6

Rapporto annuale
La situazione del Paese nel 2007
pp. XXXII+420; € 30,00
ISBN 978-88-458-1585-0
ISSN 1594-3135

Rapporto annuale
La situazione del Paese nel 2007
pp. XXXII+420+1 cd-rom; € 35,00
ISBN 978-88-458-1584-3
ISSN 1594-3135

LE NOVITÀ EDITORIALI A CARATTERE TEMATICO

AMBIENTE E TERRITORIO

Atlante statistico territoriale delle infrastrutture
Indicatori statistici, n. 6, ediz. 2008
pp. 272+1 cd-rom; € 28,00
ISBN 978-88-458-1580-5

Statistiche ambientali
Annuari, n. 10, edizione 2008
pp. 618+1 cd-rom; € 50,00
ISBN 978-88-458-1591-1

POPOLAZIONE

Evoluzione e nuove tendenze dell'instabilità coniugale (*)
Argomenti, n. 34, edizione 2008
pp. 164; € 18,00
ISBN 978-88-458-1582-9

Popolazione e movimento anagrafico dei comuni
anno 2005
Annuari, n. 18, edizione 2008
pp. 236+1 cd-rom; € 28,00
ISBN 978-88-458-1578-2

La presenza straniera in Italia: l'accertamento e l'analisi
Atti del Convegno dicembre 2005
pp. 216; € 18,00
ISBN 978-88-458-1603-1

SANITÀ E PREVIDENZA

I bilanci consuntivi degli enti previdenziali (*)
anno 2006
Informazioni, n. 5, edizione 2008
pp. 104+1 cd-rom; € 20,00
ISBN 978-88-458-1598-0

Istruzioni integrative per l'applicazione dell'Icd-10 nella codifica delle cause di morte
Metodi e norme, n. 35, edizione 2007
pp. 204+1 cd-rom; € 25,00
ISBN 978-88-458-1570-6

Statistiche della previdenza e dell'assistenza sociale (*)
I - I trattamenti pensionistici anno 2005
Annuari, n. 9, edizione 2007
pp. 128+1 cd-rom; € 18,00
ISBN 978-88-458-1553-9

Statistiche della previdenza e dell'assistenza sociale (*)
II - I beneficiari delle prestazioni pensionistiche - Anno 2005
Annuari, n. 10, edizione 2007
pp. 124+1 cd-rom; € 18,00
ISBN 978-88-458-1574-4

Stime preliminari della mortalità per causa nelle regioni italiane
anno 2004
Informazioni, n. 1, edizione 2007
pp. 116+1 cd-rom; € 19,00
ISBN 978-88-458-1381-8

CULTURA

Spettacoli, musica e altre attività del tempo libero (*) - Anno 2006
Informazioni, n. 6, edizione 2008
pp. 228+1 cd-rom; € 28,00
ISBN 978-88-458-1599-7

Statistiche culturali - Anno 2006
Annuari, n. 46, edizione 2008
pp. 164+1 cd-rom; € 23,00
ISBN 978-88-458-1571-3

L'uso dei media e del cellulare in Italia (*) - Anno 2006
Informazioni, n. 2, edizione 2008
pp. 292+1 cd-rom; € 28,00
ISBN 978-88-458-1579-9

FAMIGLIA E SOCIETÀ

Come si progetta il monitoraggio del lavoro sul campo di un'indagine sulle famiglie
Metodi e norme, n. 34, edizione 2007
pp. 332+1 cd-rom; € 32,00
ISBN 978-88-458-1566-9

Conciliare lavoro e famiglia (*)
Una sfida quotidiana
Argomenti, n. 33, edizione 2008
pp. 264; € 22,00
ISBN 978-88-458-1573-7



I consumi delle famiglie

anno 2006

Annuari, n. 13, edizione 2008

pp. 176+1 cd-rom; € 23,00

ISBN 978-88-458-1586-7

Evoluzione e nuove tendenze dell'instabilità coniugale (*)

Argomenti, n. 34, edizione 2008

pp. 164; € 18,00

ISBN 978-88-458-1582-9

L'indagine europea sui redditi e le condizioni di vita delle famiglie (Eu-Silc)

Metodi e norme, n. 37, edizione 2008

pp. 188; € 18,00

ISBN 978-88-458-1596-6

Spettacoli, musica e altre attività del tempo libero (*) - Anno 2006

Informazioni, n. 6, edizione 2008

pp. 228+1 cd-rom; € 28,00

ISBN 978-88-458-1599-7

I tempi della vita quotidiana

Un approccio multidisciplinare

all'analisi dell'uso del tempo

Argomenti, n. 32, edizione 2007

pp. 416; € 32,00

ISBN 978-88-458-1554-6

Time Use in Daily Life

A Multidisciplinary Approach to the Time Use's Analysis

Argomenti, n. 35, edizione 2008

pp. 332; € 30,00

ISBN 978-88-458-1587-4

L'uso dei media e del cellulare in Italia (*) - Anno 2006

Informazioni, n. 2, edizione 2008

pp. 292+1 cd-rom; € 28,00

ISBN 978-88-458-1579-9

I viaggi in Italia e all'estero nel 2005 (*)

Informazioni, n. 3, edizione 2008

pp. 96+1 cd-rom; € 15,00

ISBN 978-88-458-1581-2

La vita quotidiana nel 2006

Informazioni, n. 12, edizione 2007

pp. 300+1 cd-rom; € 27,00

ISBN 978-88-458-1565-2

PUBBLICA AMMINISTRAZIONE

I bilanci consuntivi degli enti previdenziali (*)

anno 2006

Informazioni, n. 5, edizione 2008

pp. 104+1 cd-rom; € 20,00

ISBN 978-88-458-1598-0

Le cooperative sociali in Italia

anno 2005

Informazioni, n. 4, edizione 2008

pp. 144+1 cd-rom; € 22,00

ISBN 978-88-458-1588-1

Finanza locale: entrate e spese dei bilanci consuntivi (comuni, province e regioni)

anno 2005

Annuari, n. 12, edizione 2008

pp. 128+1 cd-rom; € 20,00

ISBN 978-88-458-1593-5

Statistiche della previdenza e dell'assistenza sociale (*)

I - I trattamenti pensionistici

anno 2005

Annuari, n. 9, edizione 2007

pp. 128+1 cd-rom; € 18,00

ISBN 978-88-458-1553-9

Statistiche della previdenza e dell'assistenza sociale (*)

II - I beneficiari delle prestazioni

pensionistiche - Anno 2005

Annuari, n. 10, edizione 2007

pp. 124+1 cd-rom; € 18,00

ISBN 978-88-458-1574-4

GIUSTIZIA E SICUREZZA

Evoluzione e nuove tendenze dell'instabilità coniugale (*)

Argomenti, n. 34, edizione 2008

pp. 164; € 18,00

ISBN 978-88-458-1582-9

CONTI ECONOMICI

Contabilità nazionale Conti economici nazionali Anni 1995-2006

Annuari, n. 11, edizione 2008

pp. 332+1 cd-rom; € 32,00

ISBN 978-88-458-1575-1

Valore aggiunto ai prezzi di base dell'agricoltura per regione

anni 2002-2007

Informazioni, n. 9, edizione 2008

pp. 200+1 cd-rom; € 23,00

ISBN 978-88-458-1602-4

LAVORO

Conciliare lavoro e famiglia (*)

Una sfida quotidiana

Argomenti, n. 33, edizione 2008

pp. 264; € 22,00

ISBN 978-88-458-1573-7

Forze di lavoro - Media 2006

Annuari, n. 12, edizione 2007

pp. 208+1 cd-rom; € 27,00

ISBN 978-88-458-1557-7

La progettazione e lo sviluppo informatico del sistema Capi sulle forze di lavoro

Metodi e norme, n. 36, edizione 2008

pp. 100; € 15,00

ISBN 978-88-458-1594-2

Statistiche della previdenza e dell'assistenza sociale (*)

I - I trattamenti pensionistici

anno 2005

Annuari, n. 9, edizione 2007

pp. 128+1 cd-rom; € 18,00

ISBN 978-88-458-1553-9

Statistiche della previdenza e dell'assistenza sociale (*)

II - I beneficiari delle prestazioni

pensionistiche - Anno 2005

Annuari, n. 10, edizione 2007

pp. 124+1 cd-rom; € 18,00

ISBN 978-88-458-1574-4

PREZZI

Il valore della moneta in Italia dal 1861 al 2007

Informazioni, n. 8, edizione 2008

pp. 170; € 18,00

ISBN 978-88-458-1601-7

AGRICOLTURA E ZOOTECNIA

Donne della terra: i loro "numeri" per e nell'agricoltura

Atti del Convegno

gennaio 2006

pp. 352; € 27,00

ISBN 978-88-458-1597-3

Le Statistiche agricole verso il Censimento del 2010: valutazioni e prospettive

Atti del Convegno

ottobre 2006

pp. 456; € 33,00

ISBN 978-88-458-1592-8

INDUSTRIA E SERVIZI

Statistiche dei trasporti

anno 2004

Annuari, n. 5, edizione 2007

pp. 280; € 22,00

ISBN 978-88-458-1543-0



Statistiche dei trasporti marittimi
anni 2002-2004
Informazioni, n. 5, edizione 2007
pp. 56+1 cd-rom; € 15,00
ISBN 978-88-458-1545-4

Statistiche del trasporto aereo
anni 2004-2005
Informazioni, n. 6, edizione 2007
pp. 44+1 cd-rom; € 15,00
ISBN 978-88-458-1547-8

Statistiche del trasporto ferroviario
anni 2004-2005
Informazioni, n. 14, edizione 2007
pp. 48+1 cd-rom; € 15,00
ISBN 978-88-458-1572-0

Statistiche sui permessi di costruire
anni 2003-2004
Informazioni, n. 7, edizione 2007
pp. 36+1 cd-rom; € 15,00
ISBN 978-88-458-1551-5

**Statistiche sull'innovazione
nelle imprese**
anni 2002-2004
Informazioni, n. 1, edizione 2008
pp. 192; € 18,00
ISBN 978-88-458-1577-5

**I viaggi in Italia e all'estero
nel 2005 (*)**
Informazioni, n. 3, edizione 2008
pp. 96+1 cd-rom; € 15,00
ISBN 978-88-458-1581-2

► **COMMERCIO ESTERO**

**Commercio estero e attività
internazionali delle imprese 2007**
1. Merci, servizi, investimenti diretti
2. Paesi, settori, regioni
+ *L'Italia nell'economia internazionale*
Rapporto ICE 2007-2008 + 1 cd-rom
+ Sintesi del Rapporto ICE
Annuari, n. 10, edizione 2008
pp. 348+440+590+48
€ 100,00 (in cofanetto)
ISBN 978-88-458-1589-8

▼ **PRODOTTI CENSUARI**

5° CENSIMENTO GENERALE DELL'AGRICOLTURA 22 OTTOBRE 2000

Caratteristiche strutturali delle aziende agricole
Fascicolo nazionale; € 25,00
Fascicoli regionali; € 22,00
Fascicoli provinciali; € 22,00

Caratteristiche tipologiche delle aziende agricole
Fascicolo nazionale; € 20,00
Fascicoli regionali; € 14,00

**Organizzazione e atti del 5° Censimento generale
dell'agricoltura**
pp. 152; € 14,00; ISBN 88-458-1287-1

APPROFONDIMENTI TEMATICI

La coltivazione della vite in Italia
Volume I - Caratteristiche generali
pp. 300; € 26,50; ISBN 88-458-1280-4
Volume II - Vitigni
pp. 248; € 22,00; ISBN 88-458-1281-2

La donna in agricoltura
pp. 316; € 14,00; ISBN 88-458-1284-7

Le imprese agricole
pp. 338; € 22,00; ISBN 88-458-1283-9

Le infrastrutture delle aziende agricole
pp. 150; € 11,50; ISBN 88-458-1279-0

La zootecnia in Italia
pp. 380; € 26,50; ISBN 88-458-1282-0

14° CENSIMENTO GENERALE DELLA POPOLAZIONE E DELLE ABITAZIONI - 21 OTTOBRE 2001

Primi risultati
pp. 300+1 cd-rom; € 25,00; ISBN 88-458-0689-8

Popolazione legale
pp. 312+1 cd-rom; € 27,00; ISBN 88-458-1069-0

**Struttura demografica e familiare della popolazione
residente - Italia**
pp. 294+1 cd-rom; € 32,00; ISBN 88-458-1388-6

**Struttura socioeconomica e condizione abitativa della
popolazione residente - Italia**
pp. 226+1 cd-rom; € 25,00; ISBN 88-458-1538-2

Popolazione residente e abitazioni nelle province italiane
Fascicoli provinciali; +1 cd-rom; € 19,00/27,00

Popolazione residente e abitazioni nelle regioni italiane
Fascicoli regionali; +1 cd-rom; € 27,00

Popolazione residente e abitazioni nei grandi comuni italiani
+1 cd-rom; € 17,00/19,00

I documenti - Conoscere il censimento
pp. 628+1 cd-rom; € 50,00; ISBN 88-458-1536-6

**Il Piano di rilevazione e il Sistema di produzione -
Conoscere il censimento**
pp. 450; € 32,00; ISBN 88-458-1539-0

APPROFONDIMENTI TEMATICI

Gli stranieri in Italia: analisi dei dati censuari
pp. 248; € 20,00; ISBN 88-458-1537-4

8° CENSIMENTO GENERALE DELL'INDUSTRIA E DEI SERVIZI - 22 OTTOBRE 2001

Imprese, istituzioni e unità locali
Fascicolo nazionale; € 31,50
Fascicoli regionali; € 22,00
Fascicoli provinciali; € 14,00

Distretti industriali e sistemi locali del lavoro 2001
pp. 158+1 cd-rom; € 17,00; ISBN 88-458-1505-6

Atti
pp. 262; € 11,50; ISBN 88-458-1278-2

Acquisto pubblicazioni



Inviare questo modulo via fax al numero **06.4673.3477** oppure spedire in **busta chiusa** a:
Istituto Nazionale di Statistica, DCCE, Commercializzazione e attività promozionali
Via Cesare Balbo, 16 – 00184 Roma

Per ulteriori informazioni telefonare al numero 06 4673.3267

Desidero ricevere le seguenti pubblicazioni

Titolo	Codice ISBN	Prezzo
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

Per un totale di euro _____

Eventuale sconto ⁽¹⁾ euro _____

Contributo spese di spedizione euro _____ **6,00** _____

Importo da pagare euro _____

⁽¹⁾ Il Sistan, gli Enti pubblici, le Biblioteche e le Università usufruiscono di uno **sconto del 10%** se acquistano direttamente dall'Istat. Per tutti gli utenti che acquistano oltre 20 volumi è previsto uno **sconto del 20%**.

DATI PER LA FATTURAZIONE

Ente/Cognome e Nome _____

Referente _____

Cod.fiscale* | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | P.IVA* | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ |

Indirizzo _____ Cap _____ Città _____

Prov. _____ tel. _____ fax _____ e-mail _____

* è necessario inserire sia il codice fiscale sia la partita IVA

DESTINATARIO DEI PRODOTTI (se diverso da quanto indicato nei dati per la fatturazione)

Ente/Cognome e Nome _____

Indirizzo _____ Cap _____ Città _____

Prov. _____ tel. _____ fax _____ e-mail _____

MODALITÀ DI PAGAMENTO. L'importo dovrà essere versato dall'acquirente, dopo il ricevimento della fattura, sul c/c postale n. 619007, oppure con bonifico bancario c/o la Banca Nazionale del Lavoro, indicando con chiarezza il numero, la data della fattura e il codice cliente. Per i versamenti tramite bonifico bancario le coordinate sono: c/c n. 218050, ABI 01005.8, CAB 03382.9; via swift: B.N.L.I. IT RR, codice CIN K, codice anagrafico 63999228/j; IBAN IT64K0100503382000000218050.

INFORMATIVA - I dati da Lei forniti saranno utilizzati esclusivamente per l'esecuzione dell'ordine e per l'invio, da parte dell'Istat, di promozioni commerciali, senza alcun impegno da parte Sua. Il trattamento dei dati avverrà nell'assoluto rispetto del d.lgs 196/2003, esclusivamente ad opera dei dipendenti dell'Istituto incaricati. Il titolare dei dati è l'Istituto nazionale di statistica, Via Cesare Balbo n. 16, 00184 Roma; responsabile del trattamento dei dati è il Direttore centrale per la comunicazione e la programmazione editoriale, anche per quanto riguarda l'esercizio dei diritti dell'interessato di cui all'articolo 7 del d.lgs. n. 196/2003. In qualsiasi momento potrà far modificare o cancellare i Suoi dati indirizzando la richiesta a Istat, DCCE, Commercializzazione e attività promozionali, Via Cesare Balbo n. 16, 00184 Roma, oppure via e-mail all'indirizzo marketing@istat.it, o inviando un fax al numero 064673.3477

Data _____

Firma _____

PV09

Altri prodotti e servizi

ABBONAMENTI 2009

L'abbonamento consente di disporre di tutte le informazioni relative al settore tematico prescelto, diffuse attraverso le pubblicazioni edite nel 2009, accompagnate, ove previsto, da supporto informatico. Gli abbonati riceveranno per posta i prodotti che saranno via via pubblicati nel/i settore/i prescelto/i, editi nell'anno di sottoscrizione dell'abbonamento. Oltre all'abbonamento ai singoli settori editoriali è prevista la modalità di abbonamento "Tutti i settori", che prevede l'invio di tutta la produzione editoriale dell'anno ad esclusione dei volumi appartenenti alle collane Tecniche e strumenti, Essays, Quaderni del Mipa e Censimenti.

L'abbonamento all'area "Generale", infine, comprende 11 numeri del *Bollettino mensile di statistica* e l'*Annuario statistico italiano*. Tutti coloro che sottoscriveranno un abbonamento anche a un solo settore riceveranno una copia del *Rapporto annuale*.

Per meglio comprendere il sistema degli abbonamenti è possibile visionare, sul sito www.istat.it, l'elenco 2007 e l'elenco 2008 delle pubblicazioni inviate agli abbonati alle edizioni 2007 e 2008.

Il coupon è scaricabile dal sito www.istat.it/SERVIZI/ABBONAMENTI.

Ulteriori informazioni possono essere richieste a:

ISTAT - Direzione centrale per la comunicazione e la programmazione editoriale - PED/D

Via Cesare Balbo, 16 - 00184 ROMA - Tel. 06.4673.3278-3280 - Fax 06.4673.3477 oppure 06.4667.8240

e-mail: marketing@istat.it

WWW.ISTAT.IT

Nel sito Internet è possibile informarsi sulla produzione editoriale più recente, richiedere prodotti e servizi offerti dall'Istat, leggere e prelevare i comunicati stampa, accedere alle banche dati, collegarsi con altri siti nazionali e internazionali.

CATALOGO ON LINE

Dalla home page del sito Internet è possibile collegarsi con il catalogo on line, che contiene l'elenco completo delle pubblicazioni edite dall'Istat a partire dall'anno 2000.

Attraverso questo utile strumento è possibile effettuare la ricerca del volume per titolo, per settore, per collana, per anno di edizione e per codice ISBN. Ogni pubblicazione è presentata attraverso una scheda che riporta, oltre alle caratteristiche tecniche, anche una breve descrizione del prodotto. Molti dei volumi presenti in questo catalogo sono scaricabili gratuitamente.

CONT@CT CENTRE

Dal sito Internet è possibile ricevere informazioni su dati e pubblicazioni Istat, avere assistenza nella ricerca delle statistiche ufficiali europee e supporto nella individuazione delle metodologie e classificazioni ufficiali comunitarie (Eurostat). Solo dopo essersi registrati compilando l'apposito *form* è possibile richiedere i seguenti servizi: certificazioni prezzi e retribuzioni, dati elementari per uffici Sistan, collezioni campionarie di dati elementari (file standard), dati censuari e cartografici, abbonamenti e dati del commercio estero, ricerche storiche e bibliografiche, elaborazioni personalizzate.

Inoltre ai giornalisti è dedicata un'area speciale per rispondere alle richieste di dati, pubblicazioni e approfondimenti su particolari tematiche.

Abbonamenti 2009



Inviare questo modulo via fax al numero **06.4673.3477** oppure spedire in **busta chiusa** a:
Istituto Nazionale di Statistica, DCCE, Commercializzazione e attività promozionali
Via Cesare Balbo, 16 – 00184 Roma

Per ulteriori informazioni telefonare ai numeri 06 4673.3278-3280

Desidero sottoscrivere i seguenti abbonamenti per l'anno 2009	ITALIA	ESTERO
Generale (Bollettino mensile di statistica, Annuario statistico italiano)	<input type="checkbox"/> euro 180,00	<input type="checkbox"/> euro 200,00
Ambiente e territorio	<input type="checkbox"/> euro 70,00	<input type="checkbox"/> euro 80,00
Popolazione (escluso Censimenti)	<input type="checkbox"/> euro 50,00	<input type="checkbox"/> euro 55,00
Sanità e previdenza	<input type="checkbox"/> euro 70,00	<input type="checkbox"/> euro 80,00
Cultura	<input type="checkbox"/> euro 50,00	<input type="checkbox"/> euro 55,00
Famiglia e società	<input type="checkbox"/> euro 140,00	<input type="checkbox"/> euro 150,00
Pubblica amministrazione	<input type="checkbox"/> euro 80,00	<input type="checkbox"/> euro 90,00
Giustizia e sicurezza	<input type="checkbox"/> euro 40,00	<input type="checkbox"/> euro 45,00
Conti economici	<input type="checkbox"/> euro 70,00	<input type="checkbox"/> euro 80,00
Lavoro	<input type="checkbox"/> euro 80,00	<input type="checkbox"/> euro 90,00
Prezzi	<input type="checkbox"/> euro 40,00	<input type="checkbox"/> euro 45,00
Agricoltura e zootecnia(escluso Censimenti)	<input type="checkbox"/> euro 40,00	<input type="checkbox"/> euro 45,00
Industria e servizi (escluso Censimenti)	<input type="checkbox"/> euro 50,00	<input type="checkbox"/> euro 55,00
Commercio estero	<input type="checkbox"/> euro 80,00	<input type="checkbox"/> euro 90,00
Tutti i settori (escluso Censimenti)	<input type="checkbox"/> euro 800,00	<input type="checkbox"/> euro 900,00

Qualunque abbonamento, anche a un solo settore, comprende l'invio di una copia del "Rapporto annuale"

Per un totale di _____
 Eventuale sconto ⁽¹⁾ _____
 Importo da pagare _____

⁽¹⁾ il Sistan, gli Enti pubblici, le Biblioteche e le Università usufruiscono di uno sconto del 10% soltanto se sottoscrivono l'abbonamento direttamente con l'Istat.

DATI PER LA FATTURAZIONE

Ente/Cognome e Nome _____
 Referente _____
 Cod.fiscale* | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | P.IVA* | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ | _ |
 Indirizzo _____ Cap _____ Città _____
 Prov. _____ tel. _____ fax _____ e-mail _____

* è necessario inserire sia il codice fiscale sia la partita IVA

DESTINATARIO DEI PRODOTTI (se diverso da quanto indicato nei dati per la fatturazione)

Ente/Cognome e Nome _____
 Indirizzo _____ Cap _____ Città _____
 Prov. _____ tel. _____ fax _____ e-mail _____

MODALITÀ DI PAGAMENTO. L'importo dovrà essere versato dall'acquirente, dopo il ricevimento della fattura, sul c/c postale n. 619007, oppure con bonifico bancario c/o la Banca Nazionale del Lavoro, indicando con chiarezza il numero, la data della fattura e il codice cliente. Per i versamenti tramite bonifico bancario le coordinate sono: c/c n. 218050, ABI 01005.8, CAB 03382.9; via swift: B.N.L.I. IT RR, codice CIN K, codice anagrafico 63999228/j; IBAN IT64K0100503382000000218050.

INFORMATIVA - I dati da Lei forniti saranno utilizzati esclusivamente per l'esecuzione dell'ordine e per l'invio, da parte dell'Istat, di promozioni commerciali, senza alcun impegno da parte Sua. Il trattamento dei dati avverrà nell'assoluto rispetto del d.lgs 196/2003, esclusivamente ad opera dei dipendenti dell'Istituto incaricati. Il titolare dei dati è l'Istituto nazionale di statistica, Via Cesare Balbo n. 16, 00184 Roma; responsabile del trattamento dei dati è il Direttore centrale per la comunicazione e la programmazione editoriale, anche per quanto riguarda l'esercizio dei diritti dell'interessato di cui all'articolo 7 del d.lgs. n. 196/2003. In qualsiasi momento potrà far modificare o cancellare i Suoi dati indirizzando la richiesta a Istat, DCCE, Commercializzazione e attività promozionali, Via Cesare Balbo n. 16, 00184 Roma, oppure via e-mail all'indirizzo marketing@istat.it, o inviando un fax al numero 064673.3477

Data _____

Firma _____

PV09

PIÙ INFORMAZIONI. PIÙ VICINE A VOI.

I Centri d'informazione statistica

Per darvi più servizi e per esservi più vicino l'Istat ha aperto al pubblico una rete di Centri d'informazione statistica che copre l'intero territorio nazionale. Oltre alla vendita di prodotti informatici e pubblicazioni, i Centri rilasciano certificati sull'indice dei prezzi, offrono informazioni tramite collegamenti con le banche dati del Sistema statistico nazionale (Sistan) e dell'Eurostat (Ufficio di statistica della Comunità europea), forniscono elaborazioni statistiche "su misura" ed assistono i laureandi nella ricerca e selezione dei dati.

Presso i Centri d'informazione statistica, semplici cittadini, studenti, ricercatori, imprese e operatori della pubblica amministrazione troveranno assistenza qualificata e un facile accesso ai dati di cui hanno bisogno. D'ora in poi sarà più facile conoscere l'Istat e sarà più facile per tutti gli italiani conoscere l'Italia. Per gli orari di apertura al pubblico consultare il sito www.istat.it nella pagina "Servizi".

ANCONA Via Castelfidardo, 4
Telefono 071/5013011 Fax 071/5013085

BARI Piazza Aldo Moro, 61
Telefono 080/5789317 Fax 080/5789335

BOLOGNA Galleria Cavour, 9
Telefono 051/6566111 Fax 051/6566185-182

BOLZANO Via Canonico Michael Gamper, 1
Telefono 0471/418400 Fax 0471/418419

CAGLIARI Via Firenze, 17
Telefono 070/34998700-1 Fax 070/34998732-3

CAMPOBASSO Via G. Mazzini, 129
Telefono 0874/604854-8 Fax 0874/604885-6

CATANZARO Viale Pio X, 116
Telefono 0961/507629 Fax 0961/741240

FIRENZE Lungarno Cristoforo Colombo, 54
Telefono 055/6237711 Fax 055/6237735

GENOVA Via San Vincenzo, 4
Telefono 010/5849701 Fax 010/5849742

MILANO Via Porlezza, 12
Telefono 02/806132214 Fax 02/806132205

NAPOLI Via G. Verdi, 18
Telefono 081/4930190 Fax 081/4930185

PALERMO Via G. B. Vaccarini, 1
Telefono 091/6751811 Fax 091/6751836

PERUGIA Via Cesare Balbo, 1
Telefono 075/5826411 Fax 075/5826484

PESCARA Via Caduta del Forte, 34
Telefono 085/44120511-2 Fax 085/4216516

POTENZA Via del Popolo, 4
Telefono 0971/377261 Fax 0971/36866

ROMA Via Cesare Balbo, 11/a
Telefono 06/46733102 Fax 06/46733101

TORINO Via Alessandro Volta, 3
Telefono 011/5166758-64-67 Fax 011/535800

TRENTO Via Brennero, 316
Telefono 0461/497801 Fax 0461/497813

TRIESTE Via Cesare Battisti, 18
Telefono 040/6702558 Fax 040/6702599

VENEZIA-MESTRE Corso del Popolo, 23
Telefono 041/5070811 Fax 041/5070835

La Biblioteca centrale

È la più ricca biblioteca italiana in materia di discipline statistiche e affini. Il suo patrimonio, composto da oltre 500.000 volumi e 2.700 periodici in corso, comprende fonti statistiche e socio-economiche, studi metodologici, pubblicazioni periodiche degli Istituti nazionali di statistica di tutto il mondo, degli Enti internazionali e dei principali Enti e Istituti italiani ed esteri. È collegata con le principali banche dati nazionali ed estere. Il catalogo informatizzato della biblioteca è liberamente consultabile in rete sul sito Web dell'Istat alla voce Biblioteca (www.istat.it).

Oltre all'assistenza qualificata che è resa all'utenza in sede, è attivo un servizio di ricerche bibliografiche e di dati statistici a distanza, con l'invio dei risultati per posta o via fax, cui i cittadini, gli studenti, i ricercatori e le imprese possono accedere.

È a disposizione dell'utenza una sala di consultazione al secondo piano.

ROMA Via Cesare Balbo, 16 Telefono 06/4673.2380 Fax 06/4673.2617

<https://contact.istat.it/>

Orario: piano secondo da lunedì a giovedì 9.00 - 16.00 venerdì 9.00 - 14.00

Integrazione di dati campionari Eu-Silc con dati di fonte amministrativa

Il volume vuole essere uno strumento utile a chi si pone l'obiettivo di migliorare la qualità del dato attraverso tecniche di integrazione dei dati rilevati su più fonti, siano esse statistiche e/o amministrative. Sebbene la trattazione sia orientata al contesto della produzione dei dati sui redditi delle famiglie è tuttavia possibile ricavare utili linee guida applicabili anche in altri settori delle statistiche sociali. L'articolazione del processo di integrazione in fasi sequenziali - che inizia dall'analisi del fenomeno e dalla ricognizione delle fonti, si sviluppa attraverso la scelta della chiave per l'abbinamento, l'armonizzazione delle unità e delle variabili, l'analisi delle incoerenze e della completezza del dato, e termina con la riconciliazione dei valori tra le diverse fonti - rappresenta un modello universale che può, dunque, essere esteso a qualsiasi altra indagine. I capitoli illustrano gli aspetti concettuali, i requisiti tecnici e le soluzioni metodologiche adottate nelle varie fasi che precedono l'integrazione dei dati campionari Eu-Silc con i dati fonti amministrativa.

Combining Survey and Administrative Data in Eu-Silc Project

This book helps practitioners gain a deeper insight on the strategies used to improve data quality for both survey and administrative data, through integration techniques applied to the different data sources.

Even though a lot of the practical advice given in the publication is specifically addressed to data producers, it still remains a useful guideline for practitioners involved in all sectors of social statistics. The system perspective used includes all phases; from the analysis of the phenomenon and the sources, and is developed through the choice of the best matching key, the record linkage, the problem solving approach for harmonizing units and variables and the analysis of data incoherence and coverage and finishes with the reconciliation of values reported in the different sources. It represents a universal model, and therefore can be applied to any survey. The chapters discuss concepts and technical features and provide methodological solutions adopted by Eu-Silc, to combine survey data with administrative data.

ISBN 978-88-458-1612-3



1M01200903800000

€ 17,00