

LA NUOVA INDAGINE SULLE SPESE PER CONSUMI IN ITALIA





LA NUOVA INDAGINE SULLE SPESE PER CONSUMI IN ITALIA

ISBN 978-88-458-1856-1

© 2015
Istituto nazionale di statistica
Via Cesare Balbo, 16 - Roma

Salvo diversa indicazione la riproduzione è libera,
a condizione che venga citata la fonte.

Immagini, loghi (compreso il logo dell'Istat), marchi
registrati e altri contenuti di proprietà di terzi
appartengono ai rispettivi proprietari e non possono
essere riprodotti senza il loro consenso.



INDICE

Premessa	7
1. Il processo di ristrutturazione dell'indagine	9
1.1 Le linee guida europee per la rilevazione della spesa per consumi	9
1.2 La rilevazione della spesa per consumi in Italia	11
1.3 Il sistema delle sperimentazioni per una revisione della metodologia di indagine	15
1.3.1 <i>Alcune considerazioni preliminari</i>	18
1.3.2 <i>Il dettaglio delle voci di spesa</i>	20
1.3.3 <i>I riferimenti temporali</i>	21
1.3.3.1 <i>La definizione dei periodi di riferimento per le spese correnti</i>	22
1.3.3.2 <i>La definizione dei periodi di riferimento per le altre spese</i>	26
1.3.4 <i>I contenuti informativi</i>	29
1.3.4.1 <i>Le core variables europee</i>	29
1.3.4.2 <i>La rilevazione del reddito</i>	29
1.3.4.3 <i>La situazione lavorativa</i>	31
1.3.4.4 <i>La rilevazione delle quantità autoconsumate e delle quantità acquistate</i>	33
1.3.4.5 <i>I luoghi di acquisto</i>	36
1.3.4.6 <i>L'approfondimento sulle caratteristiche dei viaggi e delle escursioni</i>	37
1.3.5 <i>La componente longitudinale</i>	39
1.3.6 <i>L'uso di filtri</i>	40
1.3.6.1 <i>I percorsi filtrati</i>	40
1.3.6.2 <i>Il controllo degli zeri di spesa</i>	42
1.3.7 <i>La tecnica di rilevazione mista CAPI/CADI</i>	42
1.3.8 <i>Conclusioni</i>	43
2. Il disegno della nuova indagine	45
2.1 La transizione a tecniche di rilevazione CA nel contesto della modernizzazione delle indagini campionarie sulle famiglie	46
2.2 Il disegno di campionamento e la selezione del campione	48
2.2.1 <i>Il disegno di campionamento</i>	48



2.2.1.1 Schema generale del disegno di campionamento	48
2.2.1.2 Stratificazione e selezione del campione	49
2.2.2 L'utilizzo delle Liste Anagrafiche Comunali per la selezione del campione	52
2.3 Il disegno della rilevazione	54
2.3.1 Il ciclo di rilevazione	54
2.3.2 Gli strumenti	55
2.3.2.1 L'agenda e la scheda contatti	55
2.3.2.2 La struttura dei questionari	59
2.3.3 La rete di rilevazione	69
2.3.3.1 L'organizzazione della rilevazione e dei rapporti con la rete	69
2.3.3.2 La formazione dei rilevatori	70
2.3.3.3 Il monitoraggio della qualità dell'indagine	79
2.4 L'acquisizione e il trattamento dei dati	89
2.4.1 L'acquisizione controllata delle informazioni	90
2.4.1.1 Il piano delle compatibilità e i percorsi del questionario elettronico	90
2.4.1.2 L'acquisizione CADI del diario	92
2.4.2 Il trattamento dei dati	94
2.4.2.1 La fase di trattamento mensile	95
2.4.2.2 La fase di trattamento trimestrale: le variabili qualitative	96
2.4.2.3 La fase di trattamento trimestrale: le variabili quantitative	97
2.4.2.4 La fase di trattamento trimestrale: la valutazione monetaria delle quantità autoconsumate	100
2.4.2.5 La fase di trattamento annuale	101
2.5 La costruzione delle stime	101
2.5.1 La stima della spesa	101
2.5.2 Il sistema di ponderazione: scelta dei vincoli e trattamento degli outliers	103
2.5.2.1 La probabilità di inclusione e il peso diretto	104
2.5.2.2 La correzione per mancata risposta	105
2.5.2.3 La calibrazione a fonti esterne	106
2.5.3 Errori di campionamento	108
2.6 La riconciliazione tra indagine PAPI e indagine CAPI	111
2.6.1 Il raccordo tra le voci di spesa	111
2.6.2 Il raccordo dei periodi di riferimento	115
2.6.3 Un quadro di sintesi	116
3. La ricostruzione delle stime in serie storica	119
3.1 Il processo della ricostruzione	121
3.2 L'approccio micro al trattamento dei dati	122
3.2.1 Le voci elementari di spesa nelle due indagini	122
3.2.2 Il sistema di ponderazione e di calibrazione	126
3.3 L'impatto del raccordo micro	128
3.3.1 Un ulteriore passaggio micro: il riporto al mese	130
3.4 L'approccio macro di ricostruzione	133
3.5 La ricostruzione e il confronto con i modelli teorici	135
3.6 L'esito della ricostruzione: i risultati principali	139

Indice

3.6.1 <i>La ricostruzione della spesa</i>	140
3.6.2 <i>La ricostruzione delle incidenze di povertà</i>	146
3.7 Conclusioni	149
Glossario	153
Riferimenti bibliografici	157

5

140
146
149
153
157

PREMESSA¹

Il livello e la struttura della spesa sostenuta dalle famiglie per acquisire beni e servizi destinati al consumo costituiscono informazioni fondamentali per analizzare gli standard di vita e i comportamenti di consumo familiare nei differenti ambiti sociali e territoriali di un paese, costituendo, in Italia, gli elementi necessari per le stime ufficiali della povertà relativa e assoluta. Questi dati rappresentano, inoltre, la base informativa per la stima del sistema di ponderazione dell'indice dei prezzi al consumo e contribuiscono a definire la stima dei Consumi finali delle famiglie nell'ambito del sistema dei Conti Nazionali.

Indagini di questo tipo sono realizzate da decenni in tutti gli stati dell'Unione Europea (in Italia dal 1968) e, nonostante si differenzino ancora sotto diversi aspetti (frequenza, timing, contenuto, struttura, ecc.), a partire dal 1988, ogni cinque anni, Eurostat ne raccoglie e pubblica i risultati.

A partire da una serie di linee guida e raccomandazioni messe a punto dal Comitato del programma statistico europeo già negli anni '80, nell'ultimo decennio sono state individuate le dimensioni su cui concentrare gli sforzi di armonizzazione indirizzati a migliorare la comparabilità di questi dati. Per assecondare tale processo, il nostro paese ha realizzato una profonda ristrutturazione dell'indagine che ne ha ammodernato l'impianto metodologico tenendo conto, tra l'altro, anche delle trasformazioni intervenute negli stili di vita dei cittadini rispetto alle modalità di acquisto dei beni e servizi.

Il presente volume intende testimoniare questo complesso cammino di innovazione che si è avvalso delle competenze di un affiatato team di esperti delle diverse fasi di produzione e ricostruzione del dato. Il primo capitolo si sofferma, in particolare, sulla lunga serie di sperimentazioni che ha affiancato il processo di ristrutturazione dell'indagine e che ha consentito la revisione della metodologia riguardo, tra gli altri, alla classificazione dei beni e servizi oggetto di consumo, al periodo di riferimento delle spese, al processo di *data capture*, all'ampliamento dei contenuti informativi, inclusa l'introduzione delle variabili dell'indagine Viaggi e vacanze. Il secondo capitolo affronta il disegno della nuova indagine a partire dal campionamento fino alla costruzione delle stime, passando per la definizione dei nuovi strumenti di rilevazione, l'organizzazione del lavoro sul campo e il sistema di ponderazione. Infine, il terzo capitolo descrive il processo di ricostruzione in serie storica dei principali aggregati di spesa e della povertà (relativa e assoluta) per il periodo 1997-2013, direttamente confrontabili con i dati diffusi a partire dal 2014, che tengono conto anche dell'effetto delle nuove popolazione intercensuarie.

Il lavoro di ristrutturazione dell'Indagine sulle spese si è innestato nell'ambizioso programma di rinnovamento che l'Istat si è dato per aumentare l'efficienza del sistema statistico e la sua capacità di dare risposta alle esigenze degli utenti, contenendo, al contempo, l'onere sui rispondenti. Si tratta di un processo che prevede la reingegnerizzazione, la standardizzazione e l'industrializzazione dei processi, come

1 Il volume è stato curato da Donatella Grassi e Nicoletta Pannuzi. La premessa è a cura di Cristina Freguja.



pure l'innovazione dei prodotti, l'integrazione delle diverse fonti di dati (indagini campionarie e fonti amministrative) e la valorizzazione di quelli provenienti dagli archivi amministrativi.

Un simile approccio, non può che tradursi in uno sforzo continuo per trovare soluzioni metodologiche e tecnologiche volte a consolidare i risultati raggiunti e per spostare sempre più in là la frontiera dell'innovazione. Basti pensare al progetto di integrazione delle informazioni su consumo, reddito e ricchezza, sia attraverso metodologie di *linkage* e *matching* statistico, sia mediante l'arricchimento tematico dei questionari di indagine, al fine di costruire un dataset integrato di informazioni individuali in un'ottica di sostituibilità/integrazione della fonte amministrativa (trattata statisticamente) ai dati di natura campionaria.

Integrata nell'ampio sistema di produzione statistica e grazie anche alle innovazioni tecnologiche via via disponibili, l'Indagine sulle spese si gioverà sempre più della sinergia con le altre fonti del settore sia da indagine sia da archivio, rafforzando ulteriormente il proprio ruolo nell'ambito delle statistiche sulle condizioni socio-economiche delle famiglie.

1. IL PROCESSO DI RISTRUTTURAZIONE DELL'INDAGINE¹

1.1 Le linee guida europee per la rilevazione della spesa per consumi

L'indagine sui consumi (Household Budget Survey) viene condotta in tutti Paesi membri dell'Unione Europea, con lo scopo di rilevare la spesa sostenuta dalle famiglie per soddisfare i propri bisogni. In generale, tali indagini rappresentano la base informativa per la stima del sistema di ponderazione dell'indice dei prezzi al consumo e, in molti Paesi, forniscono anche informazioni rilevanti per la stima dei Consumi finali delle famiglie nell'ambito del sistema dei Conti Nazionali.

Nel decennio scorso, a livello europeo, è stato avviato un processo di *Gentlemen Agreement*, volto alla messa a punto di definizioni e metodologie condivise, che ha portato tra l'altro alla nuova classificazione Coicop dei beni e servizi oggetto di consumo (*Classification of Individual COntsumption by Purpose*). A partire dallo standard internazionale messo a punto dalla Divisione Statistica delle Nazioni Unite², nel quale le spese per consumo sono classificate in base allo scopo per il quale vengono effettuate, la classificazione adottata a livello europeo (Eurostat) è stata ulteriormente dettagliata. Nello specifico, ai tre livelli di disaggregazione (4 digit corrispondenti a 117 classi) contenuti nella classificazione UNSD ne è stato aggiunto un quarto (5 digit corrispondenti a 303 categorie).

I capitoli di spesa che definiscono il campo di osservazione dei consumi delle famiglie sono 14: i primi 12 sono relativi ai beni e servizi oggetto di consumo da parte delle famiglie, il 13esimo è relativo ai consumi delle istituzioni sociali private a servizio delle famiglie³ e l'ultimo a quelli delle amministrazioni pubbliche per le famiglie⁴. A titolo di esempio, la divisione 01 (Prodotti alimentari e bevande analcoliche) include due gruppi (01.1 – Cibo e 01.2 Bevande analcoliche); nell'ambito del primo gruppo sono previste nove classi (tra cui 01.1.1 - Pane e cereali, 01.1.2 – Carne, 01.1.2 – Pesce) e nel secondo due (01.2.1 – Caffè, tè, cioccolato, 01.2.2 – Acqua minerale, succhi di frutta e altre bibite analcoliche); la prima classe del primo gruppo è suddivisa a sua volta in otto categorie (tra cui 01.1.1.1 – Riso, 01.1.1.2 – Farina, 01.1.1.3 – Pane).

1 Il capitolo è stato redatto da D. Grassi ad eccezione del par. 1.3.4.3, redatto da P. Scalisi, e del par. 1.3.4.6, redatto da M. Di Torrice.

2 <http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=5>

3 13.1- Abitazione, 13.2- Salute, 13.3- Tempo libero e cultura, 13.4- Istruzione, 13.5- Protezione sociale, 13.6- Altri servizi.

4 14.1- Abitazione, 14.2- Salute, 14.3- Tempo libero e cultura, 14.4- Istruzione, 14.5- Protezione sociale.

Prospetto 1.1 – Raccomandazioni Eurostat per la conduzione delle HBS nei Paesi dell’Unione⁵

ASPETTO DELLA RILEVAZIONE	RACCOMANDAZIONE
Spesa per consumi	<p>Deve essere intesa come l’insieme delle spese sostenute per tutti i beni e servizi destinati al soddisfacimento dei bisogni familiari e individuali.</p> <p>Include:</p> <ol style="list-style-type: none"> 1. la stima del valore del canone mensile delle abitazioni di proprietà a disposizione della famiglia comprensive delle pertinenze (fitti figurativi) in quanto rappresentano il costo del servizio “abitare” che la famiglia consuma nel corso del tempo; 2. il valore monetario dei prodotti autoconsumati dalla famiglia, provenienti cioè dalla propria impresa (individuale o familiare), dalla coltivazione diretta (inclusa la raccolta di prodotti spontanei), dall’allevamento (inclusa la caccia e la pesca) o ricevuti come compenso per la propria attività lavorativa. <p>Non rappresentano invece spesa per consumi tutte quelle spese sostenute dalle famiglie per:</p> <ol style="list-style-type: none"> 1. l’attività professionale; 2. trasferimenti o donazioni; 3. accrescimento del patrimonio; 4. tasse e tributi.
Disegno di campionamento	Deve essere probabilistico e rappresentativo dei principali domini di stima
Unità di rilevazione	Deve essere la famiglia, definita rispetto a due criteri: coabitazione e condivisione delle spese
Capofamiglia	Deve essere il membro adulto della famiglia che maggiormente contribuisce al reddito familiare
Classificazione delle voci di spesa	Deve essere seguita la Coicop più recente (2013)
Core variables	<ol style="list-style-type: none"> 1. Sesso 2. Età in anni compiuti 3. Paese di nascita 4. Cittadinanza al momento dell’intervista 5. Stato civile 6. De facto marital status 7. Tipologia familiare 8. Paese di residenza (Nuts1) 9. Regione di residenza (Nuts2) 10. Grado di urbanizzazione 11. Condizione lavorativa autodichiarata 12. Posizione nella professione 13. Professione (ISCO88) 14. Settore di attività economica (NACE rev 2) 15. Titolo di studio più elevato conseguito (ISCED-97) 16. Reddito familiare netto <p>Per ciascuna di esse Eurostat identifica anche gli items minimi di risposta. La rilevazione delle <i>core variable</i>, considerate rilevanti nell’ambito delle HBS, deve essere effettuata attraverso gli strumenti di rilevazione che ciascun Paese ritiene più idonei.</p>
Valuta	La spesa per consumi familiari deve essere misurata in Euro e in PPS (<i>purchasing power standard</i>).
Cadenza del rilascio di dati	Quinquennale
Trasmissione delle informazioni	Ogni Paese partecipante deve trasmettere i dati dell’anno di riferimento e un Quality Report. La trasmissione dei dati deve essere effettuata secondo quanto descritto nei documenti redatti in occasione di ogni rilascio quinquennale (Description of the data transmission for HBS,)

Il fine dettaglio della classificazione costituisce un indubbio vantaggio in termini di comparabilità tra paesi e tra ambiti di applicazione; tuttavia, nel contesto delle indagini sulle famiglie⁶, presenta alcuni aspetti di criticità riconducibili essenzialmente alla richiesta di distinguere voci di spesa difficilmente scorporabili (ad esempio le spese per assicurazione sul viaggio nel caso di acquisto di un pacchetto vacanze “tutto compreso”), con dettaglio eccessivo (ad esempio la spesa per bevande alcoliche distinguendo rispetto al contenuto di alcool) o relative a beni e servizi considerati illegali

⁵ http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/en/hbs_esms.htm

⁶ <http://unstats.un.org/unsd/iiss/Classification-of-Individual-Consumption-According-to-Purpose-COICOP.ashx>

o socialmente riprovevoli (ad esempio droghe o prostituzione). Per tali motivi, i singoli Paesi possono derogare da quanto previsto nella classificazione (non esiste infatti un regolamento specifico), ma sono invitati ad attenersi alle linee guida e alle raccomandazioni messe a punto dal Comitato del programma statistico europeo nel 1989.

Non vige dunque per i Paesi membri l'obbligo di effettuare la rilevazione, ma quello di attenersi ad alcuni principi di base qualora intendano effettuarla e trasmettere le informazioni ad Eurostat (Prospetto 1.1). Tali principi riguardano la definizione di spesa per consumi familiari, la classificazione delle voci di spesa, alcuni aspetti del disegno di campionamento, l'unità di rilevazione e la determinazione della persona di riferimento; viene inoltre raccomandata la rilevazione di alcune specifiche variabili socio-economiche familiari e individuali (*core variables*) necessarie a descrivere il contesto nell'ambito del quale il comportamento di spesa si esplica, permettendo una migliore identificazione delle popolazioni oggetto di studio e dei sottogruppi che la costituiscono. Inoltre, se sistematicamente introdotte in tutte le indagini sociali, esse consentono analisi integrate di dati provenienti da fonti diverse.

Fatti salvi tali principi di base, ogni Paese può disegnare la rilevazione nel modo che ritiene più opportuno, stabilendo in autonomia la tecnica di campionamento e la numerosità campionaria, la tecnica di rilevazione, sia per quanto concerne gli strumenti sia per quanto concerne la rete di rilevazione, i metodi di trattamento delle informazioni rilevate, il sistema di ponderazione; ciò ovviamente limita la comparabilità, nello spazio e nel tempo.

1.2 La rilevazione della spesa per consumi in Italia

In Italia, l'indagine sui consumi delle famiglie viene svolta con regolarità sin dalla fine degli anni '60; nel corso del tempo la metodologia di indagine ha subito diversi mutamenti, anche rilevanti, che in alcuni casi hanno inficiato la confrontabilità dei dati nel tempo. L'ultima ristrutturazione è quella del 1997 e ha riguardato tutte le fasi del processo di produzione del dato, al fine di garantire una più elevata qualità delle informazioni rilasciate, ottenere stime rappresentative anche a livello regionale e armonizzare definizioni e metodologie alle allora vigenti direttive Eurostat (Eurostat, 1993; Istat, 2000).

Tutto ciò ha determinato l'interruzione della serie storica dei dati sulle spese per consumi; attualmente il confronto diretto tra le stime si limita agli anni 1997-2013 in quanto, da allora, il disegno della rilevazione è rimasto pressoché inalterato (Prospetto 1.2).

Prospetto 1.2 – La rilevazione della spesa per consumi in Italia – Anni 1997-2013

DESCRIZIONE	
Schema di campionamento	Casuale a due stadi (comuni-famiglie) con stratificazione delle unità di primo stadio. Il territorio italiano viene suddiviso in 228 strati in base alla dimensione demografica del comune e alla regione di appartenenza: 107 strati sono costituiti da un unico comune e 121 da più comuni tra i quali ne vengono selezionati casualmente tre (uno per ciascun mese del trimestre)
Domini di studio territoriali	Ripartizione, Regione
Domini di studio temporali	Trimestre, Anno
Numerosità campionaria teorica	Circa 28 mila famiglie residenti in 470 comuni italiani, con sostituzione delle unità di secondo stadio nell'ambito di un bacino di famiglie sostitutive pari al numero totale di famiglie base da intervistare
Periodicità della rilevazione	Continua (su base mensile)
Unità di rilevazione	Famiglia di fatto definita come l'insieme di persone coabitanti e legate da vincoli di matrimonio, parentela, affinità, adozione, tutela o affettivi.
Persona di riferimento	L'intestatario del foglio di famiglia anagrafico
Rispondente	Tutti i componenti della famiglia; se uno dei componenti non può partecipare alla rilevazione, sono ammesse risposte <i>proxy</i>
Tecnica di somministrazione	Mista in modalità PAPI (diario autocompilato e intervista diretta)
Modalità di svolgimento della rilevazione	Per ciascuna famiglia selezionata sono previste due visite a domicilio da parte del rilevatore: <i>prima visita</i> : nei giorni precedenti l'inizio del periodo di compilazione, il rilevatore consegna il diario alla famiglia; <i>seconda visita</i> : nei primi giorni del mese successivo a quello di compilazione del diario, il rilevatore effettua la visita per ritirare il diario e svolgere l'intervista finale
Rete di rilevazione	Costituita dai rilevatori, selezionati dagli uffici di statistica dei Comuni, formati e coordinati dall'Istat, tramite i propri Uffici Regionali con la supervisione del Servizio Condizioni economiche delle Famiglie
Principale variabile oggetto di studio	Spesa per consumi delle famiglie intesa come l'insieme di tutte le spese sostenute dalle famiglie per acquistare beni e servizi destinati al diretto soddisfacimento dei bisogni dei propri componenti (inclusi gli autoconsumi, il fitto figurativo e i regali).
Classificazione delle voci di spesa	Procome - Coicop ⁶
Periodo di riferimento delle variabili di spesa	7 giorni Alimentari, autoconsumi, servizi di ristorazione e mense, articoli e servizi non alimentari correnti
	Ultimo mese Abbigliamento, arredamenti, articoli per la casa, servizi domestici, sanità, vacanze, articoli sportivi e sport, cultura, istruzione, trasporti
	Ultimi 3 mesi Combustibili e utenze, manutenzioni abitazioni, beni durevoli, mezzi di trasporto e di comunicazione
	Ultimi 12 mesi Spese eccezionali e assicurazione veicoli
Core variables	<ol style="list-style-type: none"> 1. Sesso 2. Età in anni compiuti 3. Paese di nascita 4. Stato civile 5. Stato civile (di fatto) 6. Tipologia familiare 7. Paese di residenza (Nuts1) 8. Regione di residenza (Nuts2) 9. Grado di urbanizzazione 10. Condizione lavorativa autodichiarata 11. Posizione nella professione (distintamente per dipendenti e indipendenti) 12. Settore di attività economica (ATECO 2002) 13. Titolo di studio più elevato conseguito (ISCED-97)

Tra i principali obiettivi dell'indagine vi sono la fornitura trimestrale di dati alla Contabilità nazionale (CN), per la stima dei Consumi finali delle famiglie (Istat, 2004a), e la fornitura annuale di dati al Servizio Prezzi, per la ponderazione del paniere.

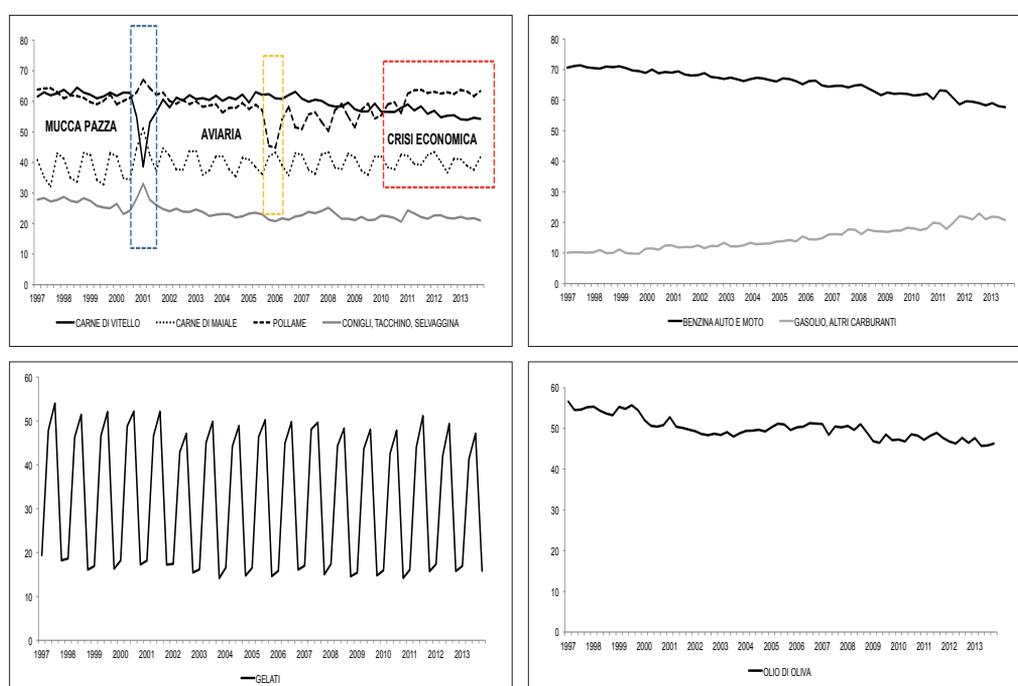
Nel tempo tuttavia l'accresciuta richiesta di informazioni nell'ambito della ricerca sociale ed economica ha decisamente ampliato gli obiettivi dell'indagine: i dati di spesa vengono utilizzati per la stima della povertà relativa e assoluta (Istat, 2009), per la stima dell'inflazione per classi di spesa delle famiglie (Istat 2007; Istat 2014a) e rappresenteranno anche parte del contenuto di un database integrato, attraverso

1. Il processo di ristrutturazione dell'indagine

tecniche di *linkage* e di *matching*, su reddito, consumi, ricchezza e benessere (Donatiello et al., 2014, Garofalo, 2014). Contestualmente, i comportamenti di spesa e di acquisto da parte delle famiglie sono cambiati, sia in termini di abitudini, sia rispetto alla scelta dei beni, servizi e dei luoghi di acquisto.

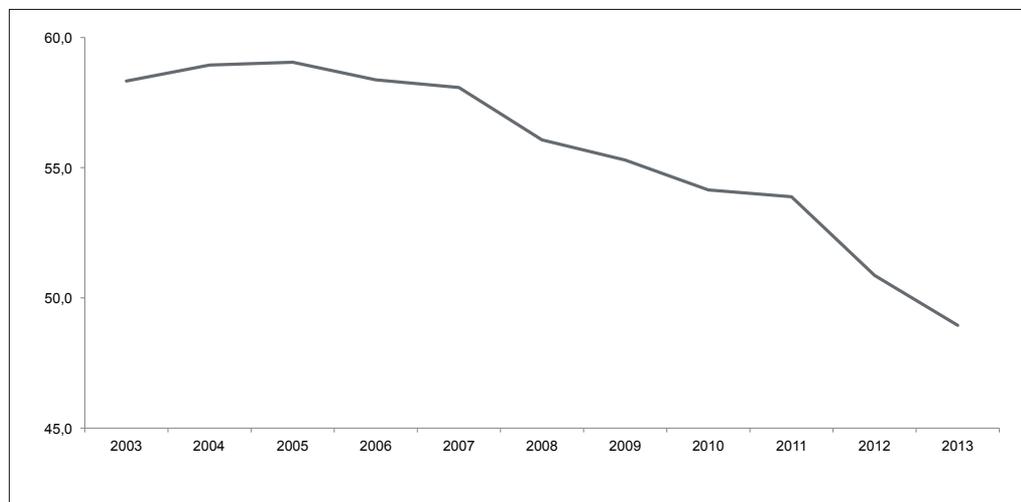
Pur continuando a cogliere in maniera precisa gli eventi congiunturali, le specifiche stagionalità delle spese e i cambiamenti nelle scelte di allocazione della spesa (Grafico 1.1), l'indagine, con il passare degli anni, ha mostrato la progressiva inadeguatezza di alcuni strumenti, come ad esempio, l'uso di una finestra temporale di 7 giorni per rilevare le spese alimentari o correnti.

Grafico 1.1 – Frequenze di spesa settimanali per l'acquisto di alcuni beni rilevati tramite diario autocompilato – Dati per trimestre – Anni 1997-2013 (valori percentuali)



Sono infatti sempre di meno le famiglie che effettuano acquisti tutti i giorni o tutte le settimane: il numero medio di acquisti in una settimana, in aumento fino al 2005 (quando arriva a 59), inizia a diminuire già a partire dal 2006, prima cioè dell'inizio della crisi, e raggiunge 48,9 nel 2013, con una perdita media di circa 10 acquisti a settimana (Grafico 1.2).

Grafico 1.2 – Numero medio di acquisti nei 7 giorni di compilazione del diario tra le famiglie che hanno effettuato almeno un acquisto – Anni 2003-2013



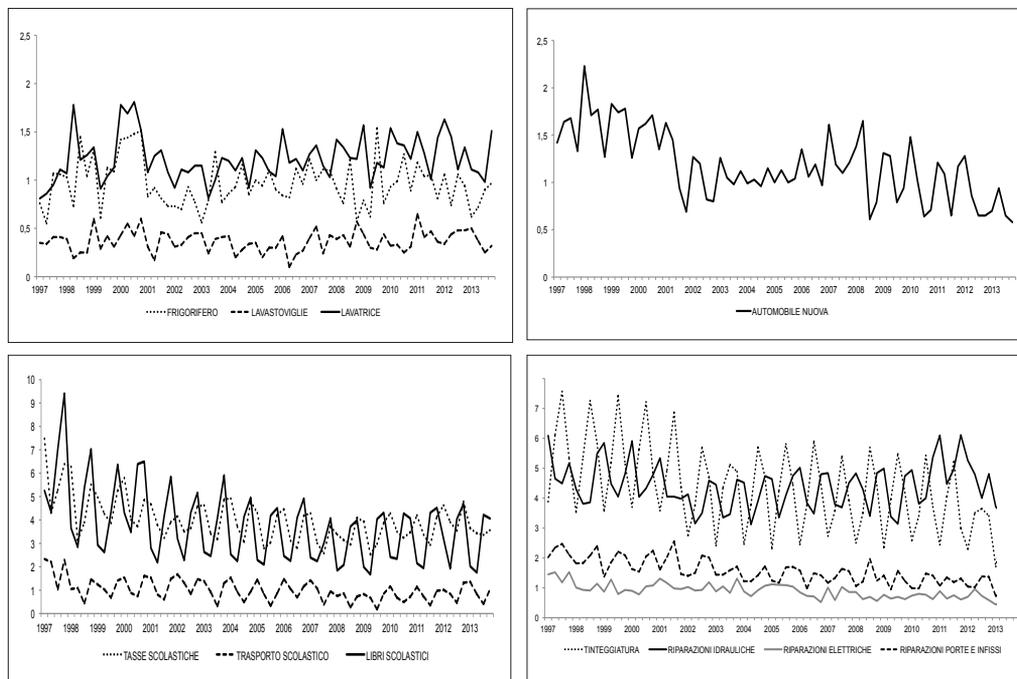
I cambiamenti nelle abitudini di acquisto si sono riflessi in una diminuzione delle frequenze di spesa che, in alcuni casi, ha comportato un aumento della variabilità delle stime (Grafico 1.3); si è quindi palesata la necessità di aggiustamenti – nelle definizioni o nei periodi di riferimento – anche al fine di irrobustire le stime.

Nel tempo, inoltre, la gamma di prodotti disponibili sul mercato e i canali distributivi sono cambiati, determinando modifiche importanti nei comportamenti di acquisto delle famiglie, spesso in maniera diversificata per sottogruppi di popolazione. Si pensi ad esempio al fatto che nel giro di 15 anni la frequenza di acquisto dei telefoni cellulari è raddoppiata, nonostante la crisi, mentre quella dei videoregistratori si è pressoché azzerata; si pensi anche ai nuovi prodotti lanciati sul mercato, come *e-reader* o *e-book*, alle antenne satellitari o ai pacchetti per le comunicazioni “tutto compreso”.

A tutto ciò si aggiungono gli importanti mutamenti sociali, economici e culturali che hanno caratterizzato l'Italia negli ultimi anni, tra i più importanti l'aumento della presenza straniera e le trasformazioni del mercato del lavoro.

Infine, il processo di armonizzazione condotto a livello internazionale ha messo sempre più in evidenza la necessità di aggiornare gli strumenti di rilevazione, anche in un'ottica di recepimento delle raccomandazioni dell'Istituto di statistica europeo.

Grafico 1.3 – Frequenze di spesa trimestrali (mensili perle tasse, il trasporto e i libri scolastici) per alcuni beni e servizi rilevati tramite quesiti retrospettivi – Anni 1997-2013 (valori percentuali)



Tutto quanto detto ha quindi reso necessario modificare l'impianto e gli strumenti di indagine, al fine di adeguarli alla nuova realtà e ai rinnovati obiettivi. Gli strumenti per la rilevazione delle informazioni statistiche necessitano, infatti, di continua manutenzione e periodiche revisioni in modo da mantenere, nel tempo, la loro capacità di cogliere fenomeni in mutamento. D'altra parte, sussiste la necessità di effettuare misure ripetute dello stesso fenomeno che siano confrontabili nel tempo, al fine di dar conto della dinamica e dei cambiamenti intervenuti. La possibilità di confrontare le informazioni nel tempo rappresenta, di per sé, un valore aggiunto all'analisi del fenomeno e la rottura di serie deve essere necessariamente valutata e spiegata.

La ricerca di un giusto compromesso tra le due esigenze ha comportato, a sedici anni dalla precedente ristrutturazione, una revisione completa del processo e una ricostruzione in serie storica dei principali aggregati, cercando così di mantenere quanto più possibile la continuità con il passato.

1.3 Il sistema delle sperimentazioni per una revisione della metodologia di indagine

La revisione del processo ha avuto tra i principali obiettivi quello di calare la rilevazione della spesa per consumi nel più generale contesto, nazionale ed internazionale, della rilevazione delle condizioni socio-economiche delle famiglie; di aggiornare cioè il dizionario comune tra i diversi ambiti di analisi, di creare le condizioni affinché l'indagine diventasse parte di un più ampio sistema di informazione e di garantire, assieme ad elevati standard di qualità e robustezza delle stime, la continuità informativa con i dati rilevati negli anni precedenti.

Prospetto 1.3 – Il sistema delle sperimentazioni

Denominazione indagine	Dimensione campionaria	Rete di rilevazione	Classificazione	Tecnica di rilevazione
Indagine sulle frequenze di spesa - (Maggio-novembre 2004)	Circa 2.000 famiglie in 38 comuni	Comunale	La stessa utilizzata nel diario della PAPI 2004	Diario giornaliero autocompilato per un mese dalle famiglie per la registrazione degli eventi di spesa
Indagine sperimentale per la rilevazione delle spese correnti - (Ottobre-novembre 2006)	Circa 2.000 famiglie in 128 comuni	Comunale	Coicop 2003	Diario giornaliero autocompilato per 14 giorni per rilevare le spese correnti e degli autoconsumi con 5 unità di misura
Indagine sperimentale per la definizione dei periodi di riferimento delle spese familiari - (Ottobre 2006)	Circa 1.500 famiglie in 64 comuni	Comunale	Coicop 2003	Intervista diretta per la rilevazione delle caratteristiche familiari e delle spese e del mese di effettuazione con ampliamento dei periodi di riferimento e inserimento di un filtro iniziale per ogni voce di spesa
Indagine sperimentale per la rilevazione delle quantità alimentari acquistate - (Aprile-giugno 2007)	Circa 1.000 famiglie in 38 comuni	Comunale	Coicop 2003 più dettagliata per rilevare le quantità acquistate tenendo conto dei nutrienti	Diario giornaliero autocompilato per 14 giorni per la rilevazione delle spese alimentari, delle quantità acquistate, del luogo di acquisto, degli autoconsumi <ul style="list-style-type: none"> - Diario giornaliero autocompilato per 14 giorni per la rilevazione delle spese correnti degli autoconsumi, dei luoghi di acquisto - Intervista iniziale diretta per la rilevazione delle caratteristiche familiari
Indagine pilota sui consumi delle famiglie (Ottobre-dicembre 2008)	Circa 3.500 famiglie in 470 comuni	Comunale	Coicop 2003	<ul style="list-style-type: none"> - Intervista finale diretta per la rilevazione delle spese familiari non rilevate tramite diario, il mese di effettuazione e il luogo di acquisto - Diario giornaliero autocompilato per 14 giorni per rilevare le spese correnti degli autoconsumi, dei luoghi di acquisto - Intervista iniziale diretta per la rilevazione delle caratteristiche familiari
Indagine pilota sui consumi delle famiglie (Ottobre-dicembre 2009)	Stesse famiglie intervistate nel 2008	Comunale	Coicop 2003	<ul style="list-style-type: none"> - Intervista finale diretta per la rilevazione delle spese familiari non rilevate tramite diario, il mese di effettuazione e il luogo di acquisto - Diario settimanale autocompilato per 14 giorni per rilevare le spese correnti degli autoconsumi, dei luoghi di acquisto - Intervista iniziale diretta per la rilevazione delle caratteristiche familiari
Indagine pilota sui consumi delle famiglie (maggio-giugno 2010)	Circa 3.500 famiglie in 470 comuni	Comunale	Coicop 2003	<ul style="list-style-type: none"> - Intervista finale diretta per la rilevazione delle spese familiari non rilevate tramite diario, il mese di effettuazione e il luogo di acquisto - Diario settimanale autocompilato per 14 giorni per rilevare le spese correnti degli autoconsumi, dei luoghi di acquisto - Intervista iniziale diretta CAPI per la rilevazione delle caratteristiche familiari (tutte le core variables) - Intervista finale diretta CAPI per la rilevazione delle spese familiari non rilevate tramite diario, il mese di effettuazione e il luogo di acquisto. Viene introdotto un focus sulle caratteristiche e sulle spese di viaggio
Indagine pilota sui consumi delle famiglie (maggio 2011)	Circa 1.000 famiglie in 32 comuni	Professionale	Coicop 2003	<ul style="list-style-type: none"> - Diario settimanale autocompilato per 14 giorni per rilevare le spese correnti degli autoconsumi, dei luoghi di acquisto - Intervista iniziale diretta CAPI per la rilevazione delle caratteristiche familiari (tutte le core variables) - Intervista finale diretta CAPI per la rilevazione delle spese familiari non rilevate tramite diario, il mese di effettuazione e il luogo di acquisto. Focus sulle caratteristiche e sulle spese di viaggio
Indagine sulle spese delle famiglie (maggio-dicembre 2012)	Circa 14.000 famiglie in 500 comuni	Professionale	Coicop 2003	

L'obiettivo è stato dettagliato nei 10 sotto-obiettivi di seguito riportati, al fine di approcciare in modo approfondito tutte le criticità della rilevazione:

1. definizione della strategia di campionamento e di costruzione delle stime, al fine di rappresentare meglio specifici sottogruppi di popolazione;
2. armonizzazione della classificazione delle voci di spesa alla più recente classificazione Coicop;

1. Il processo di ristrutturazione dell'indagine

3. definizione dei periodi di riferimento delle voci rilevate, sia tramite diario sia in modo retrospettivo, al fine di cogliere meglio i comportamenti di spesa e rendere più robuste le stime;
4. innovazione delle tecniche di cattura, di controllo e di correzione del dato, in vista del notevole ampliamento informativo introdotto, al fine di garantire una maggiore qualità e tempestività dell'informazione;
5. introduzione una componente panel, al fine di migliorare la precisione delle variazioni degli aggregati stimati;
6. introduzione della rilevazione delle quantità alimentari acquistate, allo scopo di ottenere informazioni utili sui comportamenti alimentari e nutrizionali delle famiglie;
7. introduzione della rilevazione del luogo di acquisto, informazione rilevante soprattutto per la rilevazione dei prezzi al consumo;
8. rilevazione delle *core variables* richieste da Eurostat e delle variabili necessarie ad eventuali *matching* statistici con altre indagini dell'Istituto;
9. valutazione degli effetti di *burden* su rispondenti e rilevatori;
10. continuità dei confronti temporali prima/dopo la ristrutturazione.

Prospetto 1.4 – Le sperimentazioni per la valutazione degli obiettivi

	Fre- quenze di spesa	Rilevazio- ne delle spese correnti	Definizione dei periodi di riferi- mento delle spe- se familiari	Rilevazio- ne delle quantità alimentari acquistate	Pilota sui consumi delle famiglie 2008	Pilota panel sui consumi delle fami- glie 2009	Pilota sui consumi delle fami- glie 2010	Pilota sui consumi delle fami- glie 2011	Spese delle famiglie
Definizione della strategia di campionamento e di costruzione delle stime					X	X			
Armonizzazione della classificazione alla più recente Coicop		X	X	X	X	X	X	X	
Definizione dei periodi di riferimento		X	X	X					
Innovazione delle tecniche di cattura, di controllo e di correzione del dato									X
Introduzione una componente panel						X			
Introduzione della rilevazione delle quantità alimentari acquistate				X					
Introduzione della rilevazione del luogo di acquisto				X					
Rilevazione delle core variables									X
Valutazione effetti di burden su rispondenti e rilevatori	X	X	X	X	X	X	X	X	X
Continuità dei confronti temporali									X

Poiché l'impatto dei cambiamenti connessi a ciascun obiettivo doveva essere valutato sia singolarmente sia nel complesso, è stato pianificato un sistema di indagini sperimentali "a cascata" (Prospetti 1.3 e 1.4), che, nell'arco di 10 anni, ha consentito di valutare, di volta in volta, gli effetti dei cambiamenti, arrivando per gradi alla definizione del nuovo disegno di indagine. Per ogni indagine sperimentale sono stati valutati, oltre all'effetto sui livelli e sulla struttura della spesa, anche la portata dell'effetto ricordo, dell'effetto *telescoping* e dell'effetto stanchezza. Nei paragrafi successivi, verranno illustrate le ipotesi che hanno delineato la cornice del sistema e il disegno delle sperimentazioni, insieme ai principali risultati ottenuti con riferimento agli specifici obiettivi.

1.3.1 Alcune considerazioni preliminari

Nell'affrontare la ristrutturazione dell'indagine, si è partiti dal presupposto che in Italia l'Indagine sulle spese rappresenta una fonte primaria per la stima dei consumi finali delle famiglie nel sistema dei Conti Nazionali; tale esigenza determina la necessità di stimare la spesa sostenuta dalle famiglie per tutte le voci inserite nella classificazione Coicop a 5 digit, con cadenza trimestrale.

Tempestività e dettaglio sono pertanto i requisiti fondamentali affinché le stime prodotte possano essere utilizzate nell'ambito del sistema dei Conti Nazionali.

L'indagine deve quindi necessariamente i) essere condotta tutti gli anni, ii) prevedere una strategia di campionamento trimestrale, iii) consentire di cogliere la stagionalità del fenomeno e iv) fornire stime rappresentative a un livello molto disaggregato.

È noto in letteratura (Crossley, Winter, 2013 con molti richiami alla letteratura sulla materia) che una maggiore disaggregazione delle voci induce stime più elevate della spesa totale, presumibilmente perché aiuta le famiglie a ricordare anche spese che altrimenti verrebbero dimenticate; un'eccessiva disaggregazione, d'altro canto, aumenta notevolmente il carico di lavoro delle famiglie, determinando un più elevato effetto "stanchezza". È dunque essenziale individuare il dettaglio ottimale, per mezzo del quale ottenere dati di qualità, che rispondano anche alle esigenze dei principali utilizzatori del dato.

Nel caso italiano, la tempestività rappresenta un requisito chiave che può essere raggiunto limitando al massimo la codifica e il trattamento a posteriori dell'informazione rilevata; l'utilizzo di questionari strutturati presenta il duplice vantaggio di standardizzare le modalità di raccolta delle informazioni e di sollecitare il ricordo, attraverso una dettagliata ed esaustiva elencazione delle voci.

Per ridurre il peso sul rispondente si può, inoltre, suddividere la rilevazione in più fasi e utilizzare tecniche miste di *data capturing*, scegliendo il metodo più idoneo a rilevare le categorie di spesa che compongono l'aggregato totale (Silberstein, Scott; 1991). A guidare tale scelta sono generalmente l'entità e la frequenza delle singole spese: le più rare e consistenti possono essere rilevate attraverso quesiti retrospettivi (sottoposti tramite intervista diretta); quelle più frequenti, generalmente anche di entità più contenuta, possono essere efficacemente rilevate attraverso un diario giornaliero, strumento più idoneo anche per la rilevazione degli autoconsumi (consumo di beni prodotti dalla famiglia stessa).

La tecnica mista con questionari strutturati è stata dunque la tecnica scelta per soddisfare i requisiti di tempestività e dettaglio richiesti.

Il reddito, la spesa per consumi, la ricchezza sono variabili che presentano componenti sia familiari sia individuali; tuttavia, se per il reddito e la ricchezza le poste individuali sono decisamente maggioritarie, la spesa per consumi si caratterizza per una preponderante connotazione familiare: solo alcune spese possono essere imputabili ai singoli individui e la componente più rilevante (le spese alimentari e per l'abitazione) non può che essere riferita alla famiglia nel suo complesso, nell'ambito della quale si mettono in atto economie di scala, con effetti diretti sia sui livelli sia sulle scelte di allocazione.

In alcuni paesi europei la rilevazione avviene attraverso l'uso di questionari individuali, in altri ai questionari individuali si associano quelli familiari. La scelta tra questionari individuali o familiari è stata ampiamente dibattuta in letteratura senza pervenire a conclusioni univoche (Crossley, Winter, 2013; Goldenberg and Ryan, 2009; Silberstein, Scott, 1991; Filippucci, Ferrante, 1997).

L'utilizzo di questionari multipli induce livelli di spesa più elevati, suggerendo, da una parte, una possibile mancata registrazione degli eventi di spesa tramite questionari familiari, ma, dall'altra, un possibile effetto "duplicazione della spesa", soprattutto per i beni e i servizi consumati da tutta la famiglia.

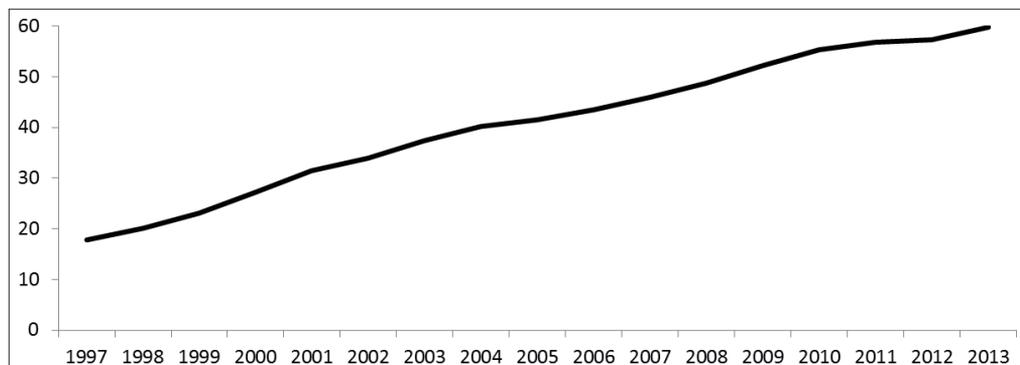
I questionari individuali mostrano inoltre una più elevata incidenza di mancata collaborazione, confermata dai risultati ottenuti con uno specifico quesito inserito nella sperimentazione del 2008: il 70% circa delle famiglie ha dichiarato di essere poco o per niente d'accordo con l'ipotesi di ricorrere a questionari individuali.

Una rilevazione con questionari a livello familiare è stata dunque la scelta effettuata, per garantire soddisfacenti livelli di partecipazione ed eliminare il possibile "effetto duplicazione".

L'uso di strumenti e questionari cartacei determina una maggior probabilità di errore, sia nella fase di compilazione, sia nella fase di registrazione, rispetto all'utilizzo di lettori ottici o di tecniche di rilevazione *Computer Assisted (CA)*. Oltre a ridurre i possibili errori di compilazione/digitazione, l'uso di tali tecniche riduce anche i tempi di rilascio del dato, essendo la fase di rilevazione contestuale a quella di acquisizione. Per tale motivo, il passaggio a una tecnica CAPI è sembrata la naturale modernizzazione della rilevazione tramite intervista diretta.

Il "diario elettronico", che possa cioè essere compilato tramite personal computer, sembra invece uno strumento ancora difficilmente gestibile dalle famiglie. Nel 2013 (Grafico 1.4), quasi il 60% della famiglie possiede un personal computer (quota triplicata rispetto al 1997) e, tra queste, quasi tutte dispongono di accesso ad internet; le percentuali tuttavia diminuiscono drasticamente per le famiglie di anziani (meno di un quinto possiede un personal computer), quando la persona di riferimento ha un basso titolo di studio (circa un terzo) e nel Sud (poco più della metà). Inoltre, la richiesta alla famiglia di usare nuove tecnologie determina, ancora una volta, alti tassi di mancata partecipazione che, in questo caso, sarebbero affetti anche da un'elevata distorsione. Allo stato attuale, si è preferito procedere con un diario cartaceo, rimandando a un prossimo futuro le valutazioni circa il ricorso a tecniche miste.

Grafico 1.4 – Famiglie che possiedono PC – Anni 1997-2013 (valori percentuali)



1.3.2 Il dettaglio delle voci di spesa

La classificazione delle voci di spesa per consumi ha subito nel tempo diverse revisioni. L'iniziale classificazione, denominata Procome, basata sull'European Combined Nomenclature e rivista da Eurostat nel 1993, che poneva l'attenzione soprattutto sui beni alimentari e sulla loro comparabilità a livello internazionale, nel 1997 viene sostituita con l'adattamento della classificazione Coicop (utilizzata per i conti nazionali e per gli indici dei prezzi) alla rilevazione sulle famiglie, la cosiddetta Coicop-HBS. Tale classificazione, come già accennato, prevede quattro livelli: i) Divisioni (primi due digit), ii) Gruppi (terzo digit), iii) Classi (quarto digit) e iv) Categorie (quinto digit).

Da allora le revisioni sono state periodiche: nel 1999, nel 2003, nel 2011 e nel 2013 sono stati apportati aggiornamenti con un ampliamento progressivo del numero di voci, passate dalle 192 del 1997 alle 303 del 2013.

La classificazione rappresenta lo standard di riferimento per la trasmissione delle informazioni ad Eurostat; i singoli Paesi possono, tuttavia, prevedere ulteriori disaggregazioni (al fine di cogliere meglio i comportamenti di spesa tenendo anche conto del disegno complessivo di indagine) dalle quali si deve comunque poter risalire alla classificazione richiesta. Nonostante la Coicop sia decisamente dettagliata, ulteriori disaggregazioni possono risultare utili per:

- chiarire la richiesta e facilitare il ricordo delle famiglie;
- soddisfare richieste specifiche di interlocutori privilegiati;
- utilizzare modalità e strumenti diversi, distinguendo voci che, pur appartenendo alla stessa categoria Coicop, presentano differenze importanti rispetto all'entità e alla frequenza di spesa;
- adattare la classificazione internazionale alle esigenze nazionali; la classificazione internazionale è infatti il frutto di un accordo tra paesi che hanno specifiche realtà in riferimento ai diversi domini di studio. In particolare, motivi culturali o economici possono rendere impossibile o sgradita per la famiglia la rilevazione di specifiche spese: oltre alle spese per narcotici o prostituzione, si pensi a quelle per "Test di gravidanza e contraccettivi meccanici" o, più semplicemente, a quelle per bevande alcoliche. Il dettaglio di alcune voci può infine risultare utile per conservare alcune sovrapposizioni con il passato, in un'ottica di raccordo delle serie di dati.

Per valutare l'impatto dell'introduzione della nuova classificazione, tenendo sotto controllo tutti gli aspetti di cui sopra, nel 2006 sono state condotte due sperimentazioni: l'Indagine sperimentale per la rilevazione delle spese correnti (COR), e l'Indagine sperimentale per la definizione dei periodi di riferimento delle spese familiari (DIR). Entrambe le indagini, pur accogliendo le modifiche imposte dalla classificazione Coicop del 2003, mantengono quanto più possibile inalterate le definizioni delle voci di spesa rispetto all'indagine sui consumi condotta in quell'anno al fine di fornire una misura di impatto depurata da effetti indesiderati. Al fine di alleggerire il *burden* sulle famiglie, in fase sperimentale i due campioni sottoposti a test sono stati selezionati in modo che le famiglie che compilavano il diario fossero diverse da quelle che rispondevano all'intervista diretta.

L'obiettivo primario era quello di individuare il dettaglio e la sequenza di voci ottimali per rendere chiare le definizioni, per scorporare spese che, seppur afferenti alla stessa voce Coicop, risultavano molto diverse per entità e frequenza (ad esempio, tra le spese per trasporti su strada – codice Coicop-2003 07.3.2.1 – si sono distinte le spese per biglietti singoli, rilevate tramite diario giornaliero, da quelle per abbonamenti, rilevate tramite quesito retrospettivo) o per tenere conto della sequenza concettuale di alcune voci (ad esempio, la voce relativa agli alimenti per animali è stata inserita immediatamente dopo le spese alimentari poiché le famiglie concettualmente la collocano proprio tra le spese alimentari benché la Coicop le collochi tra i beni e servizi per Ricreazione, spettacoli e cultura). Le due indagini hanno fornito i primi fondamentali elementi di valutazione per la messa a punto della successiva indagine sperimentale, condotta in modalità PAPI nel quarto trimestre del 2008, con la quale sono state rilevate, su un campione teorico di circa 3.500 famiglie, 415 voci di spesa (contro le 198 della classificazione Coicop 2003).

In tale sede, oltre ad introdurre alcune modifiche definitorie si è sottoposta a test la possibilità di rilevare le spese sanitarie attraverso quesiti retrospettivi, valutando anche il grado di compartecipazione dei cittadini alla spesa stessa; si è chiesto quindi alle famiglie di distinguere le spese sostenute per il pagamento del solo ticket, relativo a prestazioni sanitarie offerte dal Sistema sanitario nazionale (SSN), da quelle sostenute per prestazioni sanitarie fornite dal settore privato (inclusa intramoenia). L'analisi dei dati rilevati ha messo in evidenza alcune criticità; per molte delle spese sanitarie la percentuale di mancata risposta parziale è stata decisamente elevata: per il pagamento di ticket per prestazioni dentistiche la mancata risposta sfiora il 20% e supera il 50% nel caso di pagamenti di ticket per visite presso medici generici o per ricoveri in strutture pubbliche. Le famiglie dunque non comprendono bene la distinzione o non sono in grado di fornire informazioni circa il pagamento di ticket. Ne è derivata una revisione totale del quesito con l'esclusione delle distinzioni tra spese a pagamento intero o in compartecipazione con il SSN.

Per il resto la classificazione è apparsa funzionale alla rilevazione della spesa ed è stata ben gestita dalle famiglie e può pertanto essere adottata.

1.3.3 I riferimenti temporali

La corretta individuazione del periodo di riferimento temporale per ciascuna voce di spesa rilevata rende robuste le stime in termini sia di frequenza sia di ammontare.

Se ad esempio la rilevazione tramite diario (contestuale dunque al momento della spesa) è la tecnica più funzionale a misurare voci di spesa frequenti, di entità contenuta, la



durata della sua compilazione rappresenta un elemento fondamentale per la riuscita della rilevazione; se troppo lunga determina, infatti, un eccessivo carico di lavoro per il rispondente e può tradursi in pesanti distorsioni connesse con l'effetto stanchezza (Crossley, Winter, 2013); se troppo breve può produrre stime con minori distorsioni in termini di media, ma, a causa delle scarse frequenze osservate, affette da una maggiore variabilità.

Anche su tale argomento il dibattito in letteratura è ampio e dimostra come non sia possibile individuare un unico periodo ottimale per tutte le voci di spesa (Bradburn, 2010; Silberstein, Jacob, 1989).

La scelta scaturisce da una sorta di bilanciamento tra due aspetti fondamentali che caratterizzano la rilevazione della spesa: la frequenza con cui viene effettuata e la capacità da parte della famiglia di ricordarla. Se infatti il primo aspetto suggerirebbe l'opportunità di allargare la finestra temporale al diminuire della frequenza con cui la spesa viene effettuata (al fine di avere un più elevato numero di eventi osservati), tale allargamento può risultare eccessivo nel caso in cui la possibilità di ricordo sia limitata nel tempo (perché l'entità della spesa ad esempio è contenuta e tende quindi ad essere dimenticata facilmente).

È tuttavia possibile individuare alcuni criteri generali per la definizione del periodo di riferimento ottimale in base alla combinazione di alcuni elementi:

- Distribuzione delle spese nell'arco dell'anno (concentrata, uniforme, casuale)
- Percentuale di famiglie che effettuano la spesa (alta, bassa)
- Ammontare di spesa (alto, basso)
- Variabilità delle spese (alta, bassa)

In generale, un ammontare di spesa alto viene dimenticato più difficilmente e consente quindi di ampliare il periodo di riferimento rispetto a uno basso.

Le spese più facili da catturare sono quelle distribuite in modo uniforme nell'arco dell'anno, effettuate da una percentuale alta di famiglie e con una bassa variabilità. In questo caso anche una finestra temporale ridotta può assicurare stime robuste (in generale questa è l'ipotesi che viene fatta per i generi alimentari acquistati frequentemente e dalla maggior parte delle famiglie).

Le spese concentrate in un particolare periodo dell'anno, con alte percentuali di famiglie che effettuano la spesa e con variabilità bassa, possono essere rilevate in modo affidabile anche con finestre temporali ridotte (è il caso ad esempio del canone di abbonamento alla TV che, pagato generalmente nel solo mese di gennaio da circa l'80% delle famiglie, fornisce una stima affidabile anche con un riferimento breve), mentre le spese concentrate in un particolare periodo dell'anno, con basse percentuali di famiglie che effettuano la spesa e variabilità alta, necessitano di una finestra temporale più ampia.

Sulla base della lunga serie di dati a disposizione e anche della ricognizione di quanto effettuato dagli altri paesi europei, sono state definite specifiche finestre temporali per ciascuna voce di spesa.

In generale i periodi di riferimento sono stati ampliati rispetto alla rilevazione condotta fino al 2013, solo in alcuni casi sono rimasti uguali.

1.3.3.1 La definizione dei periodi di riferimento per le spese correnti

Nel 2004 è stata condotta l'Indagine sulle frequenze di spesa (FREQ), che, tramite un diario autocompilato, ha rilevato gli eventi di spesa (non l'ammontare)

1. Il processo di ristrutturazione dell'indagine

delle famiglie per un periodo di 30 giorni; ciò ha consentito di verificare la rappresentatività di una finestra temporale settimanale nel descrivere la spesa mensile o piuttosto la necessità di ampliare il periodo di rilevazione. A parità di classificazione delle voci, è stato possibile confrontare il comportamento di spesa mensile stimato a partire dal modello a 7 giorni⁷ con quello ottenuto osservando l'intero mese. L'approssimazione migliora all'aumentare della frequenza di acquisto: i beni o servizi acquistati almeno una volta alla settimana vengono ben rappresentati attraverso un'osservazione settimanale, mentre per gli altri il grado di approssimazione diminuisce all'aumentare della distanza temporale tra un evento di acquisto e il successivo.

Le voci rilevate sono state classificate in 3 gruppi:

1. Beni/servizi acquistati almeno una volta la settimana (12 voci)
2. Beni/servizi acquistati almeno una volta in due settimane (46 voci)
3. Beni/servizi acquistati almeno una volta al mese (51 voci).

Sono solo 12 dunque le voci per le quali è possibile ricostruire in modo corretto il comportamento di spesa mensile con un diario a 7 giorni; un ampliamento del periodo di osservazione appariva quindi auspicabile, anche alla luce del fatto che, nella maggior parte dei Paesi che effettuano la rilevazione viene utilizzato un diario a 14 giorni. Tale scelta appare una buona soluzione di compromesso tra l'adeguata ricostruzione del comportamento mensile e il contenimento del carico di lavoro per i rispondenti. Per le voci con bassa frequenza di acquisto (anche nei 30 giorni) è possibile utilizzare quesiti retrospettivi purché l'entità della spesa sia tale da poter essere ricordata anche per un lasso di tempo superiore alle due settimane.

La fattibilità di un ampliamento del diario a 14 giorni è stato testato sia con l'indagine sperimentale per la rilevazione delle spese correnti del 2006, sia con l'indagine pilota del 2008 e ha mostrato in entrambi i casi buone performance.

I coefficienti di regressione dei modelli sottostanti gli andamenti nel tempo delle singole voci (frequenze giornaliere) non sono mai statisticamente diversi da zero, ad indicare andamenti pressoché costanti⁸.

Anche i test di significatività⁹ sulle differenze nei valori di spesa media e nelle frequenze di spesa della prima e della seconda settimana risultano significativi solo per alcune voci.

L'analisi è stata ulteriormente approfondita confrontando le spese medie dei singoli giorni nelle due settimane rilevate. Si è proceduto cioè ad una sorta di destagio-

7 L'indagine PAPI sui consumi delle famiglie si avvale di un modello complesso per il calcolo di coefficienti di espansione delle spese dai 7 giorni rilevati al mese basato sulle effettive frequenze di acquisto campionarie (Crescenzi, Ceccarelli, 1999; Rinaldelli, 2004).

8 Al fine di fornire una misura della sostanziale stabilità osservata nell'andamento delle spese medie è stata stimata la retta di regressione rispetto al tempo ($y=at+b$ dove y rappresenta la spesa media e t il tempo) sulle spese medie giornaliere. Il valore del coefficiente a rappresenta l'inclinazione della retta interpolatrice e, se positivo (negativo) e significativamente diverso da zero, indica un andamento crescente (decrescente). L'elaborazione è stata effettuata distinguendo le famiglie rilevate a ottobre da quelle rilevate a novembre, in modo che l'effetto dei giorni festivi fosse più evidente. Il coefficiente non risulta significativamente diverso da zero né in un mese né nell'altro, evidenziando quindi un andamento pressoché costante: con il passare dei giorni non si osservano né aumenti né diminuzioni nella spesa media delle famiglie.

9 Test di ipotesi non parametrico (test di permutazione) per le spese medie e test per dati appaiati (confronto tra proporzioni) per le frequenze.

nalizzazione delle spese, calcolando, in ciascun mese di indagine e per ogni voce di spesa, le differenze settime (Δ^7) nel seguente modo¹⁰:

$$\Delta_i^7 = s(g_i^{II}) - s(g_i^I)$$

Dove:

i indica il generico giorno delle settimana

$s(g_i^I)$ è la spesa media nell' i -esimo giorno dei primi 7 giorni di rilevazione

$s(g_i^{II})$ è la spesa media nell' i -esimo giorno dei secondi 7 giorni di rilevazione

Un valore positivo di Δ_i^7 indica che la spesa media delle famiglie nell' i -esimo giorno della prima settimana è inferiore a quella dell' i -esimo giorno della seconda settimana; viceversa quando il valore è negativo.

Le differenze settime sono state utilizzate per verificare l'eventuale presenza di una progressiva riduzione delle spese registrate nel diario.

A tal fine è stata calcolata la media delle Δ_i^7 (Tavola 1.1) e si osserva come per il mese di ottobre nella prima settimana, ogni giorno, si spendono 5,10 euro in più rispetto alla seconda; per il mese di novembre il valore è pari a 4,28.

Se la differenza viene resa relativa (viene cioè rapportata alla media delle $s(g_i)$) ne risulta che, in media, la spesa nell' i -esimo giorno della seconda settimana è del 12,6% inferiore a quella della prima nel mese di ottobre e del 10,4% in quello di novembre. Si può inoltre notare che il valore delle differenze settime risulta crescente e che nel settimo giorno è addirittura positivo ad indicare che l'eventuale effetto stanchezza è più evidente nella parte centrale del periodo piuttosto che all'inizio o alla fine.

Le differenze settime possono essere calcolate per ogni voce di spesa:

$$\Delta_{ji}^7 = s(g_{ji}^{II}) - s(g_{ji}^I)$$

dove:

j indica la generica voce di spesa

i indica il generico giorno delle settimana

$s(g_{ji}^I)$ è spesa media nell' i -esimo giorno dei primi 7 giorni di rilevazione per la j -esima voce

$s(g_{ji}^{II})$ è spesa media nell' i -esimo giorno dei secondi 7 giorni di rilevazione per la j -esima voce

¹⁰ Per tale approfondimento si ringrazia Riccardo Gatto per i preziosi suggerimenti.

Tavola 1.1 - Indicatori sulle differenze settime - Ottobre, Novembre 2006

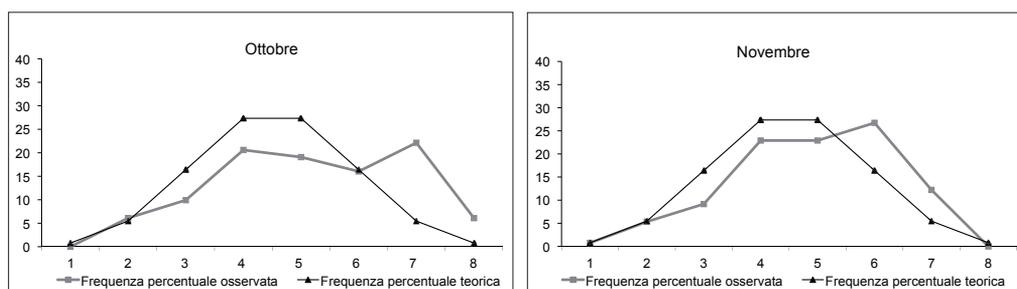
	Ottobre	Novembre
Giorno 1	-14,05	-10,07
Giorno 2	-10,79	-11,34
Giorno 3	-1,73	-9,58
Giorno 4	-6,25	-4,12
Giorno 5	-3,11	-1,89
Giorno 6	-4,16	0,78
Giorno 7	4,36	6,26
Media delle differenza settime delle spese medie giornaliere	-5,10	-4,28
Media delle spese medie giornaliere	40,53	41,0
Rapporto tra media delle differenze e numero medio di questionari con almeno una spesa	-12,6	-10,4

Fonte: Indagine sperimentale per la rilevazione delle spese correnti

Tenendo conto che, per ogni j -esima voce, il numero delle differenze settime negative può variare tra 0 (nella seconda settimana le spese medie sono sempre superiori a quelle della prima) e 7 (nella seconda settimana le spese medie sono sempre inferiori a quelle della prima), si calcola la frequenza delle voci di spesa rispetto al numero di differenze settime negative. In assenza di effetti esterni di disturbo (effetto stanchezza, effetto festività infrasettimanali, ecc.) le frequenze si distribuiscono in modo casuale e quindi come una binomiale.

La distribuzione osservata viene quindi confrontata con la distribuzione teorica binomiale attraverso un test che permette di valutare in modo sintetico la significatività della differenza nelle due distribuzioni (Grafico 1.5).

Grafico 1.5 - Distribuzione delle differenze settime negative (teorica e osservata) - Ottobre, Novembre 2006



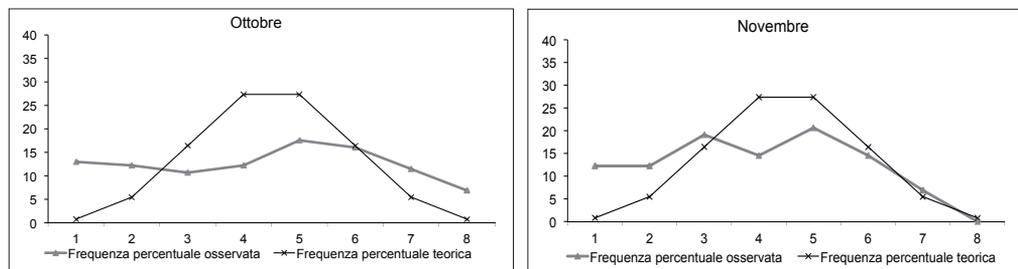
Fonte: Indagine sperimentale per la rilevazione delle spese correnti

Il test indica che la distribuzione osservata è significativamente diversa dalla binomiale in entrambi i mesi di rilevazione e, essendo asimmetrica a destra, lascia ipotizzare un lieve effetto stanchezza.

Analoghi risultati si ottengono se il procedimento viene applicato alle frequenze di spesa giornaliera.

Se tuttavia si pongono a zero le differenze settime non statisticamente significative, la distribuzione delle differenze settime negative che si ottiene risulta statisticamente diversa da una binomiale, ma per effetto di una asimmetria sinistra che esclude la possibilità di un effetto stanchezza (Grafico 1.6).

Grafico 1.6 - Distribuzione delle differenze settime significativamente negative (teorica e osservata) – Ottobre, Novembre 2006



Fonte: Indagine sperimentale per la rilevazione delle spese correnti

I valori del *daily transaction index* oscillano tra un minimo di 95 e un massimo di 103, risultando del tutto in linea con quanto osservato in altri Paesi. Si evidenzia, anche nel caso italiano, un effetto primo giorno¹¹; che può dipendere dal fatto che le famiglie, “dimenticando” di girare pagina, concentrano la registrazione delle spese sul primo giorno, ma anche da un effetto *telescoping*, per cui le famiglie tenderebbero a riportare sul primo giorno spese effettuate prima dell’inizio del periodo di rilevazione.

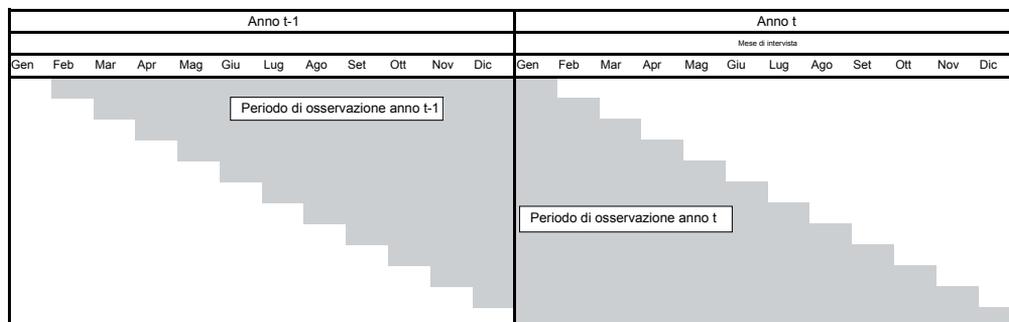
Il diario è stato quindi riorganizzato, passando da un’ottica giornaliera ad una settimanale, onde evitare che le famiglie “dimentichino” di girare pagina. Oltre a ridurre il numero di pagine, ciò consente alla famiglia di avere di fronte l’intero periodo di rilevazione rendendo più facile il ricordo di spese non immediatamente registrate.

1.3.3.2 La definizione dei periodi di riferimento per le altre spese

Come già accennato, la diversa frequenza ed entità che caratterizzano le spese rendono necessario l’uso di periodi di riferimento diversi, al fine di assicurare adeguata rappresentatività statistica alle stime. Inoltre, essendo l’indagine continua, al fine di valutare anche l’effetto della stagionalità (il mese di intervista è variabile), è stato necessario utilizzare periodi di riferimento mobili: la Figura 1.1 mostra lo schema di rilevazione nel caso si utilizzi un periodo di riferimento mobile e un mese di intervista variabile (per semplicità illustrativa si fa riferimento al periodo più ampio, pari agli ultimi 12 mesi). Tale scelta implica che il periodo di riferimento è composto da mesi di calendario diversi a seconda del mese di rilevazione e che, a conclusione dell’indagine, le spese si riferiscono di fatto agli ultimi 23 mesi (da febbraio dell’anno t-1 a dicembre dell’anno t).

¹¹ Per gli Stati Uniti Silberstein e Scott (op. cit.,1991) osservano che il primo giorno è sempre nettamente più alto del secondo e poi anche degli altri.

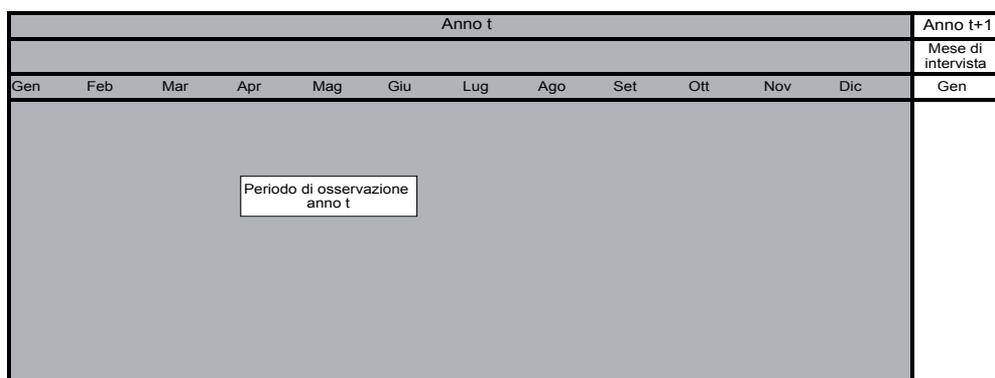
Figura 1.1 – Schema di rilevazione delle spese con periodo di riferimento mobile e mese di intervista variabile



- Per ovviare a tale inconveniente, le possibili soluzioni potrebbero essere:
- utilizzare solo le sole spese effettuate nel periodo di osservazione dell'anno t. Tuttavia, ciò implicherebbe una notevole perdita di informazione e non irrilevanti problemi di definizione del campione e di affidabilità delle stime;
 - concentrare l'indagine in un unico mese di rilevazione, chiedendo alle famiglie le spese riferite ai 12 mesi precedenti (in tal caso il periodo di riferimento sarebbe uguale per tutte). Tale soluzione, tuttavia, oltre a presentare notevoli problemi organizzativi, non assicurerebbe un'adeguata rappresentazione della stagionalità e dei diversi comportamenti di spesa nel corso dell'anno per le spese correnti che, inevitabilmente, sarebbero riferite all'unico mese di rilevazione (Figura 1.2).

Con entrambe le soluzioni, le spese si riferirebbero al solo anno t, ma comporterebbero distorsioni notevoli ed effetti importanti sul numero di eventi osservati e, quindi, sulla variabilità della stima (Eurostat, 2003).

Figura 1.2 – Schema di rilevazione delle spese con periodo di riferimento fisso (anno t) e mese di intervista fisso



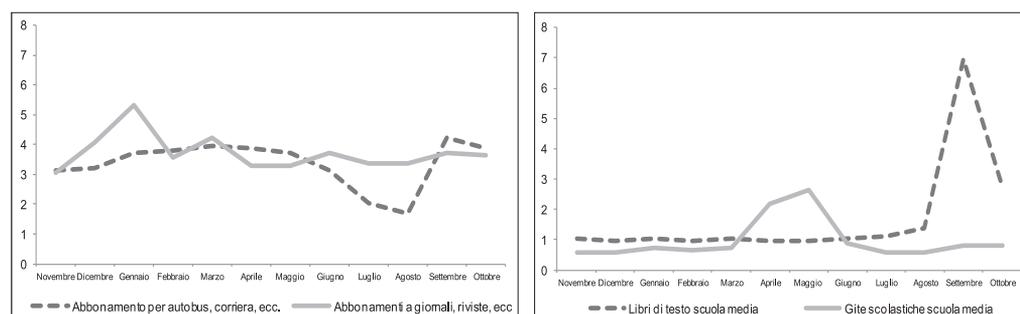
Per le voci di spesa con cadenza regolare ma non troppo frequente (ad esempio bollette, affitto, condominio), può essere utile fare riferimento al cosiddetto approccio dell'ultimo pagamento: si chiede alla famiglia di indicare l'ammontare dell'ultimo pagamento effettuato (nel periodo di riferimento) e la cadenza con cui viene generalmente fatto. Tale soluzione, fornisce l'informazione desiderata alleggerendo notevolmente il *burden* sul rispondente (Eurostat, 2003)

L'ampliamento dei periodi di riferimento nei quesiti retrospettivi è stato accompagnato da un approfondimento circa l'effettivo mese di effettuazione della spesa (calendarizzazione delle spese). Tale informazione consente, oltre che di tenere sotto controllo la stagionalità degli eventi di spesa, di effettuare analisi anche per periodi inferiori a quello di rilevazione. Ciò è particolarmente importante nel caso italiano per ricostruire le spese a livello trimestrale e fornire così le stime per i consumi finali delle famiglie (cfr. par. 2.5.1).

L'Indagine DIR del 2006 ha fornito le prime valutazioni sugli effetti delle modifiche introdotte nei periodi di riferimento per le voci rilevate con quesiti retrospettivi. In particolare, si è escluso un evidente effetto ricordo sia in termini di frequenze di spesa, sia in termini di spese medie effettive. Il 13% delle famiglie, infatti, pur dichiarando problemi a ricordare la spesa, ha poi fornito l'informazione grazie al supporto di bollette, ricevute o altri titoli di pagamento; inoltre, solo il 7% delle famiglie dichiara di aver avuto difficoltà a collocare le spese nel tempo e comunque fornisce l'informazione.

In effetti l'analisi dei trend mensili, sia nel 2006 sia nel 2008, non evidenzia importanti cadute riconducibili ad un effetto ricordo o ad un effetto *telescoping* connessi con l'ampliamento dei periodi di riferimento. Gli andamenti colgono adeguatamente la stagionalità delle singole voci di spesa (Grafico 1.7) e i livelli di spesa media sono del tutto confrontabili con quelli osservati negli stessi periodi con l'indagine PAPI sui consumi. Per quanto riguarda il numero di eventi osservati (le frequenze di spesa), questi sono molto simili quando i periodi di riferimento rimangono inalterati (ad esempio per l'energia elettrica della prima abitazione, le frequenze di spesa sono pari al 95% con la sperimentale e al 98% con l'indagine sui consumi, per la seconda abitazione sono rispettivamente del 7,7% e del 6,3%).

Grafico 1.7 – Famiglie che hanno effettuato la spesa per alcune voci e mese di effettuazione – Novembre 2006



Fonte: Indagine DIR 2006

Laddove invece il periodo di riferimento viene ampliato le frequenze sono, come ipotizzato, più elevate: nel 2008 per la tinteggiatura dell'abitazione principale si passa dal 4,6% osservato con l'indagine sui consumi (con riferimento agli ultimi tre mesi) al 21,5% osservato con la sperimentazione (negli ultimi dodici mesi), per l'acquisto di telefono cellulare dal 3,4% (negli ultimi tre mesi) al 24,3% (negli ultimi dodici mesi), per l'acquisto di lavatrice dall'1,2% al 6,9%.

In conclusione, dunque, l'ampliamento dei periodi di riferimento e la calendarizzazione delle spese sono le soluzioni adottate per avere maggiore robustezza nelle stime e permettere anche la stima trimestrale.

1. Il processo di ristrutturazione dell'indagine

1.3.4 I contenuti informativi

1.3.4.1 Le *core variables* europee

L'obiettivo delle HBS è quello di fornire una descrizione della spesa per consumi privati con un sufficiente dettaglio e in funzione delle caratteristiche socio-economiche delle famiglie (l'ampiezza e la composizione della famiglia, la zona di residenza, il livello di reddito familiare, l'età, il sesso, il grado di istruzione e la condizione lavorativa dei singoli componenti). In Italia, a partire dalla fine degli anni '80, la statistica sociale ha sperimentato una rapida evoluzione (Istat, 2006a), in Europa lo sviluppo è più recente e dettato dalla necessità da parte dei *policy makers* europei di supportare le proprie scelte con informazioni su mercato del lavoro, condizioni di vita, movimenti di popolazione, protezione sociale, percorsi formativi, condizioni di salute, criminalità e su tutti gli altri aspetti di una società complessa. Ciò ha reso sempre più evidente, da una parte, il bisogno di informazioni integrate e, dall'altra, la necessità di standardizzare concetti e definizioni nell'ambito degli strumenti statistici di rilevazione. È stato dunque implementato un set di *core variables* (Prospetto 1.1) comune a tutte le indagini sociali effettuate in tutti i Paesi dell'unione (Eurostat, 2011).

L'introduzione delle *core variables* nell'Indagine sulle spese è stata dunque testata attraverso diverse sperimentazioni, soprattutto riguardo quesiti complessi quali la condizione lavorativa oggettiva e soggettiva, la professione svolta, il reddito.

Anche la necessità di standardizzazione interna dei processi ha richiesto uno studio approfondito circa la possibilità di rilevare, assieme alla spesa, informazioni che costituissero una base comune in una prospettiva di *matching* e di *linkage* statistici tra fonti. Le condizioni abitative, il luogo prevalente di acquisto, le caratteristiche dei viaggi e dei viaggiatori sono stati oggetto di studio in fase di sperimentazione.

1.3.4.2 La rilevazione del reddito

L'obiettivo, nel rilevare il reddito, è quello di ottenere una proxy delle condizioni economiche dei rispondenti che non dipenda unicamente dal reddito individuale, ma anche da quello di tutti gli altri membri della famiglia. Per questa ragione la variabile di interesse per misurare il benessere economico risulta essere il reddito netto familiare, che fornisce informazioni sulla disponibilità economica effettiva della famiglia e quindi sulla possibilità che ha di spendere o risparmiare (Eurostat, 2011).

Eurostat, oltre ad annoverare tale variabile tra le *core variables*, raccomanda la sua introduzione nell'ambito delle HBS (Eurostat, 2012), come variabile di classificazione socio-economica della famiglia, escludendone tuttavia l'uso al fine di dar conto del livello e dell'andamento della condizione reddituale nel tempo o di effettuare comparazioni tra paesi (in tal senso la fonte di riferimento è l'indagine EU-Silc). Per le HBS, inoltre, Eurostat raccomanda di associare alla rilevazione del reddito quella del fitto imputato, al fine di rendere direttamente confrontabili le condizioni economiche degli affittuari e dei proprietari¹².

In un'ottica di armonizzazione, è stata dunque effettuata un'analisi di fattibilità per la rilevazione del reddito netto familiare e delle informazioni necessarie alla stima tramite modello dei fitti imputati.

¹² Commission Decision of 18 July 1995 (95/309/EC, Euratom).

Poiché in Italia la rilevazione del reddito, in particolare in associazione con la rilevazione delle spese, viene percepita dalle famiglie come decisamente intrusiva, il quesito è stato inserito alla fine dell'intera batteria di domande, in modo da non inficiare il resto della rilevazione.

Con l'indagine del 2012, è stato chiesto alle famiglie di indicare il reddito netto medio mensile familiare in modo puntuale o, solo nel caso non fossero in grado di indicarlo, tramite un valore approssimato.

Nel 2012, circa il 10% del campione non ha risposto al primo quesito e, tra questi, il 45% circa ha continuato a non rispondere anche alla richiesta di una indicazione approssimativa; per queste ultime famiglie la spesa mensile equivalente è più elevata rispetto a quella di chi fornisce l'indicazione sul reddito (circa il 18,5% in più): nel 27% dei casi sono famiglie con persona di riferimento lavoratore indipendente (sono il 24% tra le famiglie che forniscono l'indicazione sul reddito) e il 77% risiede nel Centro-Nord (il 59% tra le famiglie che forniscono l'indicazione).

La variabile presenta dunque delle indubbie e note difficoltà di rilevazione (Eurostat, 1997 e 2011) e risulta distorta a causa della più elevata mancata risposta per particolari fasce di popolazione. La sua rilevazione, tuttavia, è estremamente utile in un'ottica di *matching* statistico con altre fonti, nell'ipotesi che venga riprodotta la stessa distorsione.

La stima del fitto imputato si riferisce al costo che le famiglie proprietarie dell'abitazione in cui vivono (ma anche in usufrutto, in uso gratuito o con contratti di fitto agevolato) dovrebbero sostenere per prendere in affitto, ai prezzi di mercato, un'unità abitativa con caratteristiche identiche a quella in cui vivono (escludendo le spese di condominio, di riscaldamento, quelle accessorie o riferite all'eventuale presenza di mobilio).

L'indagine sui consumi in Italia ha sempre rilevato il valore del fitto imputato in modo soggettivo, chiedendo cioè alle famiglie di fornire un'autovalutazione del valore dell'affitto che avrebbero ottenuto dando in affitto la propria abitazione, tenendo conto dei prezzi del mercato nella zona di residenza. Le ipotesi alla base di tale scelta sono fondamentalmente due: a) la percentuale di famiglie che vive in abitazioni in affitto è molto contenuta (inferiore al 20%) e di conseguenza la stratificazione dei prezzi di affitto sarebbe certamente poco significativa rispetto ad alcune categorie poco rappresentate nel campione; b) la variabilità territoriale dei prezzi di affitto è talmente forte che è difficilmente stimabile attraverso altre fonti.

L'affitto figurativo stimato tramite autovalutazione comporta tuttavia noti problemi legati proprio alla sua natura soggettiva (Eurostat, 2010). I rispondenti potrebbero non conoscere in maniera adeguata le reali condizioni del mercato, proprio per il fatto che non ne usufruiscono; ciò potrebbe indurle a sovrastimarne/sottostimarne il valore (Ceccarelli, Cuttillo, Di Laurea, 2009) o a scoraggiarsi non fornendo l'informazione (nell'indagine sui consumi le famiglie che non forniscono una stima del fitto figurativo sono il 9,7% nel 2012 e il 10,5% nel 2013).

Per questo motivo Eurostat (2012) raccomanda l'uso di metodi di stratificazione con cui stimare i costi di affitto osservati su gruppi di popolazione omogenei in base ad alcune caratteristiche: i) titolo di godimento, ii) tipo di contratto, iii) tipo di fabbricato, iv) anno di costruzione, v) numero di stanze, vi) superficie dell'abitazione in metri quadri, vii) anno di occupazione, viii) dominio di popolazione. È inoltre opportuno associare a queste una serie di altre variabili che consentono di applicare metodi di

stima econometria (quali regressione edonica con o senza la procedura di Heckman per la correzione del *sample selection bias*). Le variabili raccomandate, in questo caso, sono variabili che descrivono ulteriori caratteristiche dell'abitazione e della zona di residenza (presenza di garage, indicatori di qualità e confort dell'abitazione), disponibilità di servizi in prossimità dell'abitazione, tipo di affaccio e zona climatica che vanno a loro volta affiancate alle caratteristiche familiari (ampiezza, tipologia, sesso, età, stato civile e le altre caratteristiche della persona di riferimento).

Per l'indagine sui redditi e le condizioni di vita (EU-Silc), si utilizza uno *stratification method* con procedura di Heckman, un modello econometrico, cioè, che analizza la relazione statistica fra le caratteristiche delle abitazioni (tipo di abitazione, dotazioni e servizi, sia in termini quantitativi sia qualitativi) e gli affitti pagati dagli inquilini, al fine di ottenere il valore ipotetico dell'affitto da imputare ai proprietari di casa, sulla base delle caratteristiche delle loro abitazioni (Istat, 2014).

Nell'Indagine sulle spese per consumi sono stati dunque inserite le principali variabili utilizzate dal modello econometrico, al fine di poter pervenire alla stessa stima del fitto imputato utilizzata nel contesto dell'indagine EU-Silc.

1.3.4.3 La situazione lavorativa

La situazione lavorativa degli individui è strettamente correlata con i comportamenti di spesa delle famiglie. Con l'indagine PAPI sui consumi si rilevava la condizione lavorativa soggettiva al momento dell'intervista (così come percepita dal rispondente), la posizione nella professione (distinguendo tra lavoratori dipendenti e indipendenti), il settore di attività economica (in base alle 17 sezioni dell'ATECO 2002).

La condizione lavorativa soggettiva, tuttavia, non è quella correntemente utilizzata in altre indagini; in particolare, l'Indagine sulle forze di lavoro utilizza i criteri stabiliti dall'ILO¹³ e considera occupate le persone di 15 anni e più che nella settimana di riferimento:

- hanno svolto almeno un'ora di lavoro in una qualsiasi attività che preveda un corrispettivo monetario o in natura;
- hanno svolto almeno un'ora di lavoro non retribuito nella ditta di un familiare nella quale collaborano abitualmente¹⁴.

In base agli stessi criteri ILO, vengono considerate disoccupate le persone non occupate tra i 15 e i 74 anni che:

- hanno effettuato almeno un'azione attiva di ricerca di lavoro nelle quattro settimane che precedono la settimana di riferimento e sono disponibili a lavorare (o ad avviare un'attività autonoma) entro le due settimane successive;

13 I concetti e le definizioni impiegati nell'indagine continua sulle forze di lavoro si basano sulle raccomandazioni emerse nei lavori della 13° Conferenza internazionale degli statistici del lavoro, organizzata dall'International Labour Organisation nel 1982. Il regolamento comunitario che ha recepito tali raccomandazioni è il No 577/98 del 9 marzo 1998. Si veda, al riguardo, Commission Regulation (EC) No 1897/2000 of 7 September 2000 implementing Council Regulation (EC) No 577/98 on the organisation of a Labour force sample survey in the Community concerning the operational definition of unemployment.

14 I dipendenti assenti dal lavoro (ad esempio, per ferie o malattia) sono considerati occupati se l'assenza non supera tre mesi, oppure se durante l'assenza continuano a percepire almeno il 50% della retribuzione. Gli indipendenti assenti dal lavoro, ad eccezione dei coadiuvanti familiari, sono considerati occupati se, durante il periodo di assenza, mantengono l'attività. I coadiuvanti familiari sono considerati occupati se l'assenza non supera tre mesi.

- oppure, inizieranno un lavoro entro tre mesi dalla settimana di riferimento e sarebbero disponibili a lavorare (o ad avviare un'attività autonoma) entro le due settimane successive, qualora fosse possibile anticipare l'inizio del lavoro.

Il dato così rilevato differisce sostanzialmente da quanto ottenuto attraverso la rilevazione soggettiva; i due dati risultano complementari descrivendo aspetti diversi della condizione lavorativa. La valutazione oggettiva garantisce uno standard confrontabile nello spazio e nel tempo, ma non fotografa sfumature che invece è possibile rappresentare attraverso una valutazione soggettiva che tiene conto di aspetti diversi. La condizione lavorativa soggettiva, infatti, riflette ciò che il rispondente effettivamente sente di essere. Si tratta, cioè, di una valutazione individuale e fa riferimento allo stato corrente dell'individuo o allo stato che egli percepisce come il suo stato normale (al momento dell'intervista o, più in generale, in un tempo più ampio prima e/o dopo l'intervista). Alcune persone, ad esempio, nonostante svolgano lavori part-time che li occupano per poche ore alla settimana, potrebbero continuare a considerarsi studenti, disoccupati o casalinghe, mentre altri potrebbero decidere di dichiararsi occupati part-time. In altri casi ancora, nonostante una breve interruzione dell'attività lavorativa al momento dell'intervista, le prospettive per il rispondente possono essere tali da indurlo a considerarsi comunque occupato.

Con l'indagine CAPI del 2012, è stata dunque sperimentata la possibilità di affiancare alla rilevazione soggettiva, percepita dai rispondenti, quella oggettiva, inserendo un set di quesiti armonizzato a quanto rilevato dall'Indagine sulle forze di lavoro. Ne risulta che il 3% delle persone oggettivamente occupate si dichiarano in cerca di occupazione e quasi il 2% in altra condizione. Tra gli oggettivamente disoccupati, d'altro canto, poco più dell'1% si ritiene occupato, mentre il 14% preferisce dichiararsi in altra condizione.

Rispetto a quanto richiesto da Eurostat in termini di *core variables*, con la PAPI non veniva rilevata la professione; tale variabile risulta infatti piuttosto complessa da rilevare attraverso uno strumento cartaceo. Le tecniche CA consentono invece di approfondire tale aspetto con un relativo aggravio del carico di lavoro per i rispondenti, garantendo inoltre standard qualitativi elevati. La complessità di questa domanda si ricava facilmente se si considera che la classificazione sottostante il fenomeno in esame comprende più di 20mila nomi di professioni, raggruppati, al suo interno, in 800 unità professionali. Di fronte a questo scenario sono state valutate due possibili opzioni metodologiche. La prima, sulla falsariga dell'esperienza censuaria condotta nel 2011, prevedeva di rilevare la professione degli intervistati in modalità "gerarchica", ovvero scegliendo i raggruppamenti di volta in volta più aderenti al profilo interessato e fermandosi al secondo livello di dettaglio della tassonomia, quello corrispondente alla richiesta di Eurostat per le HBS. La seconda, invece, concepita come domanda aperta, consentiva di codificare al massimo dettaglio classificatorio (5° digit), richiedendo l'impiego del navigatore delle professioni, un applicativo esterno al questionario che agevola la ricerca dell'unità professionale adatta a ricomprendere la professione dell'intervistato attraverso l'impiego di un sofisticato motore di ricerca. Tale procedura di classificazione ricalcava quella già in uso nelle altre indagini sulle famiglie, pur presentando requisiti di implementazione più impegnativi. In primo luogo, perché presupponeva una formazione specifica dei rilevatori, basata su un

modulo di approfondimento teorico, corredato da una parte esercitativa sull'uso del navigatore. In secondo luogo, perché induceva gli intervistatori, la maggior parte delle volte, a scegliere la modalità di codifica 'a posteriori', ovvero dopo aver terminato l'intervista, potendo ricercare con più tempo a disposizione il codice migliore sulla base degli elementi raccolti.

La scelta adottata ricalca quella già in uso nelle altre indagini sulle famiglie e si basa su una domanda aperta e sull'uso del navigatore delle professioni per codificare al massimo dettaglio (5° digit); pur riconoscendo il maggiore impegno richiesto alla rete, sia in termini formativi sia in termini metodologici, tale opzione consente la piena comparabilità dei dati rilevati con quelli raccolti dalle altre indagini sociali. Inoltre, disporre dei codici a 5 digit delle professioni, in base alla classificazione italiana (Cp 2011), consente di applicare agevolmente la tabella di transcodifica e di ottenere in automatico i corrispondenti codici a 2 digit della Isco richiesti da Eurostat.

La sperimentazione ha confermato la tenuta della scelta adottata (Tavola 1.2): le distribuzioni degli occupati per professione ottenute dalla sperimentazione del 2012 e dalla rilevazione sulle forze di lavoro dello stesso anno risultano infatti assolutamente confrontabili. Anche la differenza registrata per il primo grande gruppo si giustifica alla luce della composizione esigua e peculiare di questo raggruppamento.

Nonostante la condizione lavorativa oggettiva e la professione non rappresentino variabili *clou* per l'Indagine sulle spese, la loro introduzione, la metodologia scelta e il relativo sistema di codifica assicurano l'accuratezza nella rilevazione e l'armonizzazione sia a livello internazionale sia con le altre indagini socio-economiche.

Tavola 1.2 - Occupati per grande gruppo professionale - Anno 2012 (valori percentuali)

Grandi gruppi	Spese	IFL
I - Legislatori, imprenditori e alta dirigenza	4,2	2,6
II - Professioni intellettuali, scientifiche e di elevata spec.	12,7	13,0
III - Professioni tecniche	16,9	17,7
IV - Professioni esecutive nel lavoro d'ufficio	11,4	11,8
V - Professioni qualificate in attività commerciali e servizi	17,7	18,1
VI - Artigiani, operai specializzati e agricoltori	19,0	16,7
VII - Conduttori di impianti, operai di macchinari e veicoli	6,4	8,2
VIII - Professioni non qualificate	10,3	10,7
IX - Forze armate	1,3	1,2
Totale	100,0	100,0

Fonte: Indagine sulle spese delle famiglie e Indagine sulle forze di lavoro

1.3.4.4 La rilevazione delle quantità autoconsumate e delle quantità acquistate

Con l'indagine PAPI la rilevazione degli autoconsumi alimentari avveniva tramite un diario dedicato (lasciato in compilazione alle sole famiglie che durante il contatto iniziale dichiaravano di consumare generi alimentari di produzione propria) in cui le famiglie annotavano, per sette giorni e per un elenco di voci, la quantità consumata e il relativo valore monetario (stimato dalla famiglia stessa in base ai prezzi di mercato praticati nella zona di residenza).

Come per il fitto figurativo, anche la stima degli autoconsumi risente del fatto che le famiglie chiamate a fornire la valutazione del loro autoconsumo in termini monetari,

in genere, non affrontano tali spese e quindi possono non essere consapevoli dei prezzi praticati sul mercato. Si è dunque valutata la possibilità di rilevare unicamente le quantità dei beni autoconsumati, ipotizzando di ottenere tramite fonti esterne, oggettive, il valore monetario delle quantità rilevate.

In tal senso era dunque necessario assicurare la massima accuratezza nella rilevazione delle unità di misura per ciascuna voce. Obbligare le famiglie ad utilizzare un'unica unità di misura può generare problemi ed errori da parte delle famiglie non abituate a quella specifica unità, nello stesso tempo lasciare completa libertà nell'uso delle unità di misura richiede tempi elevati di transcodifica in fase di trattamento dati.

La sperimentazione del 2008 ha testato la scelta dell'unità di misura per rilevare le quantità consumate dei generi alimentari autoprodotti lasciando alle famiglie la possibilità di scegliere tra tre diverse unità per ciascun bene.

Solo l'11,8% delle famiglie ha trovato molto o abbastanza difficile scegliere l'unità di misura tra quelle elencate.

La maggior parte degli eventi di consumo sono stati registrati in grammi (in media il 67,8%), riferimento utilizzato perlopiù in modo corretto: in base all'analisi delle distribuzioni il valore medio e quello mediano presentano ordini di grandezza simili e le code delle distribuzioni mostrano valori estremi, ma plausibili (ad esempio per la carne bovina a fronte di un valore medio pari a 371 grammi e ad una mediana di 300 grammi, il 99° percentile è pari a 1.500 grammi ed il 1° a 20 grammi). Per il 19,7% degli eventi sono stati utilizzati gli ettogrammi, anche se in questo caso gli errori appaiono molto più frequenti (sempre per la carne bovina ad una mediana pari a 4 etti corrisponde una media di 45 etti, un 99° percentile che arriva a 600 etti ed un 1° percentile a 1 etto). Solo il 12,6% degli eventi vengono registrati utilizzando come unità di misura i chilogrammi, unità spesso eccessiva per indicare un consumo quotidiano, ma funzionale ad esempio alla rilevazione delle quantità di beni prodotti che vengono regalate e che per definizione rientrano tra gli autoconsumi (il valore mediano per la carne bovina è pari a 1, la media a 7, il 99° percentile a 135 e il 1° a 1).

Per i liquidi (olio, aceto, latte e vino), le unità di misura proposte alle famiglie sono state oltre ai litri, i bicchieri e i cucchiari (da transcodificare in base ad opportune tavole di corrispondenza, SINU, 2014). I risultati dipendono fortemente dal tipo di bene: per il latte sono stati utilizzati in uguale misura i bicchieri e i litri, nonostante per questi ultimi la precisione è decisamente minore; per l'olio la preferenza è nettissima sui cucchiari (61,8%), seguiti dai bicchieri (30,7%), con una buona precisione per entrambi; risultati analoghi sono stati ottenuti per l'aceto (l'89,2% degli eventi viene registrato utilizzando i cucchiari), mentre per il vino ben il 70% degli eventi viene rilevato attraverso i bicchieri, in modo sufficientemente corretto. La scelta si è quindi orientata verso l'uso di due unità di misura per ogni bene, selezionando quelle maggiormente utilizzate dalle famiglie.

Sempre con riferimento agli autoconsumi, si è testata anche l'opportunità di inserire la rilevazione all'interno del diario utilizzato per la rilevazione delle spese correnti, in una sezione dedicata.

La sperimentazione condotta nel 2007 (l'Indagine sperimentale per la rilevazione delle quantità alimentari acquistate - QUA) ha mostrato la fattibilità dell'introduzione della rilevazione delle quantità alimentari autoconsumate nel diario, a patto che questa avvenga in sezioni nettamente separate da quelle dedicate alla rilevazione delle spese. La sperimentazione del 2008 ha confermato e incoraggiato l'uso di una sezione del diario dedicata agli autoconsumi.

Diversi enti, a livello nazionale e internazionale, hanno più volte fatto emergere la necessità di rilevare le quantità alimentari acquistate allo scopo di ottenere *proxy* sui comportamenti alimentari e nutrizionali delle famiglie. A tale fine l'indagine QUA ha testato, quindi, anche la fattibilità di introdurre le quantità anche per i beni e servizi acquistati e ha evidenziato, come atteso, la difficoltà delle famiglie a gestire la rilevazione. Sono stati evidenziati enormi problemi da parte delle famiglie nell'indicare la corretta quantità ma anche la corretta unità di misura; la richiesta di registrare, insieme alla spesa sostenuta, anche la quantità acquistata, ha determinato un evidente effetto stanchezza, con conseguente diminuzione delle frequenze di spesa. Per la maggior parte delle voci di spesa confrontabili, la frequenza di spesa rilevata con l'indagine QUA risulta decisamente più bassa di quella rilevata, nello stesso periodo, con l'indagine sui consumi (Tavola 1.3). Il 22% delle famiglie intervistate ha dichiarato di aver avuto problemi nella compilazione del diario, per l'86% di queste i problemi si riferivano proprio al dover riportare le quantità acquistate. I dati di spesa confermano l'evidenza; ad esempio, il valore mediano della quantità di latte acquistata è pari ad un litro, mentre il valore medio supera i 120 litri, a causa del fatto che molte famiglie hanno erroneamente indicato 1000 (valore riportato sulla confezione) in corrispondenza dei litri piuttosto che dei millilitri.

Tavola 1.3 – Famiglie che hanno effettuato la spesa per alcune voci di spesa alimentare - Il trimestre 2007
(valori percentuali)

Voce di spesa	Indagine QUA (primi 7 giorni di rilevazione)	Indagine sui consumi
Pasta	70,9	72,6
Riso	35,1	41,7
Farina	26,1	33,7
Biscotti, merendine	70,3	65,2
Gelati	28,0	49,6
Carne di maiale	27,4	40,0
Pesce fresco o surgelato	48,1	66,3
Pesce, frutti di mare o crostacei, essiccati, affumicati, sotto sale: salmone affumicato, aringa affumicata, ecc.	6,6	14,4
Olio di oliva	32,3	49,8
Olio di semi	17,7	17,5
Burro	23,2	33,9
Formaggi	76,7	80,7
Uova	41,4	59,5
Pomodori (freschi o surgelati)	61,9	71,9
Pomodori pelati, in polpa, in passata	26,7	27,7
Altri prodotti, preparati dietetici e piatti pronti a base di ortaggi e legumi	4,2	15,0
Agrumi	45,0	53,0
Banane	49,5	62,5
Mele	59,4	64,2
Pere	34,0	39,9
Frutta con nocciolo	24,8	34,9
Uva, fragole, frutti di bosco	27,7	38,8
Frutta secca e essiccata: prugne, mandorle, nocciole, noci, fichi, arachidi, pinoli, semi di zucca e altri semi, ecc.	7,5	10,1
Acqua minerale, acqua potabile venduta in contenitori o acquistata alla fonte	56,2	65,4
Vino	28,4	40,0
Birra	19,2	31,5

Fonte: Indagine sui consumi e indagine sperimentale per la rilevazione delle quantità alimentari acquistate (QUA).

L'introduzione delle quantità acquistate ha anche generato confusione rispetto alle quantità autoconsumate: la percentuale di famiglie con autoconsumi di generi alimentari ha superato il 34%, percentuale molto più elevata di quella registrata con l'indagine sui consumi (circa il 5%), ma anche con le altre sperimentazioni (il 20% con la sperimentazione del 2008, il 21,4% con la sperimentazione del 2010). Si è stabilito, dunque, di non rilevare le quantità acquistate attraverso l'Indagine sulle spese, al fine di salvaguardare la corretta rilevazione delle quantità autoconsumate e della spesa sostenuta per gli acquisti.

La classificazione utilizzata per la rilevazione degli autoconsumi ricalca esattamente la Coicop-HBS in modo che il valore monetario associato alla quantità consumata possa essere sommato alla spesa rilevata, stimando così il valore totale dei beni consumati (siano essi autoprodotti o meno).

1.3.4.5 I luoghi di acquisto

Ogni famiglia mette in atto comportamenti di spesa distinti – acquistando beni e servizi differenti per caratteristiche e quantità, attraverso diverse tipologie distributive – in base al proprio reddito disponibile, alla condizione professionale e dell'età dei componenti, alla zona in cui vive, alle proprie preferenze. L'effetto della dinamica dei prezzi sulla capacità di acquisto di ogni famiglia è quindi diverso a seconda delle abitudini di spesa (si veda ad esempio Istat 2014a sugli indicatori della dinamica dei prezzi al consumo per alcune tipologie di famiglie).

La possibilità di disporre di informazioni circa i luoghi in cui le famiglie acquistano i prodotti, oltre a consentire un'analisi più approfondita dei comportamenti di spesa delle famiglie, è funzionale alla messa a punto di metodologie adeguate per il calcolo di indici della dinamica dei prezzi per tipologia distributiva.

Le sperimentazioni sono dunque servite a valutare la fattibilità di rilevare presso le famiglie, almeno per una parte delle voci rilevate, il luogo in cui è stato effettuato l'acquisto. Le tipologie distributive e le loro definizioni sono state concordate con i ricercatori che si occupano della rilevazione dei prezzi, al fine di renderle fruibili anche in quel contesto¹⁵. Per le voci rilevate tramite diario si è chiesto alle famiglie, a metà

15 **Negozi** tradizionale: è un locale generalmente di dimensioni contenute, raggiungibile direttamente dalla strada o inserito in un centro commerciale, adibito all'esposizione e alla vendita di merci al minuto, dove normalmente il cliente viene servito (ad esempio forni, macellerie, ferramenta, gioiellerie, tabaccai, farmacie, giornali, fiorai). **Mercato rionale e ambulanti**: è un luogo, spesso all'aperto ma anche all'interno di una struttura coperta, con un'area più o meno vasta che ospita anche punti vendita fissi. **Ambulante**: si intende un commerciante che non esercita la propria professione in un luogo stabile, ma sposta la propria postazione di vendita (bancarella, banco, furgone attrezzato, furgone-bar, ecc.) sul territorio. **Ipermercato, supermercato e grande magazzino non specializzato** *Ipermercato*: è un'area di vendita molto ampia e con una vasta gamma di merci di largo consumo (alimentare e non alimentare), spesso collocato alla periferia di una città, presso vie di grande comunicazione, fornito di servizi complementari per la clientela (parcheggio, ristorazione, ecc.), dove generalmente ci si serve da soli. *Supermercato*: si differenzia dall'ipermercato per le dimensioni generalmente inferiori, per il fatto di essere più facilmente raggiungibile e di non essere fornito di servizi complementari per la clientela. *Grande magazzino non specializzato*: locale di ampie dimensioni, raggiungibile direttamente dalla strada o inserito in un centro commerciale, nel quale è possibile acquistare più tipi di prodotti (ad esempio abbigliamento insieme ad arredamento, prodotti cosmetici e giocattoli). **Hard discount**: è un supermercato che vende una gamma limitata di prodotti di largo consumo non pubblicizzati e a prezzi più contenuti rispetto a quelli mediamente praticati da supermercati e ipermercati. **Grande magazzino specializzato e catene di negozi**: si differenzia dal grande magazzino non specializzato per il fatto di vendere esclusivamente un particolare tipo di prodotto, con una vasta gamma di offerta (ad esempio

1. Il processo di ristrutturazione dell'indagine

e alla fine del periodo di compilazione, di indicare, per un dato elenco di beni, dove li avessero acquistati con maggiore frequenza nei 7 giorni appena trascorsi (era possibile indicare al massimo due luoghi). Per le voci rilevate con l'intervista diretta il luogo di acquisto è stato richiesto contestualmente alla spesa sostenuta.

Sulla base dei risultati di una prima sperimentazione (l'indagine QUA del 2007), è stata condotta l'indagine del 2008 che ha evidenziato come, in generale, le famiglie siano assolutamente in grado di indicare il luogo di acquisto: più del 90% ha dichiarato di non aver avuto problemi nel comprendere le definizioni (molto o abbastanza chiare) ed oltre il 95% fornisce, per i beni acquistati, sia l'ammontare della spesa sia il luogo in cui il bene è stato acquistato, senza un significativo effetto stanchezza.

La rilevazione dei luoghi d'acquisto può essere dunque affiancata, con le modalità testate, alla rilevazione della spesa.

1.3.4.6 L'approfondimento sulle caratteristiche dei viaggi e delle escursioni

L'ipotesi di includere un modulo specifico sul turismo nelle HBS era stata in passato esplorata dal World Tourism Organization (Unwto) in collaborazione con l'International Labour Organization (Ilo)¹⁶, che avevano individuato, nel turismo, una potenziale area di interesse nel *framework* delle indagini sul reddito e sulle spese delle famiglie. L'Unwto riteneva, in particolare, che quest'ultima fosse lo strumento ideale per rilevare anche le spese turistiche, poiché si tratta di un'indagine svolta nella maggioranza dei Paesi e appositamente disegnata per rilevare le spese¹⁷.

Tuttavia, una serie di difficoltà si incontrano per realizzare fattivamente tale potenzialità:

- le rilevazioni del fenomeno turistico devono necessariamente riferirsi alle definizioni concettuali e metodologiche delle International Recommendations on Tourism Statistics (Irts) e della metodologia comunitaria, ove il “turista” viene distinto dal più ampio insieme dei “visitatori” applicando il concetto di “ambiente abituale”, generalmente non definito nell'ambito delle HBS;
- l'approccio per la rilevazione delle spese generalmente utilizzato nelle HBS non consente di distinguere le spese effettuate durante un viaggio dagli altri tipi di spese della famiglia, poiché non viene richiesto di specificare dove¹⁸ queste spese siano state effettuate;
- la rilevazione delle spese a livello familiare non consente di distinguere chi abbia effettuato il viaggio e quindi sostenuto la relativa spesa, tra i singoli componenti della famiglia;
- le spese connesse con un viaggio di lavoro, con l'eccezione degli esborsi a titolo personale, non devono essere rilevate con l'indagine HBS, mentre sono oggetto

catene di negozi sportivi, di elettrodomestici, di elettronica, di arredamento). **Internet:** ci si riferisce a tutti gli acquisti effettuati tramite internet per qualsiasi tipo di bene o servizio. **Altro luogo:** ci si riferisce a tutti i luoghi diversi da quelli precedentemente descritti, dove è possibile acquistare beni e servizi come ad esempio biglietterie ferroviarie, ambulatori medici, uffici postali, agenzie di servizi, aziende agricole, parchimetri e parcheggi orari, distributori automatici di sigarette.

16 Tourism satellite Account (TSA) implementation project “Enzo Paci Papers on Measuring the Economic Significance of Tourism” vol. 3, WTO 2003, Madrid, ISBN 92-844-0642-0.

17 “Measuring domestic tourism and the use of household surveys –A working paper”, WTO Department of Statistics and Economic Measurement of tourism, August 2005 (updated on July 2010).

18 Se nell'ambito territoriale del Comune di residenza della famiglia o fuori dai confini comunali.



di rilevazione nell'ambito della domanda turistica (anche se sostenute dal datore di lavoro)¹⁹.

In definitiva, il problema da affrontare è come ricondurre delle spese per l'acquisto di singoli prodotti in spese da riferire all'evento "viaggio", secondo le definizioni in uso per il turismo. In considerazione di tali limiti, pur riconoscendo la necessità di quantificare l'impatto economico del turismo, soprattutto relativamente alle spese per turismo *domestico*²⁰, l'Ilo prefigurò la possibilità di estendere lo scopo delle HBS per rilevare il luogo di acquisto per le spese effettuate in viaggio²¹. Conseguentemente, Unwto propose di includere un "tourism module" nelle indagini sulle spese/consumi delle famiglie. La scelta italiana di applicare quanto suggerito dall'Unwto, realizzando l'innesto dell'intero questionario utilizzato per la rilevazione della domanda turistica nella nuova Indagine sulle spese, consente di superare i suddetti limiti e di analizzare le spese delle famiglie sia nello svolgimento della loro vita quotidiana, sia durante l'evento eccezionale "viaggio", aprendo nuove opportunità di ricerca sul rapporto tra spese destinate al turismo e altre spese familiari²².

Si è dunque testata l'opportunità di inserire l'approfondimento sui viaggi e sulle escursioni nell'intervista iniziale e nell'intervista finale. Durante la visita iniziale vengono rilevate alcune informazioni sui viaggi effettuati negli ultimi 12 mesi, mentre con l'intervista finale si raccolgono informazioni sui viaggi con pernottamento e sulle escursioni (visite senza pernottamento) effettuati da ogni individuo della famiglia e conclusi nel mese di compilazione del diario. Per soddisfare i bisogni informativi nazionali ed internazionali²³, vengono raccolte tutte le caratteristiche qualitative dei viaggi (durata, motivo, destinazione, mezzo di trasporto, organizzazione del viaggio, partecipanti, tipo di alloggio, spese per trasporto, per alloggio e altre spese sostenute), mentre un set più limitato di informazioni viene rilevato con riferimento alle escursioni.

Le informazioni aggiuntive richieste dal focus su viaggi ed escursioni comportano un aumento del *response burden* in fase di intervista finale contenuto (attorno ai 10 minuti). Ciononostante, il loro inserimento alla fine di un'intervista abbastanza corposa, ha richiesto una lunga fase di progettazione, volta a rendere la sezione viaggi più snella e coerente con le precedenti sezioni del questionario. Durante la sperimentazione del 2011 e in tutte le edizioni di indagine realizzate successivamente, si è potuto constatare, che le famiglie non hanno avuto difficoltà nel sostenere l'intera intervista finale.

19 Secondo le Irts, i viaggi per motivi professionali vanno inclusi nella definizione di spostamento turistico.

20 Per turismo domestico si intende quello effettuato sul territorio nazionale dai residenti in quello specifico territorio.

21 Seventeenth International Conference of Labour statisticians, Household Income and expenditure statistics, Report II, Geneva 2003.

22 Tale scelta è stata motivata anche dalla necessità di una reingegnerizzazione della precedente indagine CATI sulla domanda turistica "Viaggi e vacanze", il cui campione, estratto dalle liste di abbonati alla telefonia fissa non riservata, aveva perso progressivamente rappresentatività. (Di Torrice et alii, 2014, op. cit.).

23 Le statistiche sul turismo sono regolamentate dal Regolamento 692/2011 del Parlamento europeo e del Consiglio.

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:192:0017:0032:EN:PDF>.

1.3.5 La componente longitudinale

L'introduzione di una componente longitudinale nella rilevazione della spesa per consumi comporta indiscutibili vantaggi sia in termini di robustezza delle stime, sia in termini di possibilità di analisi; consente inoltre valutazioni circa la dinamica della propensione al consumo delle famiglie, agevolando l'analisi di impatto di politiche economiche ed eventi congiunturali (U.S. Bureau of the Census, 1965; National Research Council, 2013; Parker, Souleles, Carroll 2014). Una parte dell'errore non campionario può essere, inoltre, eliminata attraverso controlli di compatibilità tra le risposte fornite nelle diverse occasioni di indagine, così come viene facilitata la collocazione nel tempo degli eventi, in particolare quando il riferimento temporale è pari all'intervallo di tempo previsto tra una *wave* e la successiva. Ripetere la misura sulle stesse unità consente anche di ridurre notevolmente i tempi di somministrazione dell'intervista in quanto alcune sezioni del questionario potrebbero essere somministrate utilizzando modalità "a conferma".

Le indagini longitudinali tuttavia presentano problematiche metodologiche aggiuntive rispetto alle trasversali e richiedono un notevole sforzo sia in fase di predisposizione e conduzione della rilevazione, sia per la costruzione delle stime (Lynn, 2009).

Il maggiore costo di gestione è dovuto soprattutto alla dispersione sul territorio delle unità nel tempo. In caso di intervista diretta, l'inseguimento dei rispondenti neutralizza i vantaggi economici derivanti da un disegno campionario stratificato. In generale dunque n replicazioni in una indagine longitudinale risultano più costose di n indagini *cross-section* (Lynn, 2009).

Un ulteriore aspetto da considerare e tenere sotto controllo è certamente il *panel conditioning*: la conoscenza di quanto dichiarato nelle precedenti *wave* può condizionare rispondente e/o rilevatore e introdurre distorsione nel dato rilevato (Muircheartaigh, 1996). Il rispondente inoltre apprende, col tempo, i meccanismi sottostanti al questionario e può mettere in atto strategie per il contenimento dei tempi di intervista o per evitare quesiti non graditi.

Una delle maggiori criticità di un disegno *panel* rimane tuttavia l'erosione del campione (*panel attrition*) che, in genere, si accompagna con una crescente distorsione campionaria difficile da correggere con i sistemi di ponderazione. Il fenomeno della non risposta nelle occasioni di indagine successive alla prima generalmente è più alto e aumenta all'aumentare dell'impegno richiesto al rispondente, in termini di tempo o di concentrazione. Per l'Italia, l'analisi di fattibilità (Rinaldelli, 2005; Rinaldelli, Marrone, 2005), ha evidenziato come, data la notevole pesantezza della rilevazione e la necessità di diffondere annualmente le principali stime di spesa, il disegno ruotato più idoneo possa essere uno schema *1,3,1* (la famiglia partecipa in un trimestre di rilevazione, non partecipa per i successivi tre e poi partecipa nuovamente per un trimestre). Tale schema mantiene, infatti, inalterata la varianza di livello delle stime e riduce quella delle variazioni tra un anno e il successivo.

Nel 2009 è stato quindi sperimentato il ritorno sulle famiglie che avevano partecipato alla sperimentazione del 2008. È stato possibile reintervistare solo il 70% del campione intervistato nel 2008, con distorsioni evidenti sia in termini territoriali (la caduta è stata più accentuata nel Nord, 35%, rispetto al Mezzogiorno, 20%), sia rispetto alla struttura per età (maggiore caduta nelle fasce più anziane e tra le persone di età compresa tra i 40 e i 50 anni).

Inoltre, solo il 33% dei rispondenti all'indagine PAPI sui consumi delle famiglie dichiara di essere disposto a partecipare nuovamente all'indagine e solo se a distanza di almeno un anno.

In definitiva sembra dunque che una rilevazione sulla spesa per consumi risulti molto pesante per i rispondenti, mostrando segnali di *panel attrition* già alla prima *wave*. Non è apparso pertanto opportuno inserire una componente *panel* nel disegno della nuova indagine.

1.3.6 L'uso di filtri

L'uso dei filtri in un questionario consente di differenziare l'intervista rispetto alla caratteristiche dei rispondenti, individuando percorsi diversi attraverso il salto di sezioni o, viceversa, attraverso il passaggio per specifiche sezioni di approfondimento.

Nella rilevazione della spesa per consumi è inoltre importante distinguere se, per le singole voci, l'assenza di spesa sia dovuta a una mancata risposta o piuttosto al fatto che la famiglia non abbia effettivamente acquistato quel bene o servizio. Lo strumento di rilevazione deve dunque essere capace di distinguere, quanto più possibile, le due occorrenze. Il controllo del fenomeno della mancata risposta parziale deve essere un obiettivo fondamentale nella fase di progettazione di un'indagine.

1.3.6.1 I percorsi filtrati

La rilevazione della spesa, quella che non avviene tramite diario, si basa su una sequenza molto lunga e ripetitiva di quesiti. L'opportuno uso di filtri – in particolare per le spese sostenute da un numero limitato di famiglie – può ridurre notevolmente il tempo di somministrazione, ma, in una indagine come quella sulle spese, può comportare importanti e indesiderati effetti di sottoregistrazione. Difficilmente, infatti, è possibile individuare gruppi di spese che, in base a caratteristiche socio-demografiche della famiglia, certamente non vengono sostenute, così come è noto, se ne è già discusso, che l'elencazione delle singole voci di spesa aiuta il rispondente a ricordare spese che altrimenti potrebbero essere dimenticate. L'uso di filtri per gruppi di spese può quindi fortemente inficiare il ricordo della spesa, con conseguente abbattimento delle frequenze (Crossley, Winter, 2013).

Di seguito si riporta l'evidenza empirica di tale affermazione. Nell'indagine sui consumi delle famiglie del 2013 l'8,5% delle famiglie ha sostenuto spese per istruzione nel mese precedente l'intervista; la percentuale è più elevata, ovviamente, tra le famiglie con figli: il 16,8% contro l'1,6% delle famiglie senza figli.

L'ampliamento delle voci di spesa e dei periodi di riferimento (da 1 a 12 mesi) introdotti con l'indagine DIR del 2006 hanno indotto un sostanziale aumento della frequenza di spesa: la quota di famiglie che effettua spese per istruzione negli ultimi 12 mesi è pari al 29,2% e corrisponde ad un 14,6% di famiglie che effettua la spesa nell'ultimo mese (nell'indagine sui consumi nel 2006 è risultato pari al 14%).

Sembrava dunque possibile/opportuno inserire un quesito-filtro che consentisse al rispondente che non aveva effettuato spese di saltare un'ampia batteria di domande (nel caso specifico costituita da 37 voci di spesa).

Poiché la classificazione Coicop impone di distinguere le spese per istruzione in base al corso di studi cui si riferiscono (ISCED 2013) si è testata la possibilità di utilizzare un filtro per evitare le sezioni riferite a specifici livelli di istruzione e un ulteriore

filtro per limitare alle sole famiglie che dichiarano di avere sostenuto la spesa i quesiti relativi all'ammontare e al mese di effettuazione della stessa.

La sperimentazione del 2008 ha previsto l'inserimento di tre quesiti-filtro (Prospetto 1.5), per portare il rispondente alla corretta sezione e poi un filtro per ogni voce di spesa per la rilevazione del mese di effettuazione e dell'importo di spesa. L'adozione di tale schema ha prodotto una stima delle famiglie che hanno sostenuto spese per istruzione nell'ultimo mese pari all'8,1%, nonostante l'ampliamento del numero di voci rilevate (nell'indagine sui consumi nel 2008 è risultata pari a 10,1%).

Prospetto 1.5 – Schema di rilevazione delle spese per l'istruzione nell'Indagine pilota sui consumi delle famiglie - Ottobre-Dicembre 2008

(1)	(2)	(3)	(4)	(5)	(6)		
	<i>(Se si)</i>		<i>(Se si)</i>		<i>(Se si)</i>		
Ha sostenuto spese per istruzione?	Per quale ordine scolastico?	Scuola infanzia	Per quale voce?	Tasse o rette scuola pubblica	Quando e quanto ha speso?		
				Tasse o rette scuola privata	Quando e quanto ha speso?		
				Mensa scolastica	Quando e quanto ha speso?		
				Corsi e laboratori tenuti in orario scolastico	Quando e quanto ha speso?		
				Scuolabus	Quando e quanto ha speso?		
				Gite scolastiche	Quando e quanto ha speso?		
				Scuola primaria	Per quale voce?	Tasse o rette scuola pubblica	Quando e quanto ha speso?
						Tasse o rette scuola privata	Quando e quanto ha speso?
		Mensa scolastica	Quando e quanto ha speso?				
		Corsi e laboratori tenuti in orario scolastico	Quando e quanto ha speso?				
		Scuolabus	Quando e quanto ha speso?				
		Gite scolastiche	Quando e quanto ha speso?				
		Libri di testo	Quando e quanto ha speso?				
		Lezioni private	Quando e quanto ha speso?				
		Scuola secondaria di primo grado	Per quale voce?	Tasse o rette scuola pubblica	Quando e quanto ha speso?		
				Tasse o rette scuola privata	Quando e quanto ha speso?		
				Mensa scolastica	Quando e quanto ha speso?		
				Corsi e laboratori tenuti in orario scolastico	Quando e quanto ha speso?		
				Scuolabus	Quando e quanto ha speso?		
				Gite scolastiche	Quando e quanto ha speso?		
Libri di testo	Quando e quanto ha speso?						
Lezioni private	Quando e quanto ha speso?						
Scuola secondaria di secondo grado	Per quale voce?	Tasse o rette scuola pubblica	Quando e quanto ha speso?				
		Tasse o rette scuola privata	Quando e quanto ha speso?				
		Mensa scolastica	Quando e quanto ha speso?				
		Corsi e laboratori tenuti in orario scolastico	Quando e quanto ha speso?				
		Scuolabus	Quando e quanto ha speso?				
		Gite scolastiche	Quando e quanto ha speso?				
		Libri di testo	Quando e quanto ha speso?				
		Lezioni private	Quando e quanto ha speso?				
Formazione professionale	Per quale voce?		Tasse o rette	Quando e quanto ha speso?			
			Libri di testo	Quando e quanto ha speso?			
Istruzione universitaria e post-universitaria	Per quale voce?		Tasse o rette università e post-università pubblica	Quando e quanto ha speso?			
			Tasse o rette università e post-università privata	Quando e quanto ha speso?			
			Mensa universitaria	Quando e quanto ha speso?			
			Libri di testo	Quando e quanto ha speso?			
			Lezioni private	Quando e quanto ha speso?			

Appare dunque evidente come un uso eccessivo di quesiti filtro può indurre un innalzamento della non risposta con evidenti ripercussioni sui valori di spesa. Nonostante dunque la ripetitività della rilevazione l'uso di sezioni filtrate per la rilevazione della spesa deve essere necessariamente limitato.

1.3.6.2 Il controllo degli zeri di spesa

Come già accennato, i rispondenti, pur ricordando un evento di spesa, possono averne dimenticato l'entità, la corretta collocazione nel tempo o il luogo in cui l'hanno effettuata e ciò è tanto più vero quanto più ampio è il periodo di riferimento. Nell'ottica di un generale ampliamento dei periodi di riferimento, è dunque opportuno prevedere, attraverso l'inserimento di opportuni quesiti filtro per ogni voce di spesa, che venga rilevato il numero totale di eventi spesa, anche in assenza di informazioni sull'importo e sul mese di effettuazione, così da avere un controllo sulla mancata risposta parziale.

Con la sperimentazione del 2008 è stato dunque inserito un filtro iniziale per ogni voce di spesa. Le famiglie e i rilevatori hanno gestito lo strumento molto bene: le frequenze di acquisto sono del tutto confrontabili e in assenza di filtro risultano addirittura più contenute. Non sembra dunque che la presenza del filtro induca le famiglie o i rilevatori a non dichiarare spese al fine contenere artificialmente i tempi di intervista (Kreuter et al. 2012), almeno non più di quanto succeda in assenza di filtro (chiedendo cioè direttamente l'ammontare della spesa sostenuta).

Ne risulta che, in media, sul complesso delle voci, circa il 5% delle famiglie pur dichiarando di aver effettuato una spesa non ne fornisce l'ammontare. Tali famiglie possono dunque essere distinte da quelle che invece sostengono a priori di non aver sostenuto la spesa.

In conclusione, dunque, l'ipotesi di inserire un filtro di controllo per ogni voce di spesa è risultata assolutamente auspicabile.

1.3.7 La tecnica di rilevazione mista CAPI/CADI

L'indagine sui consumi delle famiglie in Italia si è sempre avvalsa della tradizionale rete di rilevazione comunale in cui i rilevatori vengono selezionati direttamente dagli uffici di statistica dei comuni-campione che hanno anche il compito di supervisionare le attività su campo. L'indagine è stata sempre condotta con tecnica PAPI (per motivi legati sia alle risorse sia all'organizzazione della rete di rilevazione) nonostante l'uso di tecniche *Computer Assisted*, oltre a garantire noti vantaggi comuni a tutte le rilevazioni, risponda anche a esigenze più specifiche dell'Indagine sulle spese, quali i) guidare il rilevatore e la famiglia attraverso ciò che deve essere incluso/escluso dalle spese familiari, ii) controllare alla fonte l'inserimento di valori di spesa implausibili, iii) agevolare la collocazione temporale delle spese soprattutto nel caso di periodi di riferimento mobili, iv) verificare alla fonte gli zeri di spesa in base a quesiti filtro e regole di compatibilità, iv) supportare il controllo della compilazione del diario da parte del rilevatore attraverso un *software* specifico; v) supportare il *data entry* del diario che può dunque essere demandato al rilevatore stesso con ovvi vantaggi in termini di qualità e tempestività (Barcherini, Falorni, Grassi, 2014).

1. Il processo di ristrutturazione dell'indagine

Nel 2011 i questionari sono diventati elettronici ed è stata avviata una sperimentazione che ha avuto la doppia funzione di testare le tecniche *Computer Assisted* e il passaggio da una rete comunale a una professionale.

Nonostante la sperimentazione abbia messo in luce alcune difficoltà nella gestione della rete di rilevazione (il tasso di risposta ha superato appena il 50%) e abbia evidenziato la necessità di modificare l'implementazione informatica di alcune sezioni, l'uso dei questionari elettronici ha sicuramente soddisfatto le aspettative: le stime trimestrali, grazie all'eliminazione delle fasi di revisione manuale di tutti i questionari cartacei e di acquisizione centralizzata dei dati su supporto informatico, possono essere prodotte con un anticipo di 15 giorni (il gap temporale tra rilevazione e stima è passato da 60 a 45 giorni) e si basano su dati grezzi più corretti.

L'utilizzo di tali tecniche, dunque, rappresenta una opportunità estremamente importante per la rilevazione della spesa.

1.3.8 Conclusioni

Con il sistema delle sperimentazioni, sono stati testati e messi a punto gli obiettivi, le metodologie e gli strumenti di una rinnovata Indagine sulle spese delle famiglie. Nel Prospetto 1.6 si riassumono brevemente i risultati ottenuti rispetto agli obiettivi della ristrutturazione inizialmente enunciati.

Prospetto 1.6 – Sintesi dei principali risultati ottenuti con le sperimentazioni

OBIETTIVO	PRINCIPALI RISULTATI
Strategia di campionamento e di costruzione delle stime	<ul style="list-style-type: none"> – È possibile introdurre un nuovo dominio di stratificazione (la tipologia comunale) al fine di tenere sotto controllo l'impatto sulle stime dei diversi livelli e strutture di consumo in contesti territoriali differenti; – È possibile definire nuovi vincoli che consentono di ponderare rispetto a specifici sottogruppi di popolazione (ad esempio stranieri).
Classificazione delle voci di spesa alla più recente COI-COP	La classificazione proposta da Eurostat (2013) può essere utilizzata, fatte salve alcune voci che a livello nazionale presentano difficoltà di rilevazione (test di gravidanza e contraccettivi di barriera, trasporto per vie d'acqua interne e marittime, servizi pubblici di assicurazione connessi con la salute, birre lager, altre birre alcoliche, birre a basso contenuto alcolico o non alcoliche, altre bevande a base di birra).
Periodi di rilevazione diario	Il rapporto costi/benefici sembra favorevole a una compilazione a 14 giorni: l'ampliamento non sembra inficiare in modo significativo la qualità dei dati ma sembra cogliere meglio i comportamenti di spesa delle famiglie.
Periodi di riferimento delle spese rilevate in modo retrospettivo	È possibile ampliare i periodi di riferimento dei quesiti retrospettivi ed associarvi un calendario delle spese, valutando opportunamente l'ampiezza ottimale in base a frequenza, variabilità ed entità delle spese, al fine di rendere più robuste le stime accrescendo il numero di eventi osservati.
Rilevazione dell'evento spesa	È possibile inserire quesiti filtro per ogni voce di spesa al fine di rilevare l'evento spesa anche in assenza di informazioni sulla spesa.
Introduzione di una componente panel	L'indagine sulle spese determina un forte <i>burden</i> sui rispondenti; un ritorno, anche a distanza di un anno, sulle stesse famiglie determina un evidente effetto di <i>panel attrition</i> .
Rilevazione delle quantità alimentari acquistate	La rilevazione delle quantità alimentari acquistate, allo scopo di ottenere informazioni utili sui comportamenti alimentari delle famiglie, non può essere associata a quella della spesa e degli autoconsumi se non inficiando fortemente la rilevazione.
Rilevazione del luogo prevalente di acquisto	È possibile prevedendo schemi riepilogativi per beni e servizi acquistati frequentemente (e rilevati con diario) rilevare il luogo prevalente di acquisto.
Rilevazione delle <i>core variables</i>	Se associata all'impiego dei nuovi strumenti di rilevazione, che permettono di gestire meglio i percorsi del questionario alleggerendo anche il peso sui rispondenti, la rilevazione di un set di variabili, utilizzabili anche per <i>matching statistics</i> con altre indagini, è possibile e anche raccomandabile.
Innovazione delle tecniche di cattura, di controllo e di correzione del dato	L'impiego di nuove tecniche riduce gli interventi a posteriori e aumenta la tempestività delle stime.

2. IL DISEGNO DELLA NUOVA INDAGINE¹

I risultati delle sperimentazioni hanno consentito di ridisegnare l'Indagine sulle spese per consumi, tenendo conto anche delle necessità informative emerse a partire dalla precedente ristrutturazione e delle esigenze di armonizzazione sia alle altre indagini sulle famiglie (in particolare l'indagine su Viaggi e Vacanze e l'indagine EU-Silc su reddito e condizioni di vita) sia agli standard messi a punto a livello europeo (con particolare riferimento alla classificazione Coicop e alla definizione delle *core variables*).

Sono state riviste tutte le fasi del processo di produzione del dato: la rilevazione – in termini di disegno di campionamento, strumenti, tecnica e rete di rilevazione – la correzione, il sistema di ponderazione, il calcolo delle stime e la diffusione.

La nuova indagine CAPI sulle spese pur mantenendo, come in precedenza, un disegno campionario a due stadi con stratificazione delle unità di primo stadio, ha incluso tra le variabili di stratificazione la tipologia comunale; è continua nel corso dell'anno con la stessa numerosità campionaria, fissata a 28.000 famiglie: il campione annuale viene suddiviso in dodici sottocampioni mensili di circa 2.300 famiglie e, nell'ambito di ciascun campione mensile, le famiglie vengono divise in due gruppi in modo da coprire tutti i giorni dell'anno. Rispetto alla vecchia indagine, a seguito dell'ampliamento del periodo di compilazione del diario (da 7 a 14 giorni), il primo gruppo compila il diario nella prima metà del mese di riferimento (periodo di rilevazione 1), il secondo gruppo nella seconda metà (periodo di rilevazione 2); per ogni famiglia, il mese di riferimento e il periodo di rilevazione vengono stabiliti attraverso il disegno campionario e la dimensione temporale può essere dunque considerata come un'ulteriore variabile di stratificazione del campione (cfr. par. 2.2.1).

La gestione dei periodi di rilevazione e dei contatti con le famiglie viene supportata da due strumenti di rilevazione: l'agenda elettronica e la scheda contatti. In tal modo il rilevatore individua le famiglie eleggibili e il corretto rispondente, gestisce i contatti ed esegue le azioni da svolgere in caso di caduta della famiglia campione (cfr. par. 2.3.2.1). Nella vecchia indagine, insieme all'elenco delle famiglie campione (elenco base) veniva estratto un elenco di famiglie di riserva (di pari numerosità) da utilizzare in caso di caduta della famiglia dell'elenco base (scegliendo quella con caratteristiche più simili alla famiglia base caduta). Nella nuova indagine, invece, per ogni famiglia campione (appartenente all'elenco base) vengono selezionate, sempre dalle liste anagrafiche e in maniera casuale, tre famiglie sostitutive che presentano caratteristiche simili alla famiglia dell'elenco base, in termini di zona di residenza

¹ A cura di D. Grassi per la parte introduttiva e parr. 2.5.1 e 2.6; M.C. Romano per il par. 2.1; C. De Vitiis e M. Russo par. 2.2.1; C. Ceccarelli par. 2.2.2; I. Arigoni per i parr. 2.3.1, par. 2.3.2.1 per l'introduzione e la parte "L'adattamento alle specificità dell'indagine sulle spese", par. 2.3.3.3 per la parte "Gli indicatori di monitoraggio indiretti", V. Triolo par. 2.3.2.1 per la parte "L'architettura classica"; S. Barcherini parr. 2.3.2.2, parte introduttiva del par. 2.4, parr. 2.4.1 e 2.4.2.5; P. Scalisi parr. 2.3.3.1, par. 2.3.3.2 per la parte "I debriefing", par. 2.3.3.3 per la parte "I controlli telefonici"; N. Bali par. 2.3.3.2 per la parte "Il modulo sulla comunicazione e il contenimento dei rifiuti"; P. Marrone par. 2.3.3.3 per la parte "L'ottica delle attività di monitoraggio della rilevazione", M. Morricone parr. 2.3.3.2 per la parte "Attività e strumenti per la formazione" e par. 2.3.3.3 per la parte "Gli indicatori di monitoraggio diretti", C. Falorni parr. 2.4.2.1 e 2.4.2.2; V. de Martino par. 2.4.2.3; L. Rondini par.2.4.2.4; A. Cutillo par. 2.5.2; A. Masi par. 2.5.3.

(stessa via) e numero di componenti; le famiglie sostitutive vengono cioè associate ad ogni singola famiglia campione, non lasciando al rilevatore alcun margine di discrezionalità. Tutte le famiglie dell'elenco base e la prima famiglia delle tre sostitutive vengono preavvertite del loro imminente coinvolgimento nella rilevazione tramite lettera a firma del Presidente dell'Istat, che illustra i principali obiettivi dell'indagine, riporta gli articoli di legge che regolano l'obbligo di risposta e il rispetto della privacy, comunica il numero verde messo a disposizione delle famiglie per qualsiasi richiesta di chiarimento.

La tecnica di rilevazione è rimasta mista, passando però da PAPI a CAPI per quanto riguarda le interviste dirette: il diario cartaceo compilato dalle famiglie è stato ampliato da 7 a 14 giorni, le interviste sono diventate due (parte dell'intervista finale è diventata oggetto di un'intervista iniziale al fine di alleggerirla) ed è stata aggiunta una visita di controllo (supportata anch'essa da un questionario elettronico) a metà del periodo di compilazione del diario (cfr. par. 2.3.1).

L'intervista iniziale rileva alcune informazioni invariante nel breve termine, come le caratteristiche socio-demografiche dei componenti la famiglia, le caratteristiche dell'abitazione principale, il possesso di alcuni beni durevoli e alcune caratteristiche dei viaggi effettuati. Con il diario cartaceo vengono rilevate le spese correnti e le quantità dei beni autoconsumati, con l'aggiunta della rilevazione dei luoghi di acquisto. L'intervista finale rileva le spese meno frequenti o eccezionali, inclusi i luoghi di acquisto, completando l'approfondimento sulle caratteristiche e sulle spese di viaggio effettuati nell'ultimo mese (cfr. par. 2.3.2.2).

La rilevazione in modalità CAPI è stata possibile affidando la gestione delle interviste a una rete di rilevazione professionale, con conseguente abbandono della rete di rilevazione comunale utilizzata in precedenza. Ciò ha comportato una rivisitazione delle fasi relative alla formazione (cfr. par. 2.3.3.2) e al monitoraggio dell'attività dei rilevatori (cfr. par. 2.3.3.3).

Le possibilità offerte da un sistema *Computer Assisted* di controllare la congruità e la correttezza delle risposte fornite direttamente in fase di rilevazione hanno permesso di alleggerire e comprimere temporalmente anche la fase di controllo e correzione dei dati (cfr. par. 2.4).

L'ampliamento dei contenuti informativi ha consentito infine di migliorare la fase di produzione delle stime, sia per quanto concerne il sistema di ponderazione (cfr. par. 2.5.2) – con l'inserimento di un più ampio numero di vincoli e la revisione dei metodi di trattamento degli *outliers* – sia per quanto riguarda le metodologie di stima, con particolare riferimento al metodo di valorizzazione monetaria degli autoconsumi (cfr. par. 2.4.2.4) e al periodo di riferimento per la diffusione delle stime (cfr. par. 2.5.1).

2.1 La transizione a tecniche di rilevazione CA nel contesto della modernizzazione delle indagini campionarie sulle famiglie

Nell'ambito del Programma pluriennale Stat2015 (disegnato dall'Istat in linea con le raccomandazioni europee e che prevede la reingegnerizzazione, la standardizzazione e l'industrializzazione dei processi di produzione in un'ottica di sistema), a partire dal 2010 è stato avviato un ampio processo di modernizzazione delle indagini campionarie su famiglie e individui. Un asse portante di questo processo è rappre-

sentato dalla estensione delle tecniche *computer-assisted* all'intero sistema di indagini campionarie sulle famiglie, in risposta anche alle sollecitazioni di Eurostat che, ormai da molti anni, incoraggia i paesi all'utilizzo di tali tecniche per la rilevazione dei dati.

Nel percorso di modernizzazione del *data collection*, che nell'ultimo quinquennio ha subito una forte accelerazione, è stato possibile sfruttare l'ampia esperienza maturata negli anni Novanta, sia nella realizzazione di indagini CATI (in particolare nelle indagini Multiscopo sulla Sicurezza dei cittadini e su Viaggi e vacanze realizzate a partire dal 1997), sia nell'ambito dell'Indagine sulle forze di lavoro, condotta, a partire dal 2003, con tecnica mista CAPI/CATI.

La modernizzazione delle tecniche di rilevazione ha significato innanzitutto promuovere il ricorso alla tecnica CAPI, in sostituzione della tecnica PAPI, sia per indagini alla loro prima edizione, sia per indagini consolidate, come nel caso dell'Indagine sulle spese delle famiglie, per le quali è stato pertanto necessario un ridisegno dell'intero processo di indagine, in coerenza con la nuova tecnica adottata. Ove reso possibile dal target di rispondenti e dalla metodologia di indagine, è stato favorito il ricorso ad una doppia tecnica *computer-assisted*, adottando mix che consentissero di mantenere alti i livelli di partecipazione alle indagini e, al contempo, di contenere i costi².

Per la realizzazione di queste attività, a giugno 2010, nell'ambito della Direzione centrale per le indagini su condizioni e qualità della vita, è stata costituita una Task Force avente il compito di predisporre tutte le operazioni necessarie all'effettuazione della transizione delle indagini sulle famiglie a tecnica CAPI, inclusa l'Indagine sulle spese delle famiglie. Contestualmente sono state indette due procedure di gara internazionali per l'affidamento a società private delle attività di rilevazione. Il primo bando di gara ha previsto, tra le altre attività, la realizzazione nel 2011 dell'Indagine pilota sulle spese delle famiglie da effettuarsi con sistema CAPI e autocompilazione del diario cartaceo (1.000 interviste). Il secondo bando di gara ha previsto la realizzazione, sempre con sistema CAPI e autocompilazione del diario cartaceo, dell'Indagine sulle spese delle famiglie, per un totale di circa 98 mila interviste nel quadriennio dal 2012 al 2015 (circa 28 mila l'anno).

Il primo passo è stato, quindi, la predisposizione della documentazione tecnica e amministrativa necessaria ad espletare le procedure di gara, all'insegna dell'armonizzazione sia formale che contenutistica con le altre indagini dell'Istituto condotte in modalità CA. Non a caso si è scelto di includere in un medesimo bando di gara l'espletamento di servizi relativi a più indagini; ciò consentiva, oltre che di interfacciarsi con una stessa impresa, di adottare strategie comuni e condividere le scelte effettuate in ciascuna specifica indagine, armonizzando strumenti e metodologie e limitando le differenze tra i processi alle specifiche esigenze legate alle peculiarità delle indagini. In particolare, l'armonizzazione delle diverse fasi di preparazione e supervisione del lavoro sul campo, ha comportato una standardizzazione dei moduli formativi destinati ai rilevatori e agli operatori delle società private coinvolti nella gestione del *fieldwork*, che ha aiutato anche gli intervistatori ad orientarsi tra le varie indagini e coglierne meglio le specificità. Analogo discorso ha riguardato il sistema di indicatori di monitoraggio, strutturato in modo da prevedere un set di informazioni standard comuni a

2 Per una rassegna delle principali esperienze, si veda Freguja C. e Romano M.C., 2014, op. cit..



tutte le indagini e, al contempo, dotato della flessibilità necessaria ad accogliere le specifiche esigenze di controllo e supervisione dettate dalle diverse metodologie di indagine. Gli stessi questionari elettronici sono stati pensati, nelle varie loro componenti e soprattutto con riferimento ai moduli comuni alle diverse indagini (gestione dei contatti, *core variables*, etc.), nell'ottica della massima standardizzazione.

L'Indagine sulle spese ha sicuramente beneficiato dell'impostazione armonizzata dell'intero processo, d'altro canto, proprio per la sua complessità metodologica, ha spesso posto nuove sfide e, di conseguenza, imposto la ricerca di soluzioni innovative che hanno contribuito ad accrescere l'efficienza dell'intero processo produttivo, oltre che essere a loro volta fruibili da altre indagini.

In sintesi, la transizione a nuova tecnica oltre a sviluppare nuovi *skill*, essendo stata condivisa con altre indagini del sistema di indagini campionarie sulle famiglie, ha consentito l'interscambio di vecchi e nuovi *know how* derivanti dalle singole esperienze di indagine, con significativi e proficui riflessi sia sulla qualità dell'intero processo di produzione, sia sulla crescita professionale delle risorse umane in esso coinvolte.

2.2 Il disegno di campionamento e la selezione del campione

2.2.1 Il disegno di campionamento

Le innovazioni apportate in termini di campionamento hanno riguardato l'introduzione della tipologia comunale come variabile di stratificazione delle unità di primo stadio e la modifica dello schema di assegnazione casuale del periodo di compilazione del diario a seguito del suo ampliamento da 7 a 14 giorni.

2.2.1.1 Schema generale del disegno di campionamento

Una delle tecniche di indagine utilizzate, come in passato, è l'intervista diretta; pertanto, per ragioni legate sia ai costi di indagine sia a problematiche di tipo organizzativo, il campione deve essere concentrato in un limitato numero di comuni; ne deriva che, come in passato, il tipo di disegno campionario ottimale rimane quello a due stadi di selezione, in cui le unità di primo stadio sono i comuni e le unità di secondo stadio, o unità finali di campionamento, sono le famiglie.

L'indagine, come già specificato in precedenza, deve fornire stime trimestrali a livello nazionale e stime annuali a livello regionale (distinguendo le province di Bolzano e Trento).

I comuni sono stratificati, all'interno di ciascuna regione, per tipologia comunale e dimensione demografica. La tipologia comunale è stata introdotta sia in un'ottica di armonizzazione con quanto fatto per altre indagini sulle famiglie, sia per garantire robustezza alle stime, in quanto i livelli e le abitudini di spesa sono fortemente legati all'ampiezza del comune in cui si vive (non a caso anche la stima della povertà assoluta tiene conto della tipologia comunale di residenza). In particolare, si distingue tra:

- area metropolitana (comuni con oltre 250 mila abitanti);
- grandi comuni (comuni della periferia dell'area metropolitana e comuni con 50 mila abitanti e più);
- piccoli comuni (comuni con meno di 50 mila abitanti).

La seconda variabile di stratificazione – espressa in termini di popolazione residente – viene utilizzata per suddividere i comuni, all'interno di ogni incrocio definito da regione e tipologia comunale, tra: (i) *Auto Rappresentativi* (AR), comuni con popolazione al di sopra di una soglia opportunamente definita (cfr. par 2.2.1.2.), ognuno dei quali rappresenta uno strato; (ii) *Non Auto Rappresentativi* (NAR), comuni con popolazione al di sotto della suddetta soglia. Per le famiglie appartenenti ai comuni AR si adotta, pertanto, un *disegno ad uno stadio stratificato*, mentre per l'insieme dei comuni NAR si usa un *disegno a due stadi con stratificazione delle unità di primo stadio*. Da ogni comune campione, sia esso AR o NAR, si estrae un numero predeterminato di famiglie anagrafiche. È utile precisare che, mentre per l'indagine sulle spese, la famiglia rappresenta l'unità finale di campionamento e anche l'unità di analisi, per il modulo di approfondimento sulle caratteristiche dei viaggi e delle escursioni, poiché le variabili di interesse sono rilevate a livello individuale, la famiglia costituisce un grappolo di individui.

Nella vecchia indagine le specifiche esigenze relative alla produzione degli indici sui prezzi al consumo avevano richiesto che tutti i capoluoghi di provincia fossero AR; poiché tale vincolo determinava un allontanamento dalla condizione di autoponderazione del campione, venute meno le esigenze del servizio prezzi, si è ritenuto opportuno eliminarlo.

La dimensione campionaria dell'indagine non è stata oggetto di revisione; pertanto, come nella precedente, in ogni trimestre si intervistano circa 7.000 famiglie, per un totale di 28.000 famiglie all'anno.

Il disegno di campionamento, definito con riferimento a un generico trimestre dell'anno, viene replicato in modo identico per i quattro trimestri e viene anche effettuata una stratificazione mensile del campione trimestrale. Di conseguenza, la dimensione temporale può essere considerata un'ulteriore variabile di stratificazione del campione.

La partecipazione dei comuni campione nell'arco dell'anno è la seguente:

- ogni comune AR è coinvolto nell'indagine tutti i mesi (e il numero di famiglie campione annuale è suddiviso nei 12 mesi);
- ciascuno dei tre comuni campione di ogni strato NAR partecipa all'indagine quattro mesi all'anno a distanza di tre mesi (Prospetto 2.1).

Prospetto 2.1 – Stratificazione mensile dei comuni NAR

Comune	Mese di partecipazione			
1	gennaio	aprile	luglio	ottobre
2	febbraio	maggio	agosto	novembre
3	marzo	giugno	settembre	dicembre

2.2.1.2 Stratificazione e selezione del campione

L'obiettivo della stratificazione è quello di formare gruppi (o strati) di unità caratterizzate, relativamente alle variabili oggetto d'indagine, da massima omogeneità interna e massima eterogeneità fra i gruppi. Il raggiungimento di tale obiettivo si traduce, in termini statistici, in un guadagno nella precisione delle stime, ossia in una riduzione dell'errore campionario a parità di numerosità campionaria.



Nell'Indagine sulle spese per consumi, i comuni vengono stratificati in base alla loro dimensione demografica e nel rispetto delle seguenti condizioni:

- autoponderazione del campione al livello di regione;
- definizione di un numero minimo di famiglie da intervistare per comune;
- stratificazione dei comuni sulla base dell'ampiezza demografica;
- formazione di strati di comuni di ampiezza approssimativamente costante in termini di popolazione residente.

Il procedimento di stratificazione è attuato all'interno di ogni dominio territoriale individuato, per ciascuna regione geografica, dalle tre tipologie comunali considerate.

Con riferimento al generico incrocio di regione e tipologia comunale, indicato nel seguito con il termine *dominio*, si denoti con: l ($l=1, \dots, L$) l'indice di strato di comuni; c ($c=1, \dots, C$), l'indice di comune; N_l , il numero di famiglie residenti nello strato l ; N_{lc} il numero di famiglie residenti nel comune c dello strato l ; P_l , il numero di individui residenti nello strato l ; $P_{(lc)}$ il numero di individui residenti nel comune c dello strato l ; m , il numero di comuni campione in ogni strato.

Per la definizione del campione di comuni e di famiglie relativo ad un trimestre sono state dapprima effettuate le seguenti scelte:

- a. definizione del numero complessivo di famiglie campione a livello nazionale;
- b. definizione del numero n_r di famiglie campione per ciascuna regione;
- c. distribuzione delle numerosità regionali tra i domini definiti dalla tipologia comunale in misura proporzionale alla popolazione residente;
- d. scelta del numero minimo di famiglie, \bar{n}_r , da intervistare in ciascun comune campione.

Dalla scelta di \bar{n}_r e n_r dipende la suddivisione dei comuni in AR e NAR e la formazione degli strati attraverso i seguenti passi:

1. calcolo della frazione di campionamento regionale $f_r = n_r / N_r$ essendo N_r il numero di famiglie residenti nella regione r ;
2. determinazione del valore della *soglia*³ λ_r mediante la relazione

$$\lambda_r = \frac{\bar{n}_r \delta_r}{f_r} \quad (1)$$

in cui δ_r è il numero medio di componenti per famiglia a livello regionale; risulta evidente da tale espressione che la soglia per la definizione dei comuni AR cresce al crescere di \bar{n}_r ;

3. suddivisione dei comuni nei due sottoinsiemi AR e NAR sulla base della soglia λ_r ;
4. ordinamento decrescente dei comuni NAR all'interno di ogni dominio in funzione della loro dimensione demografica;
5. suddivisione dei comuni NAR in strati la cui dimensione è approssimativamente uguale al prodotto $m \times \lambda_r$;
6. selezione di m comuni campione da ciascuno strato l ($l=1, \dots, L$) con probabilità proporzionale all'ampiezza, mediante la procedura di selezione sistematica,

3 Dal momento che il campione di famiglie di ogni regione è distribuito in modo proporzionale tra i domini e, di conseguenza, la frazione di campionamento è costante all'interno della regione, la soglia risulta anch'essa definita a livello regionale, così come la suddivisione dei comuni in AR e NAR.

2. Il disegno della nuova indagine

suggerita da Madow (1949) e Murthy (1967)⁴; per il generico comune c tale probabilità è espressa dalla formula:

$$z_{lc} = m \cdot P_{lc} / P_1 \quad (2)$$

7. definizione del numero n_{lc} di famiglie da intervistare in ogni comune; dalla condizione di autoponderazione a livello regionale

$$\frac{m \cdot P_{lc}}{P_1} \frac{n_{lc}}{N_{lc}} = f_r \quad (3)$$

in cui il primo membro rappresenta la probabilità d'inclusione delle famiglie del comune c dello strato l (essendo le due frazioni rispettivamente la probabilità d'inclusione di primo e di secondo stadio), si ottiene

$$n_{lc} = \frac{f_r}{m} \frac{P_1 \cdot N_{lc}}{P_{lc}} \quad (4)$$

Al fine di raggiungere la numerosità campionaria desiderata in termini di comuni nel rispetto della condizione di autoponderazione a livello regionale, il numero minimo di interviste trimestrali per comune è stato posto pari a 14 per le regioni con popolazione superiore a 1.500.000 di abitanti e 12 per le regioni più piccole.

Relativamente alla quantità m , per consentire la stratificazione mensile del campione di comuni, si è scelta una numerosità per ogni strato NAR pari a 3. A tale proposito, è utile precisare che l'introduzione nella stratificazione della variabile relativa alla tipologia comunale ha comportato problematiche nella costruzione degli strati NAR per alcune tipologie comunali. In particolare, negli strati costituiti da meno di 3 comuni, questi sono stati definiti tutti come AR ed è stato assegnato loro una numerosità campionaria, in termini di famiglie, pari a $\hat{n}_l > n_l$ per garantire la dimensione campionaria di strato; negli strati costituiti esattamente da 3 comuni, questi sono stati definiti come NAR (Tavola 2.1).

Tavola 2.1 – Numero di comuni che partecipano all'indagine per tipo e finestra temporale di partecipazione

	Comuni campione		
	AR	NAR	TOTALI
Mese	52	150	202
Trimestre	52	450	502
Anno	52	450	502

L'estrazione delle famiglie campione viene effettuata, mediante selezione casuale, dalle liste anagrafiche comunali (cfr. par. 2.2.2); per ciascun comune campione vengono quindi selezionate (cfr. punto 7) $4n_{lc}$ famiglie, successivamente ripartite nell'anno come segue:

4 Madow, W. G. op. cit.(1949), Murthy, M. N. op. cit. (1967). Tale procedura presenta le seguenti caratteristiche: (i) assegna una probabilità di inclusione espressa nella forma $\pi_{hlc} = m z_{hlc}$; (ii) la sua implementazione è estremamente semplice; (iii) conduce all'ottenimento di stime generalmente più efficienti rispetto a quelle ottenibili con altre procedure di selezione.

- il campione di famiglie di ciascun comune AR viene suddiviso in 12 campioni mensili;
- il campione di famiglie di ciascun comune NAR viene suddiviso in 4 campioni mensili.

Sia nei comuni AR sia nei NAR, il campione mensile di famiglie viene ulteriormente ripartito in due sottoinsiemi, distinguendo le famiglie che compileranno il diario nella prima metà del mese da quelle che lo dovranno compilare nella seconda. La procedura di assegnazione dei periodi di compilazione del diario prevede che per ciascun dominio, definito dall'incrocio della regione e della tipologia comunale, l'assegnazione avvenga in maniera casuale.

2.2.2 L'utilizzo delle Liste Anagrafiche Comunali per la selezione del campione

Nell'ambito delle operazioni censuarie, il Dipartimento dei censimenti generali e archivi ha raccolto le liste anagrafiche di tutti i comuni italiani a partire dal 31 dicembre 2010.

Per ogni occasione è stato creato un archivio che contiene le informazioni relative a tutti gli individui residenti sul territorio di ciascun comune italiano, distinti per famiglia o convivenza (Liste Anagrafiche Comunali-LAC). Tali informazioni sono state raccolte tramite un'applicazione *web*, messa a disposizione dall'Istituto, che ha garantito il controllo e la qualità dei dati acquisiti, nel rispetto dei più elevati standard di sicurezza, e ha consentito un efficace supporto centralizzato e un monitoraggio in "tempo reale" dello stato delle operazioni (Ceccarelli, Pezone, Rosati, 2013).

Oltre ad essere uno degli strumenti utilizzati nell'ambito del Censimento generale della popolazione e delle abitazioni del 2011, tale archivio rappresenta un elemento cardine di innovazione, contribuendo alla standardizzazione dei processi e all'armonizzazione delle metodologie su cui si basano le indagini sociali e socio-economiche dell'Istat.

Nell'indagine PAPI sui consumi, la selezione delle famiglie da intervistare veniva demandata a ogni singolo comune campione che, in base alla metodologia indicata dall'Istat, estraeva dalla propria lista anagrafica i nominativi. Gli elenchi dei nominativi venivano poi inviati all'Istat, talvolta anche in formato cartaceo, e sottoposti a controlli di correttezza prima di essere assegnati ai rilevatori.

Dal 2011, il processo di estrazione dei campioni è svolto interamente all'interno dell'Istat, tramite l'impiego della LAC che garantisce, oltre alla standardizzazione delle procedure, una notevole tempestività nell'estrazione dei nominativi delle famiglie campione, grazie all'azzeramento dei tempi prima necessari per tutte le attività di competenza dei singoli comuni coinvolti. Tale procedura di estrazione, che agisce uniformemente su un *database* centralizzato, è stata anche in grado di attenuare in modo significativo i possibili effetti distorsivi dovuti all'inclusione o esclusione non casuale di famiglie residenti in alcune zone del territorio. Inoltre, è possibile definire, sia a priori sia in corso d'opera, strategie di sovra-campionamento, soprattutto nel caso di particolari sottogruppi di popolazione dove il fenomeno delle mancate risposte totali risulta più diffuso; si pensi ad esempio alla popolazione straniera che, a causa sia della lingua sia della più elevata mobilità sul territorio,

presenta tassi di mancata partecipazione più elevati rispetto al resto della popolazione residente.

In particolare, nel caso dell'Indagine CAPI sulle spese che prevede il ricorso alle sostituzioni per le mancate risposte, l'uso delle LAC ha consentito di rimpiazzare gli elenchi sostitutivi con le quartine, in analogia con quanto fatto per altre indagini (come già accennato, ad ogni famiglia base vengono associate 3 famiglie sostitutive con caratteristiche simili), rendendo molto più mirata la scelta della famiglia sostitutiva e standardizzandone le modalità di selezione (con la PAPI il responsabile comunale doveva selezionare la famiglia sostitutiva da un unico elenco individuandone una con caratteristiche simili a quelle della famiglia da sostituire).

La disponibilità del codice fiscale per ciascun individuo (informazione presente nelle LAC) rende inoltre possibile l'integrazione e il confronto con i dati provenienti da fonti amministrative, in particolare con la banca dati reddituale.

Infine, ad ogni estrazione, il campione di famiglie viene confrontato con i campioni già estratti in altre indagini e con quello dell'Indagine sulle spese relativo all'anno precedente. In tal modo, si evita che la famiglia partecipi a più indagini nel corso dell'anno o all'indagine per due volte consecutive, riducendo così la mancata partecipazione e la molestia statistica.

L'esperienza già maturata, tuttavia, evidenzia due criticità nell'uso delle LAC. La prima è la complessità intrinseca delle procedure e delle elaborazioni necessarie affinché l'archivio possa essere utilizzato, soprattutto per quanto riguarda l'intero processo di acquisizione e aggiornamento della LAC. La seconda, parzialmente legata alla prima, è la discrepanza temporale fra la data di acquisizione delle LAC e quella dell'estrazione: più elevata è la discrepanza, tanto maggiore è la probabilità di estrarre famiglie che hanno cambiato abitazione e che, per questo, risulteranno irreperibili.

Le LAC rappresentano anche un possibile strumento per il trattamento della mancata risposta totale, fenomeno sempre più diffuso nelle rilevazioni statistiche del quale si deve tenere conto, anche da un punto di vista teorico, a partire dalla definizione della strategia di indagine (Särndal e Lundström, 2005).

La complessità delle indagini, soprattutto in campo socio-economico, la concomitanza di differenti tecniche di rilevazione, l'utilizzo di diverse reti di rilevazione impongono risposte metodologiche articolate, che prevedono la disponibilità di informazioni anche sui non rispondenti: lo sfruttamento di fonti aggiuntive può limitare gli effetti distorsivi dei meccanismi di mancata risposta totale sulle stime finali. Con l'uso degli stimatori post-stratificati e soprattutto degli stimatori calibrati (Deville e Särndal 1992), l'impiego di informazioni ausiliarie da fonti esterne per la costruzione delle stime da indagini campionarie è ormai prassi consolidata, anche e soprattutto nel trattamento delle mancate risposte totali (Kott 2006; Ceccarelli, Cutillo 2007; Ceccarelli et al. 2008).

È noto che i passaggi che dal campione teorico portano alla determinazione delle stime finali o, meglio, alla determinazione dei coefficienti di riporto all'universo, possono essere riassunti in due macro fasi ben distinte:

- una volta raccolte le informazioni e determinato il coefficiente di riporto all'universo da disegno, si procede ad adeguare il campione realizzato a quello teorico; si tiene cioè conto della mancata risposta totale in modo da ripristinare sia la numerosità campionaria di partenza sia la struttura del campione inizialmente estratto;

- grazie all'uso di stimatori, più o meno complessi, che sfruttano informazioni ausiliarie, si arriva alla determinazione dei coefficienti di riporto all'universo finali che eliminano anche la distorsione dovuta alla sotto copertura delle liste di campionamento.

Nel contesto della prima fase si inseriscono le sperimentazioni sull'utilizzo delle LAC per la determinazione dei coefficienti correttivi della mancata risposta totale.

2.3 Il disegno della rilevazione

2.3.1 Il ciclo di rilevazione

Le 28.000 famiglie del campione annuale, come già detto, vengono suddivise in dodici sottocampioni (di circa 2.300 unità), ciascuno dei quali partecipa alla rilevazione in uno specifico mese di riferimento che coincide con il mese di calendario in cui le famiglie compilano il diario. La durata del lavoro su campo per ogni singola famiglia è tuttavia molto più ampia e si estende su tre mesi di calendario (Figura 2.1). Le famiglie coinvolte nell'indagine, infatti, oltre a compilare il diario, ricevono tre visite del rilevatore (esclusi, naturalmente, tutti i contatti, telefonici o a domicilio, necessari ad organizzarle) secondo un preciso calendario:

- la prima visita (visita iniziale) deve avvenire in un intervallo di 20 giorni che termina, tassativamente, il giorno precedente quello di inizio del periodo di compilazione del diario; in occasione di tale visita, il rilevatore somministra l'intervista iniziale e consegna alla famiglia il diario cartaceo, fornendo tutte le spiegazioni necessarie per la corretta compilazione ed evidenziando con cura le date di inizio e fine di tale operazione;
- la seconda visita (visita intermedia) è effettuata tra il quarto e l'undicesimo giorno di compilazione del diario, con l'obiettivo di verificare il corretto andamento della rilevazione, fornire spiegazioni o aiuto e, se necessario, sollecitare le famiglie "pigre", recuperando così eventuali interruzioni;
- la terza visita (visita finale) deve essere effettuata nei primi 15 giorni del mese successivo a quello di compilazione del diario; in tale occasione, il rilevatore effettua un nuovo *check* del diario, recuperando informazioni mancanti o incomplete con l'ausilio della famiglia, e somministra l'intervista finale.

Figura 2.1 – Schema di rilevazione (per il generico mese M di riferimento)

	Mese M-1		Mese M (Mese di riferimento)														Mese M+1											
	12	...	31	1	...	5	...	11	...	15	16	17	18	...	22	...	31	1	...	5	6	7	...	10	11	...	15	16
GRUPPO 1	Primo contatto e intervista iniziale		Compilazione del diario														Controllo e ritiro del diario											
			Visita intermedia														Intervista finale											
																	Inizio data entry diari											
																	Chiusura data entry diari											
GRUPPO 2			Primo contatto e intervista iniziale														Controllo e ritiro del diario											
																	Intervista finale											
																	Inizio data entry diari											
																	Chiusura data entry diari											

Al fine di coprire tutti i giorni di calendario, le famiglie del sottocampione mensile vengono suddivise in due gruppi: le famiglie del Gruppo 1, che compilano il diario per 14 giorni nella prima metà del mese di riferimento, e quelle del Gruppo 2, che lo com-

pilano per 14 giorni nella seconda metà. Ne deriva che anche l'intervista iniziale e la visita intermedia di controllo avvengono in periodi differenti. L'intervista finale, invece, è prevista nello stesso periodo per entrambi i gruppi, ossia all'inizio del mese successivo a quello di compilazione del diario. In tal modo si garantisce che i riferimenti temporali (cfr. par. 1.3.3.2) dei quesiti retrospettivi presenti nell'intervista finale siano identici per tutti (per tutte le famiglie che compilano il diario nel mese M, l'ultimo mese sarà M stesso, gli ultimi tre mesi faranno riferimento a M-2, M-1 e M, gli ultimi dodici mesi andranno da M-11 a M). In definitiva, dunque, le famiglie del Gruppo 1 vengono coinvolte nella rilevazione in un arco temporale di circa 65 giorni, mentre per quelle del Gruppo 2 il periodo di coinvolgimento è di circa 45 giorni.

I diari compilati sono acquisiti in CADi dagli stessi rilevatori, che, entro il 17esimo giorno del mese di effettuazione dell'intervista finale (M+1), devono inviare tramite *server* le interviste complete alla società incaricata della raccolta.

2.3.2 Gli strumenti

Al fine di ricostruire la spesa per consumi delle famiglie è necessaria un'architettura piuttosto complessa, che metta in relazione diversi strumenti, ognuno dei quali ha l'obiettivo di rilevare un particolare aspetto dell'informazione necessaria. La rilevazione è organizzata in più fasi facenti parte di un unico sistema di *data capturing* in modo da agevolare l'attività di raccolta del dato, minimizzando la possibilità di errore.

Il passaggio a tecniche CA ha consentito di gestire meglio e più efficacemente i tempi serrati della rilevazione, di guidare il rilevatore nella giusta sequenza di attività da svolgere e di controllare tempestivamente l'acquisizione delle informazioni. In tal senso, l'agenda e la scheda contatti rappresentano il nodo centrale del sistema (cfr. par. 2.3.2.1) che, oltre a standardizzare le modalità di contatto con le famiglie attraverso regole formali e sostanziali, gestiscono la tempistica della rilevazione, mettendo in collegamento tutti gli strumenti utilizzati per la raccolta dei dati (cfr. par. 2.3.2.2).

2.3.2.1 L'agenda e la scheda contatti

L'indagine sulle spese delle famiglie risulta decisamente onerosa e complessa per i rilevatori. È necessario dunque supportare e monitorare tutte le operazioni da essi effettuate al fine di tenere sotto controllo la qualità dei dati.

Con l'indagine PAPI la fase di assegnazione dei nominativi era gestita attraverso modelli cartacei (Mod. Istat BF/5) sui quali venivano annotati, con opportuni codici, gli esiti definitivi degli stessi. La gestione dei contatti era, dunque, demandata completamente al rilevatore, seppur opportunamente formato e supportato dal manuale di istruzioni.

L'uso di strumenti CA ha consentito di condurre e monitorare le operazioni di contatto con la famiglia in maniera automatica, evitando così la diretta e autonoma gestione da parte del rilevatore.

L'architettura classica

Nelle indagini CAPI l'agenda e la scheda contatti costituiscono un sistema integrato di supporto al rilevatore per la gestione del contatto con le famiglie, a partire dalla presentazione degli scopi dell'indagine fino alla conclusione dell'intervista (Istat, 2006b).



Tale sistema ha permesso di standardizzare il contatto con le famiglie, vincolando tutti i rilevatori al rispetto delle stesse procedure, anche in situazioni rare o problematiche. Il sistema, inoltre, garantisce un monitoraggio efficace e tempestivo dell'indagine, rappresentando la base per il calcolo degli indicatori di qualità. Infine, rappresenta lo strumento che permette il colloquio *client-server* tra *field* e centro di rilevazione, potendo connettersi dalla periferia verso l'unità di elaborazione centralizzata per scambiare dati, quali i nominativi da intervistare (in entrata) o i risultati della rilevazione (contatti, visite e interviste, in uscita).

Entrando più specificamente nelle caratteristiche del sistema, l'agenda si configura come un vero e proprio *organizer elettronico* per la pianificazione del lavoro sul campo, mentre la scheda contatti è un *questionario elettronico* che guida il rilevatore nel contatto (telefonico o al domicilio) con la famiglia, governando i modi e i tempi per accedere ai questionari di indagine (Figura 2.2).

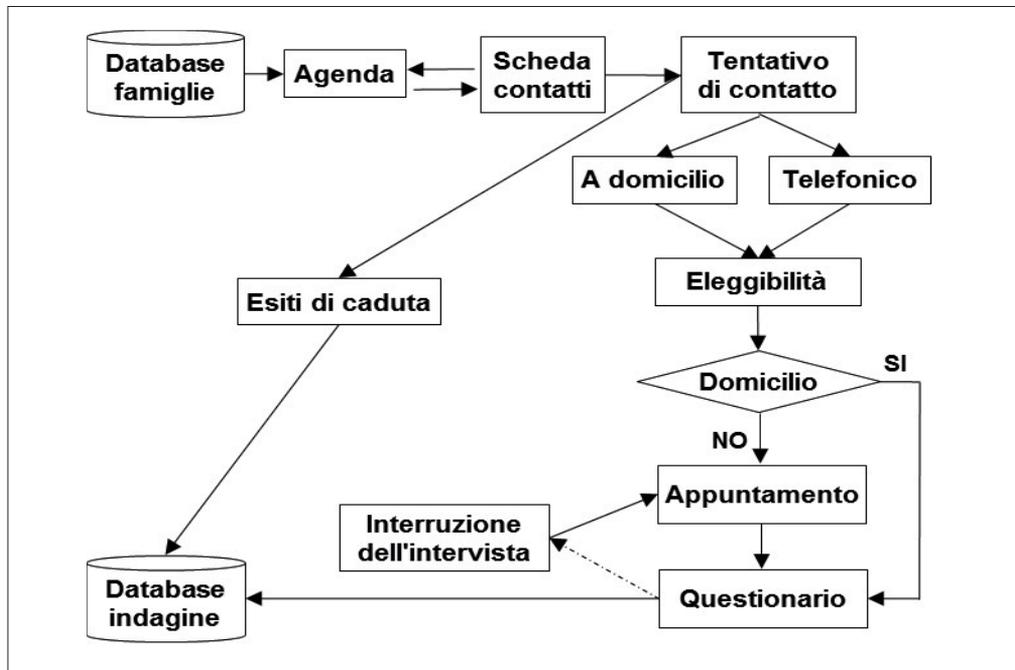
L'agenda permette di visualizzare: l'elenco delle famiglie da intervistare (all'inizio costituito dalle sole famiglie base, in quanto le corrispondenti famiglie sostitutive sono attivate in automatico in caso di sostituzione) e le informazioni necessarie per contattarle, la fase di intervista, lo stato delle interviste, gli appuntamenti, le eventuali note; in sostanza, è lo strumento che permette agli intervistatori di organizzare e monitorare in qualsiasi momento il proprio lavoro.

L'agenda consente l'accesso alla scheda contatti che governa:

- la selezione del tipo di contatto (telefonico o a domicilio);
- la gestione dei tentativi di contatto, attraverso percorsi guidati;
- il riconoscimento dell'eleggibilità della famiglia, attraverso la somministrazione di alcune domande chiave;
- la sostituzione dei nominativi in caso di caduta;
- la registrazione dell'esito del tentativo di contatto, in base al percorso seguito;
- la registrazione della data e dell'ora del contatto.

La registrazione da parte dei rilevatori di tutti i contatti avuti con le famiglie, indipendentemente dall'esito, rappresenta la necessaria premessa per il monitoraggio del lavoro sul campo e il calcolo degli indicatori di qualità. È così possibile, in qualsiasi momento, conoscere la numerosità delle interviste sospese (famiglie mai toccate, interviste temporaneamente interrotte), cadute e complete.

Figura 2.2 – Schema del flusso della scheda contatti

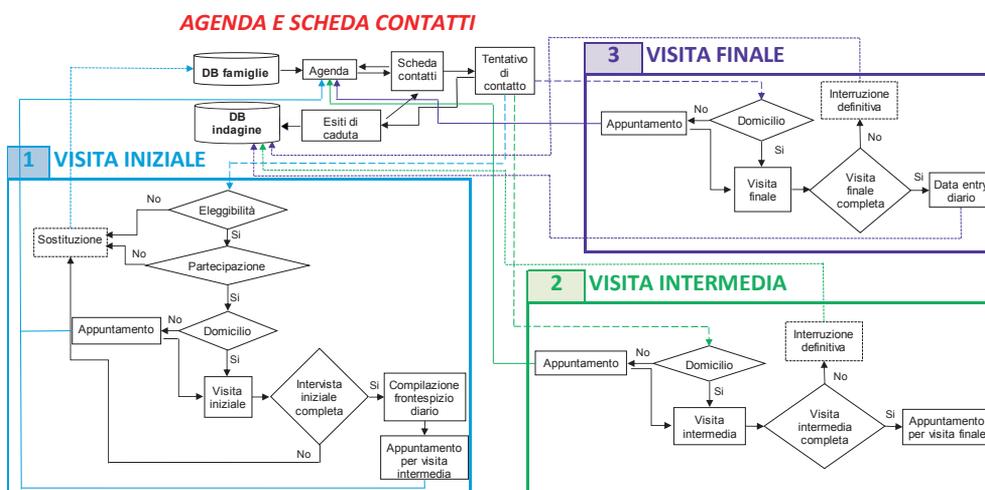


L'adattamento alle specificità dell'indagine sulle spese

Nell'indagine CAPI sulle spese delle famiglie, per tener conto delle specificità della rilevazione, l'agenda e la scheda contatti hanno subito degli adattamenti rispetto alla formulazione standard già descritta.

Questa scelta nasce dall'esigenza sia di tenere sotto controllo le tre fasi della rilevazione, in termini temporali e di contenuto, sia di creare un ponte di collegamento tra interviste CAPI e diari cartacei (Figura 2.3).

Figura 2.3 – Schema del flusso della scheda contatti nell'indagine sulle spese delle famiglie



L'agenda e la scheda contatti rappresentano dunque il nucleo centrale della rilevazione, attraverso il quale si accede alle diverse fasi in una corretta sequenza temporale.



Proprio al fine di gestire i tempi delle tre distinte occasioni di incontro con le famiglie, oltre alle consuete informazioni necessarie per reperirle, sono state inserite in agenda le date di inizio e fine dei periodi previsti per ciascuna visita, con l'obiettivo di ridurre al minimo il rischio di errori e di evidenziare eventuali sovrapposizioni nei tempi di lavorazione di famiglie appartenenti a periodi di rilevazione diversi. L'agenda, infatti, consente o inibisce al rilevatore l'accesso alle singole fasi (e ai relativi questionari) in base ai giorni di calendario, visualizzando l'informazione relativa allo stato dell'intervista e quindi allo stato di avanzamento dei lavori per ciascuna famiglia assegnatagli (Prospetto 2.2).

Per quanto riguarda i contenuti, oltre alle consuete funzionalità di una scheda contatti (Istat, 2006b), per l'Indagine sulle spese è stata inclusa la gestione dei diari cartacei:

- è stata prevista, nell'ambito del primo contatto, la fase di consegna del diario;
- è stata fornita al rilevatore una guida al controllo del diario durante le visite intermedia e finale;
- è stato previsto l'accesso al *data entry* per l'acquisizione dei diari compilati dalle famiglie con intervista CAPI completa.

Prospetto 2.2 – L'agenda elettronica

Chiave famiglia	Giorno e ora appuntamento	Mese di riferimento	Famiglia base/sostitutiva	Stato intervista	Nominativo	Comune	Indirizzo	Telefono	Data inizio periodo visita iniziale	Data fine periodo visita iniziale	Data inizio periodo visita intermedia	Data fine periodo visita intermedia	Data inizio periodo visita finale	Data fine periodo visita finale

A partire dalle informazioni contenute in agenda, viene generato un codice di aggancio, anonimo e univoco, che consente di integrare le diverse fasi di indagine e di associare tra loro le informazioni rilevate con tecniche differenti. Il codice di aggancio deve essere trascritto sul diario cartaceo dal rilevatore in fase di consegna alla famiglia (e cioè durante la prima visita) e poi digitato nuovamente per accedere alla maschera di acquisizione in fase di *data entry*. La trascrizione dei codici di aggancio è un'operazione tanto fondamentale quanto delicata, da condurre con estrema accuratezza, poiché un errore di redazione comprometterebbe il raccordo fra tutti i dati familiari raccolti, dunque la qualità dell'intero processo di produzione del dato.

La gestione e il controllo della correttezza della compilazione del diario rappresentano un'altra fase fondamentale e delicata del processo; per tale motivo, la scheda contatti propone un elenco delle problematiche più comunemente riscontrate nella compilazione da parte delle famiglie, affinché il rilevatore possa verificare e registrarne la presenza/assenza e procedere poi all'eventuale correzione sul modello cartaceo, prima del *data entry*; in tale modo l'informazione relativa alla presenza di errori viene mantenuta, rendendo possibile intervenire sulle cause.

2.3.2.2 La struttura dei questionari

L'ampliamento dei contenuti informativi, a causa dell'armonizzazione alla classificazione Coicop, delle rinnovate esigenze informative e dell'integrazione con l'indagine viaggi e vacanze, rende estremamente difficile la conduzione dell'intervista in un'unica occasione, come avveniva con l'indagine PAPI; la durata media dell'intervista avrebbe infatti sfiorato i 50 minuti. Si è dunque stabilito di suddividere l'intervista diretta in due fasi (cfr. par. 1.3.1): un'intervista iniziale da somministrare durante la prima visita di durata media pari a circa 20 minuti (variabili in base al numero dei componenti della famiglia) e un'intervista finale da somministrare durante la terza visita di durata media pari a circa 30 minuti.

L'intervista iniziale

L'intervista iniziale si compone di 4 sezioni (Figura 2.4).

Sezione 1 - composizione familiare. Le persone appartenenti alla famiglia di fatto vengono distinte da quelle coabitanti con la famiglia, ma legate ad essa da soli vincoli economici (ad esempio personale di servizio convivente o affittuari di porzioni di abitazione).

Tale distinzione risulta fondamentale per individuare esattamente l'unità economica di riferimento per la rilevazione della spesa per consumi; le persone legate da soli vincoli economici, pur vivendo con la famiglia, non ne condividono le risorse economiche e quindi le scelte di consumo.

Tuttavia, poiché la loro presenza modifica i profili di consumo e di spesa della famiglia di fatto (ad esempio aumenta i livelli di consumo alimentari e abitativi così come può implicare la spesa per retribuzioni di personale convivente), è necessario considerarle, nonostante le spese da loro sostenute non rientrino nella spesa per consumi della famiglia di fatto. Per quanto detto, di queste persone vengono rilevate solo le principali caratteristiche demografiche.

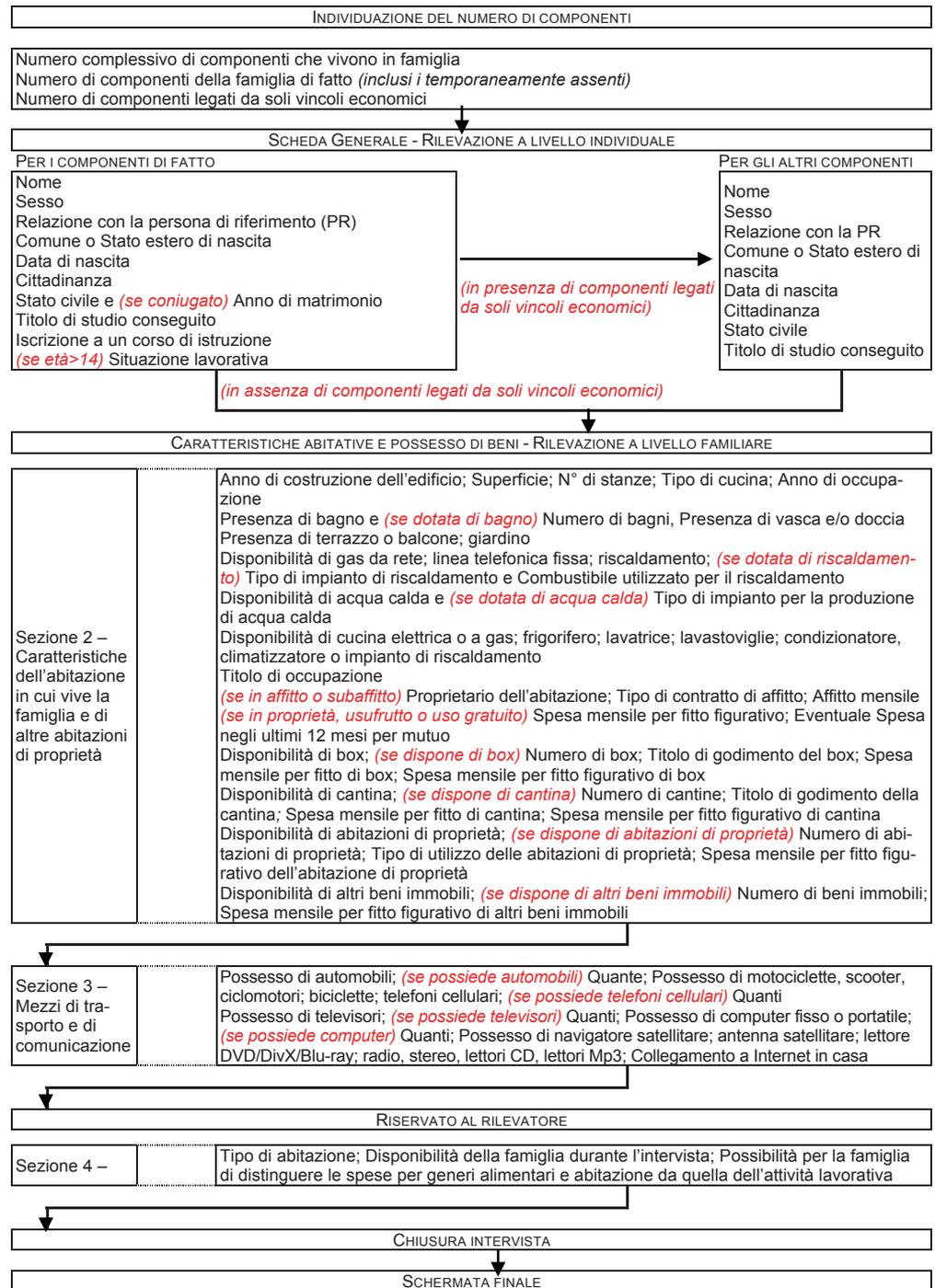
Sezione 2 - scheda generale. Si riferisce alle informazioni socio-demografiche dei singoli individui appartenenti alla famiglia: la relazione con la persona di riferimento, l'età, il sesso, la cittadinanza e il paese di nascita, lo stato civile, il titolo di studio e la situazione lavorativa. La data di nascita e il sesso, assieme al numero di componenti, guidano i percorsi e il *wording* del questionario. Sono dunque previsti diversi quesiti di controllo al fine di accertarne la correttezza assieme alla famiglia.

Il questionario CAPI ha inoltre permesso di inserire la rilevazione della condizione lavorativa oggettiva (condizione ILO) e delle professioni con l'utilizzo del motore di ricerca specifico che si basa sulla Nomenclatura e classificazione delle Unità Professionali (CP2011).

Sezione 3 - caratteristiche abitative e il possesso di alcuni beni durevoli e semi-durevoli. In questa sezione vengono rilevate alcune spese periodiche per l'abitazione principale e per eventuali altri beni immobili a disposizione della famiglia. In particolare, vengono rilevate le spese mensili per affitto o, alternativamente, il valore del fitto figurativo (sia soggettivo sia oggettivo seguendo le linee guida europee).

Sezione 4 - riservata al rilevatore. Completa il quadro sulle caratteristiche dell'abitazione principale e contiene una prima valutazione, richiesta al rilevatore, del grado di disponibilità a partecipare da parte della famiglia.

Figura 2.4 – Il flusso dell'intervista iniziale

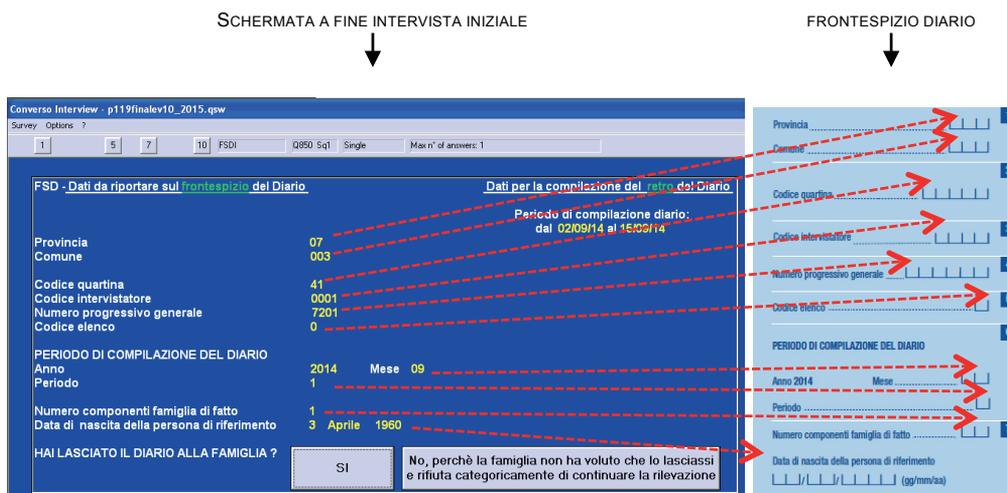


Al rilevatore vengono inoltre chieste informazioni sulla capacità della famiglia di valutare o distinguere le spese per generi alimentari e per l'abitazione da quelle inerenti l'attività lavorativa. Può accadere, infatti, che alcune famiglie, a causa di una particolare attività lavorativa, non siano in grado di quantificare le spese relative ad alcuni loro consumi perché i) non riescono a distinguerle da quelle dell'attività lavorativa oppure ii) non le sostengono in quanto fornite dal datore di lavoro come parte

del compenso. È il caso ad esempio di gestori di ristoranti o di alberghi, proprietari di aziende agricole e simili, ma anche di sacerdoti che risiedono all'interno della canonica o di custodi che risiedono all'interno del fabbricato custodito. L'informazione è di supporto per comprendere alcune omissioni sul questionario e per analizzare in modo appropriato il comportamento di spesa di alcune famiglie.

Il questionario per l'intervista iniziale termina con una schermata che riporta i codici da trascrivere, a cura del rilevatore, sul frontespizio del diario cartaceo da lasciare alla famiglia per la compilazione; tali codici sono necessari per consentire l'aggancio tra i vari strumenti utilizzati (scheda contatti, diario, questionario intervista iniziale e questionario intervista finale) (Figura 2.5). I codici si riferiscono agli identificativi della provincia e del comune di residenza della famiglia, della quartina a cui la famiglia appartiene, del rilevatore a cui è assegnata l'intervista, della posizione della famiglia nell'elenco base (numero progressivo generale) e nella quartina di appartenenza (codice elenco). Tali codici rappresentano insieme il codice univoco e anonimo assegnato ad ogni singola famiglia. Vengono inoltre riportati l'anno, il mese e il periodo di partecipazione all'indagine della famiglia, oltre al numero di componenti di fatto e alla data di nascita della persona di riferimento (intestatario della scheda familiare anagrafica); tali informazioni, seppur non strettamente necessarie per l'identificazione della famiglia, vengono riportate nel frontespizio del diario al fine di correggere eventuali errori di trascrizione che comprometterebbero l'aggancio dei diversi strumenti. In fondo alla schermata, compare anche la registrazione dell'esito dell'intervista iniziale e della consegna del diario, in modo da segnalare l'eventuale rifiuto da parte della famiglia a proseguire la collaborazione.

Figura 2.5 – La trascrizione dei codici dal questionario elettronico (intervista iniziale) al cartaceo (diario)



Il diario

Il diario viene utilizzato per rilevare le spese correnti, le spese cioè che si caratterizzano per un'elevata frequenza (spesso) associata ad entità abbastanza contenute. Si tratta dei generi alimentari e delle bevande, ma anche dei generi non alimentari di uso e acquisto quasi quotidiano (ad esempio i tabacchi, i carburanti, i biglietti per trasporti o i giornali). Le famiglie devono annotare, giorno per giorno per un periodo di



14 giorni, le spese sostenute e, nel caso vi siano le condizioni, le quantità consumate di beni autoprodotti.

Rispetto all'indagine PAPI, il diario, come già detto, è stato ampliato sia nei contenuti sia nella lunghezza del periodo di compilazione ed è stato rivisto interamente nella sua veste grafica⁵, al fine di renderlo maggiormente chiaro.

La parte iniziale è riservata a una breve presentazione dell'indagine e delle principali regole per la compilazione ed è stata notevolmente arricchita rispetto all'indagine PAPI (le pagine ad essa dedicata sono passate da due a quattro). Oltre alle spiegazioni iniziali, sono stati inseriti suggerimenti anche all'interno delle pagine da compilare (Figura 2.6).

La parte da compilare è stata suddivisa in due blocchi identici da 7 giorni, ognuno composto da 5 sezioni, e si presenta pertanto come una doppia agenda settimanale. Le sezioni riguardano:

- Spese per bar, pasti fuori, da portar via
- Spese per cibi e bevande
- Spese per generi non alimentari
- Beni prodotti dalla famiglia
- Luoghi di acquisto.

Figura 2.6 – I suggerimenti inseriti nel diario

	GIORNO 1	GIORNO 2	GIORNO 3	GIORNO 4	GIORNO 5	GIORNO 6	GIORNO 7
TABACCHI, SIGARETTE ELETTRONICHE	spesa in euro						
SIGARETTE Sigarette, escluse sigarette elettroniche	01099	02099	03099	04099	05099	06099	07099
SIGARI Sigari, inclusi sigarioli	01100	02100	03100	04100	05100	06100	07100
TABACCO Tabacco da sigaretta, da pipa, da masticare, da fumo, inclusi capsule e filo per sigarette	01101	02101	03101	04101	05101	06101	07101
SIGARETTE ELETTRONICHE E LORO ACCESSORI Sigarette elettroniche, ricaricatori, liquidi, liquori, adattatori e altri ricambi e accessori, inclusi kit complete	01102	02102	03102	04102	05102	06102	07102
MEDICINALI ARTICOLO SANITARI	Annotate anche le spese senza scontrini o ricevute						
MEDICINALI (esclusi farmaci rinfarmaci pagati anche se esenti al solo titolo) Medicinali, inclusi infusori a infusioni, prodotti omeopatici o di altra medicina non convenzionale (inclusi prodotti veterinari)	01103	02103	03103	04103	05103	06103	07103
ARTICOLO SANITARI Corsi, impasti, saponi, kit da pronto soccorso, dermatologia, fascia elastiche e simili	01104	02104	03104	04104	05104	06104	07104
CURA E SERVIZI DI ADULTI E BAMBINI	Se non sapete dove scrivere una spesa riportatela nella voce "altro" e specificate il bene o il servizio acquistato						
PARRUCCHIERE PER DONNA Esclusi trattamenti estetici	01105	02105	03105	04105	05105	06105	07105
PARRUCCHIERE E BARBIERE PER UOMO E BAMBINO Esclusi trattamenti estetici	01106	02106	03106	04106	05106	06106	07106
TRATTAMENTI ESTETICI Solertum, istituti di bellezza, estetista e simili	01107	02107	03107	04107	05107	06107	07107

Le spese per generi alimentari acquistati (ed eventualmente anche consumati) presso luoghi di ristorazione sono state inserite in una sezione a parte, anticipandole rispetto alle spese per i generi alimentari acquistati presso gli esercizi distributivi (negozi tradizionali, supermercati, mercati e simili) (Figura 2.7). Nell'indagine PAPI, infatti, l'averle collocate alla fine delle spese alimentari, in alcuni casi, ne determinava la cattiva rilevazione: ad esempio, la voce caffè (riferita all'acquisto di polvere, cialde o simili presso un esercizio commerciale) conteneva molto spesso anche la spesa per consumazioni di caffè al bar.

⁵ La nuova veste grafica del diario è stata curata da Sofia Barletta.

2. Il disegno della nuova indagine

Tenendo conto di quanto richiesto da Eurostat sulla classificazione dei luoghi di ristorazione, tali spese sono state classificate distinguendole in base a quanto del loro ammontare è riconducibile al servizio (ad esempio nel caso di pasti consumati presso ristoranti, il costo del servizio è pari, se non superiore, a quello del bene; la stessa cosa non accade se si effettuano consumazioni presso *take away*, dove la spesa sostenuta si riferisce quasi esclusivamente al costo del bene consumato). Si distingue pertanto tra i) bar, pasticcerie, gelaterie, chioschi e ambulanti; ii) rosticcerie, pizzerie al taglio, *self-service*, *fast-food*, *take away* senza servizio al tavolo; iii) ristoranti, osterie e trattorie con servizio al tavolo; iv) distributori automatici.

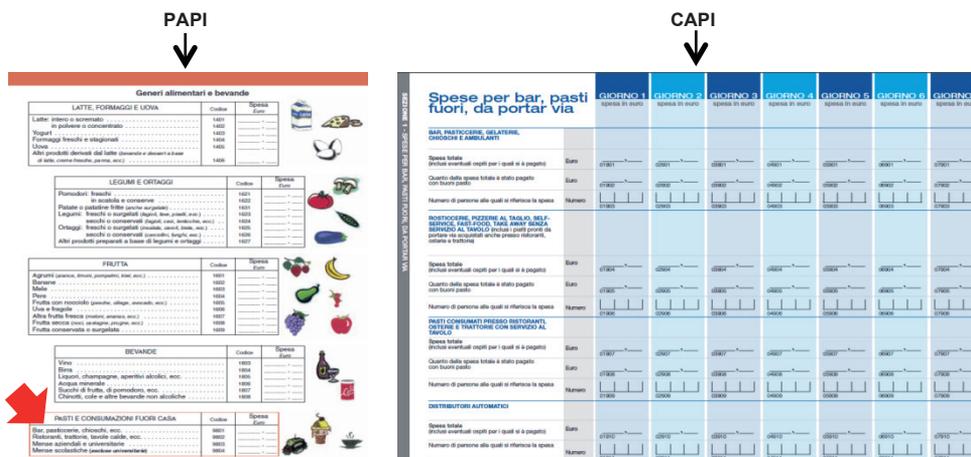
L'introduzione di una sezione separata per generi alimentari acquistati presso luoghi di ristorazione ha consentito anche di arricchirne i contenuti: oltre alla spesa sostenuta, viene rilevato il numero di persone a cui si riferisce (informazione utile per comprendere meglio le abitudini alimentari fuori casa delle famiglie) e la quota di spesa pagata con i buoni pasto (informazione utile anche nell'ambito delle stime di contabilità nazionale).

Un altro cambiamento riguarda la spesa per le mense, che non viene più rilevata con il diario, come nell'indagine PAPI, ma con l'intervista finale e, nel caso delle mense scolastiche, in associazione alle altre spese connesse con l'attività scolastica o di formazione universitaria.

Con la sezione 2 vengono rilevate le spese per alimentari e bevande, con un dettaglio di 96 voci (nella PAPI erano 62); le 71 voci indicate da Eurostat sono state in alcuni casi ulteriormente suddivise al fine di agevolare la compilazione, rendendo più esplicite le definizioni e facilitando il ricordo.

Con la sezione 3 vengono rilevati gli acquisti non alimentari di frequenza elevata (66 voci di spesa contro le 44 della PAPI), che rappresentano il complemento alle spese non alimentari (meno frequenti) rilevate con l'intervista finale.

Figura 2.7 – La rilevazione delle spese per pasti e consumazioni fuori casa. Confronto tra PAPI e CAPI



La sezione 4 rileva la quantità di beni prodotti e consumati dalla famiglia durante il periodo di compilazione del diario (cfr. par. 1.3.4.4); a differenza della indagine PAPI, nella quale la rilevazione degli autoconsumi veniva effettuata con uno specifico diario, consegnato alle sole famiglie che dichiaravano di effettuare l'autoconsumo,



nell'indagine CAPI tale rilevazione è parte integrante del diario. In modo da aiutare le famiglie a distinguere la registrazione dell'autoconsumo da quella delle spese, per i beni autoprodotti è stata utilizzata una grafica molto diversa (Figura 2.8). Le famiglie possono scegliere tra più unità di misura (grammi o chili, litri, bicchieri o cucchiaini; per le uova le quantità devono essere riportate in numero) indicando, a differenza di quanto avveniva in passato, solamente le quantità consumate o regalate e non più la stima del loro valore monetario a prezzi di mercato. Tale valore viene infatti stimato direttamente dall'Istat sulla base delle informazioni rilevate dal Servizio Prezzi. La classificazione dei beni autoconsumati è perfettamente sovrapponibile a quella delle spese, così da poter sommare la spesa e il valore monetario per i diversi beni.

Figura 2.8 – La rilevazione dei beni prodotti dalla famiglia

The image shows a detailed view of the CAPI diary form. It is organized into two main columns, each representing a week of data (GIORNO 1-7). The form is divided into several sections based on the type of goods consumed:

- Beni alimentari (alimenti):** This section includes sub-sections for 'MATERIE PRIME', 'MATERIE SECONDE', and 'MATERIE TERZIARIE'. It lists various food items like 'Margarine, confetto di burro, margine', 'Miele, zucchero', 'Uova', 'Latticini', 'Cereali', 'Legumi', 'Pasta', 'Rizzi', 'Carne', 'Pesce', 'Frutta', 'Verdure', 'Erbe aromatiche', 'Spezie', 'Altri prodotti', 'Altri beni'. Each item has columns for quantity in different units (grams, kilos, liters, etc.) for each day.
- Beni non alimentari:** This section includes 'Altri beni' and 'Altri prodotti'.
- Uova:** A separate section for recording egg consumption, with columns for quantity in units.

At the bottom right, a legend explains the symbols used in the form:

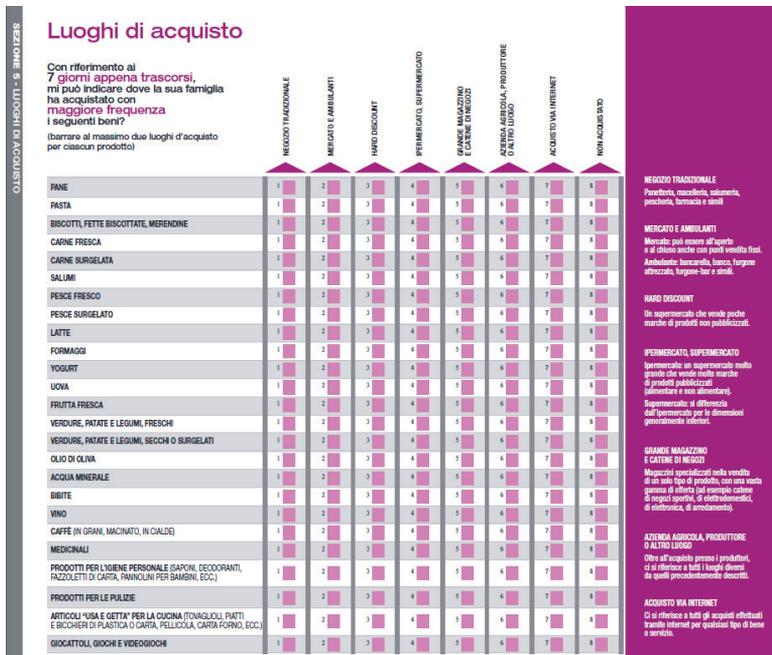
- Green box:** "Quantità di beni prodotti in proprio (regalati o consumati)" - Quantity of self-produced goods (gifted or consumed).
- Grey box:** "Non quantificati consumo di beni acquistati in negozio" - Non-quantified consumption of purchased goods in a store.

The text "giorno 1 = giorno 7" is visible at the bottom of the form, indicating that the data for the first and seventh days are the same.

La quinta e ultima sezione è riservata alla rilevazione dei luoghi di acquisto. Per alcuni beni viene chiesto alla famiglia di indicare dove li ha acquistati con maggior frequenza (Figura 2.9).

Nella parte finale del diario sono stati inseriti quattro quesiti che hanno lo scopo di indagare due aspetti diversi: i) i cambiamenti nel comportamento di acquisto; ii) la qualità della rilevazione tramite diario.

Figura 2.9 – La rilevazione dei luoghi di acquisto



Alla famiglia viene chiesto di indicare se, per alcuni beni o servizi e rispetto all'anno precedente, ha cercato di limitare la spesa, acquistando meno quantità e/o acquistando prodotti di qualità o marca diverse. Tale indicatore, se analizzato in serie storica, evidenzia come le famiglie rispondono agli eventi congiunturali o personali mettendo in atto strategie di contenimento della spesa (Figura 2.10).

A differenza dell'indagine PAPI, dove il quesito era parte dell'intervista finale, nella nuova indagine l'informazione è rilevata tramite diario, l'autocompilazione consente di evitare situazioni di imbarazzo per le famiglie che possono aver sperimentato un peggioramento della propria condizione economica.

Per valutare la qualità della rilevazione, si chiede alla famiglia i) di indicare quanto stancante è stata la compilazione del diario; ii) se in alcuni giorni, pur avendo effettuato spese, la compilazione si è interrotta; iii) chi ha annotato le spese riportate (anche eventuali persone esterne alla famiglia). In tal modo è possibile valutare la qualità della partecipazione e capire se l'assenza di spese in alcuni giorni della settimana è imputabile all'effettiva mancanza di acquisti piuttosto che a una mancata registrazione. Quindi, anche per il diario, può essere ottenuta una indicazione, seppur generica, circa gli zeri di spesa.

Figura 2.10 – Le abitudini di spesa

E per finire qualche domanda

1. Rispetto a un anno fa ha cercato di limitare la spesa per l'acquisto dei seguenti beni o servizi?

	NO	SI		NON LO COMPRABO E CONTINUO A NON COMPRARLO
	NON HO CERCATO DI LIMITARE LA SPESA	HO CERCATO DI LIMITARE LA SPESA ACQUISTANDO MENO	HO CERCATO DI LIMITARE LA SPESA ACQUISTANDO PRODOTTI DIFFERENTI O DI MARCHE DIVERSE	
ALIMENTARI	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
BEVANDE	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
ABBIGLIAMENTO E CALZATURE	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
SPESA PER LA CURA E L'IGIENE PERSONALE (prodotti di bellezza, parrucchiere, barbiere, estetista, ecc.)	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
VISITE MEDICHE ED ACCERTAMENTI PERIODICI DI CONTROLLO PREVENTIVO (dentista, mammografia, pap-test, ecc.)	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
CARBURANTI PER MEZZI PRIVATI (benzina, gasolio, ecc.)	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>
VIAGGI E VACANZE	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>

L'intervista finale

L'intervista finale si compone di tre parti (Figura 2.11): la prima si riferisce a tutte le spese sostenute dalla famiglia non rilevate tramite intervista iniziale o diario; la seconda approfondisce, a livello individuale, l'informazione sulle caratteristiche e sulle spese sostenute per viaggi e vacanze (si rimanda allo specifico volume metodologico la cui uscita è prevista nel 2016); la terza riguarda la condizione economica e reddituale della famiglia.

Nello specifico l'intervista finale si apre con una sezione orientata ad approfondire aspetti specifici nei comportamenti di spesa e viene modificata ogni anno; per il 2014 l'approfondimento ha riguardato l'acquisto di alimenti da tenere come scorta e di prodotti biologici, mentre nel 2015 il focus riguarda beni e servizi autoprodotti.

Nelle sezioni successive vengono rilevate le spese. La sequenza dei quesiti segue un doppio criterio: le singole voci sono state raggruppate in base al periodo di riferimento (ultimi 12 mesi, ultimi 3 mesi e ultimo mese) e alla categoria del bene o servizio (capitolo di spesa cui appartengono).

Le spese per trasporti, ad esempio, vengono rilevate con riferimento ai 12 mesi precedenti l'intervista nel caso di eventi rari (si pensi all'acquisto dell'automobile) o di eventi con cadenza annuale (la spesa per l'assicurazione).

Figura 2.11 – Il flusso dell'intervista finale



Via via che l'entità della spesa diventa meno elevata e/o la frequenza aumenta, si riduce il periodo di riferimento al fine di limitare l'effetto ricordo: le spese per i biglietti aerei, ad esempio, vengono rilevate con riferimento agli ultimi tre mesi, mentre

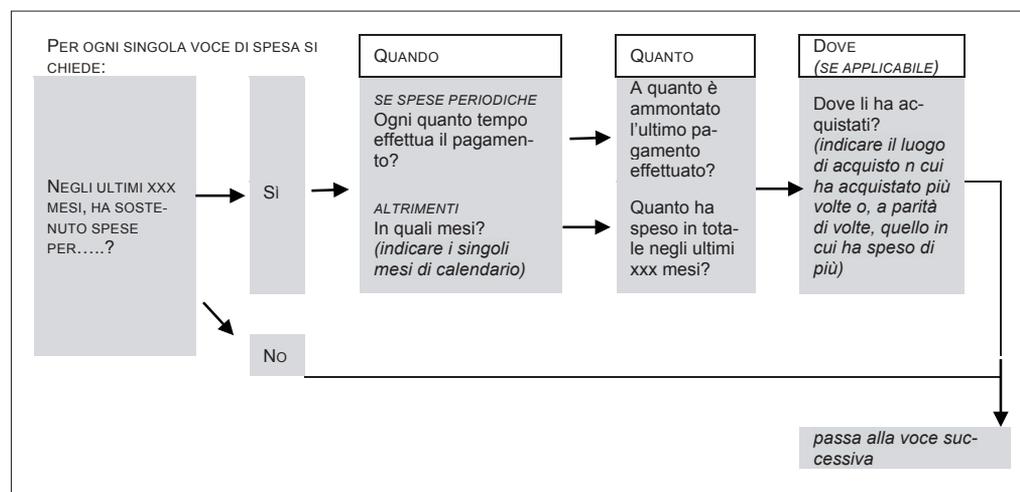


le spese per pezzi di ricambio si riferiscono all'ultimo mese. Il questionario elettronico ha permesso una gestione più efficiente dei periodi di riferimento: l'introduzione di testi mobili, che cambiano in base al mese di effettuazione dell'intervista, consente l'esatta individuazione dei periodi di riferimento e aiuta la famiglia nella collocazione temporale delle spese.

Per ogni voce di spesa è stato, inoltre, introdotto un quesito filtro con il quale distinguere le famiglie che hanno sostenuto la spesa da quelle che non l'hanno sostenuta (evitando per queste ultime la formulazione di quesiti di approfondimento sulla spesa stessa) (Figura 2.12). L'introduzione di filtri per ogni voce di spesa rilevata ha, inoltre, facilitato l'individuazione delle mancate risposte parziali (distinguendo le famiglie che non hanno effettuato la spesa da quelle che, pur avendola effettuata, non sono in grado di ricordarne l'entità, il periodo o il luogo di acquisto) e il conseguente trattamento mediante metodi di imputazione (cfr. par. 2.4.2).

Oltre ai filtri sulle singole voci di spesa, il questionario prevede anche percorsi filtrati. In particolare, vengono sottoposte a filtro le domande relative alle spese per le abitazioni secondarie (non applicabili a chi non le ha a disposizione), per istruzione (non applicabili a famiglie senza componenti in formazione), per personale di servizio (non applicabili a chi non se ne avvale).

Figura 2.12 – La rilevazione delle spese in modo retrospettivo



Il questionario elettronico ha inoltre permesso di inserire complessi controlli di coerenza tra le variabili (anche tra questionari diversi) e di attivare segnalazioni di violazioni di tali coerenze in fase di rilevazione, così da permetterne l'immediata correzione. Ad esempio, sono stati inseriti controlli sulla coerenza tra l'età dei componenti, il grado di istruzione e le spese sostenute per grado d'istruzione; tra la presenza di abitazioni secondarie e le spese per le utenze e i servizi ad esse connesse; tra la presenza di personale domestico coabitante e le spese per la retribuzione dello stesso; tra la presenza di alcuni servizi nell'abitazione e le spese sostenute per il loro utilizzo. In quest'ultimo caso, l'incoerenza si verifica quando: i) la famiglia dichiara di non disporre del servizio, ma di aver sostenuto la spesa per il suo utilizzo; ii) la famiglia dichiara di disporre del servizio, ma di non aver sostenuto spese. L'incoerenza del primo tipo, genera un quesito di controllo per verificare se quanto dichiarato

nell'intervista iniziale (assenza del servizio) è corretto e in caso negativo è possibile rettificare. Se l'incoerenza è del secondo tipo, il quesito di controllo indaga i motivi dell'assenza di spesa.

Nel 2013, la percentuale di famiglie che dichiara di non aver sostenuto la spesa per l'energia elettrica, nonostante l'abitazione in cui vivono ne è dotata, è pari all'1,4%; i quesiti di controllo della CAPI hanno consentito di verificare che l'assenza di spesa è per lo più giustificata (nella quasi totalità dei casi è attribuibile al fatto che la spesa è stata sostenuta da persone esterne alla famiglia o non è scorponabile da altre spese, quali l'affitto) e non deve esser imputata.

Nei controlli di coerenza, vi rientrano anche quelli che impediscono il proseguimento dell'intervista; ad esempio, sempre per quanto riguarda i servizi connessi con l'abitazione, l'intervista non può proseguire in presenza di una mancata risposta.

L'ultima parte del questionario raccoglie informazioni sulla situazione economica e sul reddito mensile della famiglia. Si chiede alla famiglia una valutazione sia della propria capacità di raggiungere lo standard di vita desiderato con le risorse economiche a disposizione, sia dell'ammontare di reddito netto medio mensile familiare complessivo. Qualora la famiglia abbia difficoltà a fornire il valore puntuale, le si chiede di fornire almeno un valore approssimativo, scegliendo, tra una serie di valori proposti, quello che più si avvicina al reddito familiare. L'aggiunta del quesito di rilancio, nel 2013, ha permesso di recuperare l'informazione per circa la metà delle famiglie che non avevano fornito il valore puntuale: la quota di famiglie che non lo aveva indicato era dell'8,8%, ma è scesa al 4,6% con il quesito di rilancio.

2.3.3 La rete di rilevazione

2.3.3.1 L'organizzazione della rilevazione e dei rapporti con la rete

L'affidamento a una società esterna dell'attività di rilevazione ha comportato una rivisitazione sostanziale dell'impianto organizzativo dell'indagine, ridisegnando ruoli e funzioni di tutti gli attori coinvolti. La società affidataria è responsabile del reclutamento e della gestione della rete di rilevazione, nonché della ripartizione dei carichi di lavoro, mentre i comuni hanno assunto un ruolo di affiancamento alla rete di rilevazione, di assicurazione degli intervistati e di supporto ai rilevatori nella ricerca delle famiglie. La progettazione, il coordinamento e il controllo dell'indagine sono ovviamente rimasti prerogative dell'Istat.

In questa nuova cornice organizzativa, l'attenzione dell'Istituto nel curare il rapporto con i rilevatori, in tutte le fasi del lavoro sul campo, è stata elevata. L'attività svolta dalla rete, infatti, è decisiva per conquistare la fiducia degli intervistati e rilevare in maniera corretta ed esaustiva le informazioni richieste, per assicurare cioè la qualità e l'affidabilità dei dati. In questo senso è fondamentale conquistare anche la rete, instaurando un rapporto di collaborazione e di sostegno continuo, che generi un coinvolgimento attivo e responsabile in tutte le operazioni che le competono.

L'approccio seguito nel trasmettere e consolidare le conoscenze per gestire la rilevazione si basa sul principio dell'unitarietà del disegno di indagine: ogni fase si alimenta delle fasi precedenti e, a sua volta, influenza le successive, sollecitando continue revisioni e adattamenti del disegno, sia *in itinere* sia *a posteriori*, per le successive occasioni di indagine. Allo stesso modo, le attività che vengono proposte

alla rete non sono indipendenti tra di loro, ma strettamente correlate: la formazione, ad esempio, non si esaurisce nelle riunioni di *briefing* iniziali, ma si aggiorna continuamente attraverso le occasioni di *debriefing* e le attività di assistenza e di richiamo formativo. L'esperienza condotta nei *debriefing*, in cui formazione e monitoraggio si fondono perfettamente, dimostra come, in un modello organizzativo caratterizzato da una continua contaminazione tra le varie fasi, i *feedback* tra i diversi attori favoriscono il raggiungimento degli obiettivi e assicurano un dato di qualità.

Il mutamento organizzativo di cui si è detto e il passaggio, in particolare, a una tecnica di somministrazione CA hanno modificato l'offerta formativa e il sistema di monitoraggio dell'indagine. Riguardo alla prima, il passaggio da una proposta "a cascata", dispersa sul territorio, a un'offerta unica centralizzata, gestita d'intesa con la società, ha permesso di trasmettere in maniera più efficace, e soprattutto più uniforme, le conoscenze dell'indagine. L'impiego dei computer ha ampliato il perimetro dei contenuti trattati, sollecitando una maggiore integrazione tra aspetti teorici e parti esercitative. Da questo punto di vista, è stata sperimentata una modalità cosiddetta "a doppio binario", nella quale gli approfondimenti tematici dei quesiti sono stati affrontati di pari passo con la somministrazione in aula del questionario, rendendoli più circostanziati e pertinenti. Inoltre, sono stati maggiormente valorizzati i moduli dedicati alla comunicazione, scegliendo di affiancare e sostenere i rilevatori nel compito più delicato del loro lavoro, quello di farsi accogliere dai rispondenti e motivarli a collaborare.

Riguardo al sistema di monitoraggio, invece, le informazioni raccolte dalla rete e trasmesse in tempo reale hanno consentito un trattamento tempestivo delle criticità, nonché prevenuto la reiterazione di comportamenti anomali in sede di intervista. In questo modo si è passati da un sistema di monitoraggio frammentato ed *ex-post* a un sistema di monitoraggio integrato e *in itinere*. Non solo, la maggiore integrazione tra le indagini, perseguita a livello di Istituto, ha permesso di adottare degli indicatori standardizzati, in grado di osservare le performance della rete e di condividere le risultanze dei rilevatori impiegati su più indagini.

Le pagine seguenti sono dedicate ad approfondire le novità introdotte nelle attività di formazione, gestione e monitoraggio della rete di rilevazione dell'Indagine sulle spese. Fin da ora, tuttavia, è bene sottolineare che la progettazione delle occasioni di contatto con la rete sono state impostate riconoscendo il protagonismo di tutti gli attori coinvolti nei processi formativi e nella valutazione delle prassi di ricerca. I rilevatori non sono figure da istruire e controllare, ma professionisti da includere, motivare e affiancare, sostenendoli attraverso attività continue di formazione e di monitoraggio, nelle quali possono essere a loro volta fonti di conoscenza rilevanti per migliorare la qualità dell'indagine.

2.3.3.2 La formazione dei rilevatori

Attività e strumenti per la formazione

Come anticipato, la rete dei rilevatori è gestita dalla società esterna, alla quale sono conferite le attività di reclutamento, assegnazione dei carichi di lavoro individuali e di liquidazione dei compensi relativi alle interviste realizzate. La selezione avviene nel rispetto dei requisiti minimi stabiliti dall'Istat, che sostanzialmente corrispondono all'aver la maggiore età, essere in possesso almeno di un diploma di scuola media

superiore, avere conoscenze informatiche di base ed essere automuniti. Titolo preferenziale è rappresentato dall'aver maturato in precedenza un periodo di almeno tre mesi di esperienza come rilevatore per indagini statistiche.

Ai rilevatori non viene assegnato alcun carico di lavoro prima di aver partecipato alla formazione d'aula iniziale, alla cui realizzazione concorrono l'Istat, per gli aspetti metodologici e contenutistici dell'indagine, e la società, per quanto attiene quelli tecnici del sistema CAPI.

Nel 2012, sono stati formati complessivamente 474⁶ rilevatori; nel 2013 e nel 2014, a causa di tassi di uscita⁷ pari al 42,8 e al 48,8%, è stato necessario formare rispettivamente 139 e 80 nuove risorse. Ovviamente, l'intero serbatoio di rilevatori, formati e attivi durante l'anno d'indagine, non sono mai operativi sul territorio contemporaneamente, in ragione del particolare disegno di campionamento (cfr. par. 2.2.1), che prevede un coinvolgimento differenziato dei comuni campione nell'arco dei mesi dell'anno. Mediamente, nel 2012, sono stati operativi circa 188 rilevatori al mese mentre, nel 2013, sono stati 157 e, nel 2014, sono stati 153.

Le sessioni formative, articolate in moduli che coprono tre giornate, sono state fondamentalmente di due tipi: eventi inaugurali, rivolti ad ampi gruppi di rilevatori (tra gli 80 e i 100 partecipanti) per costituire la rete iniziale e per garantire, nel passaggio al nuovo anno di indagine, la copertura territoriale in funzione della numerosità campionaria dei singoli comuni interessati; eventi periodici, rivolti a gruppi meno numerosi (tra i 25 e i 40 partecipanti) per far fronte a necessità di sostituzioni o integrazioni della rete operante nel corso dell'anno. Tutte le sessioni formative si sono svolte convogliando i rilevatori in un'unica sede, assicurando in tal modo una centralizzazione dell'azione erogatrice.

Appare perciò chiaro come la formazione abbia un notevole impatto sulle attività da espletare nell'intero processo d'indagine, richiedendo dispendio di energie e tempo. Tuttavia, l'investimento è finalizzato a prevenire una molteplicità di potenziali errori a carico degli intervistatori (mancato rispetto delle metodologie, codifiche errate, eccessivo ricorso alle sostituzioni e così via). Inoltre, non essendo la rete direttamente gestita dall'Istituto, le sessioni formative diventano anche un'insostituibile occasione per conoscere di persona i rilevatori, interagendo con loro ed accertando, per quanto possibile, il livello generale di professionalità.

Il fatto di doversi rivolgere ad intervistatori professionali e non più a referenti regionali o comunali ha determinato fin dall'inizio la necessità di progettare moduli formativi che abbandonassero la vecchia veste di riunioni di istruzione (caratterizzate da un approccio prevalentemente contenutistico), per assumerne una nuova più rivolta allo sviluppo di competenze e strategie personali, nel rispetto delle procedure e dei vincoli metodologici. Questo spostamento di prospettiva si è accompagnato ad un'esigenza di armonizzazione dell'attività formativa fornita dall'Istat in riferimento a diverse indagini sulle famiglie condotte con tecnica CA (Istat, 2014⁸), con la conse-

6 Questo numero tiene conto unicamente dei rilevatori che, dopo aver ricevuto la formazione iniziale, hanno effettuato una qualche azione sulle quartine/famiglie loro assegnate, producendo un esito familiare definitivo (di successo o insuccesso).

7 Il tasso di uscita (o tasso di *turn-over* negativo) è calcolato come il rapporto percentuale tra il numero di intervistatori che hanno cessato l'attività a conclusione del 2012 e il numero di intervistatori attivi durante lo stesso anno di indagine.

8 Bali N., Federici A., 2014, op. cit..

guente introduzione, anche per l'Indagine sulle spese, di elementi comuni quali: i) esercitazioni tecniche basate su simulazioni di interviste complete⁹; ii) programma delle giornate che ripercorre in sequenza temporale le diverse fasi/operazioni della rilevazione; iii) tecniche di somministrazione e strategie comunicative trattate con lezioni frontali e *role playing*; iv) patto d'aula iniziale¹⁰; v) schede di valutazione del corso e di autovalutazione dell'apprendimento.

L'elevata ricorrenza delle sessioni ha consentito, nel corso del 2013, un ulteriore affinamento della proposta formativa, anche sulla base delle valutazioni e dei suggerimenti forniti dai rilevatori stessi al termine delle giornate. È stata, dunque, attuata una riprogettazione attraverso tre passaggi: i) individuazione dei punti di forza e degli elementi di miglioramento; ii) predisposizione e condivisione di un programma analitico delle giornate; iii) revisione del materiale formativo e sperimentazione del nuovo programma.

Le principali criticità riscontrate hanno riguardato: ritmi eccessivamente intensi; mancanza di un'accoglienza iniziale strutturata; numero di slides elevato e con poche animazioni o immagini; discontinuità e parziale ridondanza delle istruzioni sugli aspetti tecnici del CAPI. Lo strumento sul quale si sono confrontati tutti i referenti tematici chiamati a curare i singoli moduli formativi è stato il programma analitico, stilato in modo da esplicitare per ciascun modulo: i) obiettivi di apprendimento specifici; ii) contenuti previsti; iii) intervallo orario; iv) materiali da utilizzare; v) responsabile della conduzione (Istat, società o entrambe). Nel Prospetto 2.3 è riportata una versione semplificata di questo strumento, con l'assetto dei moduli attualmente in uso.

La proposta formativa così riformulata è stata accolta positivamente dai rilevatori, che hanno fornito giudizi migliori, mostrando meno affaticamento e maggiore partecipazione (con la precedente versione, il 76% dei rilevatori ha espresso parere positivo sul fatto di aver ricevuto gli strumenti necessari per svolgere al meglio le interviste, mentre, con il nuovo programma, ad assegnare valutazioni globali elevate è stato più dell'82%). La tempistica delle giornate è stata ottimizzata, riuscendo a contenere i tempi senza sacrificare i contenuti da esporre.

Ripensare il programma in termini di obiettivi ha consentito una ricollocazione più equilibrata dei moduli. Inoltre, è stata introdotta la modalità "a doppio binario" tra teoria e pratica, alternando la descrizione di concetti e definizioni e la sperimentazione pratica di specifiche sezioni degli strumenti d'indagine. Le slides sono state notevolmente ridotte di numero e ripensate in funzione di questa modalità, inserendo stralci di telegiornali, immagini e altri video appositamente realizzati (ad es. su tecniche di riformulazione dei quesiti, rilanci, criteri di prevalenza, ecc.).

9 Gli intervistatori leggono a turno i quesiti, immaginando di avere di fronte una famiglia. Si impratichiscono con il Capi e apprendono man mano le modalità di somministrazione.

10 Il cosiddetto patto d'aula serve a rendere espliciti e a condividere i ruoli e i comportamenti che formatori e partecipanti si impegnano ad interpretare e a rispettare durante le giornate. Se ben condotto, questo momento consente di instaurare un rapporto comunicativo e collaborativo utile al raggiungimento del comune obiettivo di apprendimento.

Prospetto 2.3 – Programma per la formazione dei rilevatori

MODULI	OBIETTIVI	MATERIALI
<i>Prima mezza giornata</i>		
Apertura	Condividere gli obiettivi delle giornate. Anticipare regole (rispetto orari, domande e risposte). Indagare aspettative. Creare un clima collaborativo. (30 min)	slides, lavagna
Presentazione	Illustrare gli obiettivi conoscitivi e le finalità dell'indagine. Introdurre i principali concetti e classificazioni. (45 min)	slides, video tg
Le fasi	Fornire il quadro completo delle fasi della rilevazione. Introdurre gli strumenti. Descrivere l'apparato organizzativo della rilevazione. (45 min)	slides
Gli strumenti informatici	Fornire indicazioni per l'accesso all'area test. Descrivere le agende con i nominativi delle famiglie (sia per il Capi sia per il diario). (75 min)	manuale tecnico
Linee guida per la somministrazione	Richiamare generalità su Sezioni, script, filtri, controlli di compatibilità (<i>hard e soft</i>) ed Help in linea del questionario elettronico. Illustrare regole generali per la lettura dei quesiti e delle modalità di risposta e per la codifica. (45 min)	slides, video
<i>Seconda giornata</i>		
Scheda contatti (a)	Descrivere la guida strutturata per la gestione dei contatti con le famiglie. Mostrare il funzionamento del software. Inaugurare l'apertura della famiglia con cui si proseguirà l'esercitazione nei giorni seguenti fino ad intervista completa. (90 min)	slides, protocollo esempi, netbook
Scheda generale (a)	Descrivere scheda generale e sistema di classificazione delle professioni. Iniziare l'inserimento delle informazioni sulla famiglia fittizia oggetto dell'esercitazione e utilizzare il navigatore. (120 min)	slides, netbook
Intervista iniziale (a)	Descrivere i contenuti del questionario per l'intervista iniziale e compilare le sezioni continuando a simulare l'intervista con la famiglia fittizia. (60 min)	slides, netbook
Il diario (a)	Descrivere il diario. Fornire istruzioni sulla consegna e compilazione. Illustrare il controllo di errori di compilazione su diari precompilati, percorrendo le schermate delle apposite sezioni della scheda contatti. (150 min)	slides, diari, netbook
<i>Terza giornata</i>		
Intervista finale (a)	Descrivere i contenuti del questionario, i periodi di riferimento delle spese, i controlli di compatibilità con l'intervista iniziale. Compilare il questionario proseguendo la simulazione di intervista. (240 min)	slides, netbook
Data entry diario	Istruire all'utilizzo del software per l'imputazione. Descrivere i controlli. (45 min)	diari, netbook
Trasmissioni	Spiegare come effettuare le trasmissioni per l'invio telematico dei dati di intervista, del diario ed esiti di contatto e la ricezione dei nominativi assegnati. (60 min)	manuale tecnico, video
Indicatori di performance	Informare circa il sistema di monitoraggio del lavoro sul campo. Illustrare e descrivere i principali indicatori e i risultati della rilevazione in corso. (15 min)	slides
Incontro con un rilevatore esperto	Offrire uno spazio di confronto per anticipare criticità della rilevazione e condividere strategie nei contatti con le famiglie. (30 min)	dibattito
La comunicazione efficace	Indicare alcune regole di comunicazione importanti nei primi approcci con le famiglie. Simulare a coppie primi contatti per ottenere la partecipazione. (150 min)	slides, <i>role playing</i>
<i>Quarta mezza giornata</i>		
Focus Viaggi (a)	Presentare il focus sui viaggi e le vacanze. Esaminare la struttura di questa parte dell'intervista finale e introdurre le definizioni operative di viaggi ed escursioni. Riprendere l'intervista finale lasciata in sospeso, portando a termine la simulazione con la famiglia fittizia. (180 min)	slides, netbook
Schede di autoverifica	Mettere alla prova l'apprendimento dei principali concetti esposti durante le giornate. Riassumere brevemente le fasi della rilevazione e le operazioni da attuare. Familiarizzare con il manuale Istat "Istruzioni per il rilevatore". (60 min)	schede di autoverifica, manuale Istat
Chiusura	Schede di valutazione del corso. Saluti e appuntamento ai <i>debriefing</i> .	schede di valutazione

(a) Per questi moduli formativi è stata seguita la modalità "a doppio binario", con brevi introduzioni teoriche, seguite da esercitazioni tecniche; le esercitazioni, inoltre, vengono interrotte in corrispondenza di snodi significativi degli strumenti d'indagine, per riprendere, con slides riassuntive, i contenuti via via esposti.

L'inserimento di uno specifico modulo tecnico gestito dalla società mira a contenere le ridondanze e i rallentamenti che si producevano durante le esercitazioni e va a completare la gamma di nozioni introduttive collocate in prima giornata. Le schede di autovalutazione dell'apprendimento, che precedentemente venivano ritirate dal personale Istat a fine compilazione e utilizzate per tarare meglio gli interventi formativi, attualmente hanno assunto un ruolo prevalentemente a sostegno del consolidamento delle conoscenze e vengono proposte sotto forma di gioco in aula a conclusione del corso e lasciate a disposizione dei partecipanti. Le domande per l'autovalutazione contenute nelle schede consentono di ripercorrere gli argomenti principali, esposti durante tutte le giornate di formazione, e forniscono l'occasione per illustrare il manuale Istat "Istruzioni per il rilevatore", che costituirà in seguito il principale strumento di consultazione per il corretto svolgimento del lavoro sul campo.

La fase del programma di formazione maggiormente focalizzata sui rilevatori, infine, occupa l'ultima parte della terza giornata ed è introdotta da una descrizione degli indicatori di performance della rete, riferiti alla rilevazione in corso, attraverso cartografie regionali.

Il modulo sulla comunicazione e il contenimento dei rifiuti

Un ulteriore fattore di forza, nell'ottica dei mutamenti organizzativi legati all'utilizzo delle tecniche *Computer Assisted*, è dato dall'opportunità di realizzare una formazione mirata ad aumentare le competenze professionali per raccogliere informazioni di qualità. A questo proposito, gli intervistatori non solo devono conoscere i contenuti dell'indagine e gli aspetti normativi connessi alla rilevazione, ma anche sapere gestire le strategie organizzative e le relazioni interpersonali nei rapporti con le famiglie.

Nel programma formativo è stato quindi inserito un modulo più propriamente dedicato al processo comunicativo per la conquista della famiglia, ovvero alla gestione e al contenimento dei rifiuti orientato a:

- allenare gli intervistatori ad una strategia comunicativa diretta alla motivazione dei rispondenti attraverso modalità linguistiche e gestuali;
- favorire capacità di persuasione e convincimento per far fronte alle diverse aspettative e richieste degli intervistati;
- fornire gli elementi per stabilire un buon clima comunicativo con le famiglie, per essere disponibili e interessati, senza scadere nella confidenza eccessiva.

Potenziare le capacità comunicative aiuta inoltre gli intervistatori a far fronte alle possibili interruzioni delle interviste già avviate, a saper gestire la stanchezza dell'intervistato e ad essere preparati psicologicamente a eventuali rifiuti che potrebbero essere percepiti come sconfitte personali.

La progettazione di questo modulo tiene conto di un opportuno bilanciamento tra richiami teorici e contenuti pratici. Dopo una presentazione teorica iniziale delle principali caratteristiche della comunicazione e degli stili volti a migliorare le capacità comunicative, vengono proposte esercitazioni, *role playing* ed esemplificazioni per contenere i rifiuti. I giochi di ruolo hanno il vantaggio di essere particolarmente efficaci per condividere le esperienze vissute, poiché danno la possibilità ai rilevatori di recitare quei ruoli - intervistato e/o intervistatore - che incontrano e incarnano nella realtà.

Il *role playing* è suddiviso in tre fasi: la preparazione, il gioco di ruolo e la discussione finale. Nella fase di preparazione, l'animatore descrive lo scenario, spiega all'aula le operazioni che occorre compiere, invita i partecipanti a giocare e chiede loro di scegliere quale ruolo intendono rappresentare. Designati i ruoli, l'animatore assegna solo all'intervistato un copione¹¹, mentre all'intervistatore è assegnato il compito per ascoltare attentamente, per individuare eventuali difficoltà manifestate dall'intervistato, esortare e far argomentare le motivazioni di una eventuale resistenza, adattando il linguaggio e i contenuti all'interlocutore, e motivare infine

¹¹ Il copione è stato strutturato per rappresentare quelle tipologie familiari più resistenti a concedere l'intervista. Alcuni esempi: "sei un'anziana signora che teme che le si voglia vendere qualcosa, non conosce l'Istat e rifiuta l'intervista"; "sei un giovane professionista, molto impegnato, che tratta l'intervistatore con sufficienza e fretolosità affermando che queste interviste sono inutili, che è contrario alle interviste e non ha tempo da perdere".

l'interlocutore alla partecipazione. Segue quindi la fase del gioco di ruolo vero e proprio, in cui i partecipanti hanno l'opportunità di sperimentare le situazioni sul campo da un punto di vista diverso rispetto a quanto farebbero in realtà. La terza fase, finalizzata alla discussione di quanto avvenuto, aiuta i partecipanti, ma anche il resto dell'aula, a comprendere meglio le dinamiche interpersonali che avvengono sul campo, a considerare i comportamenti più idonei per ottenere l'intervista, a valutare il clima di fiducia e il rapporto comunicativo che si instaura tra intervistatore e intervistato e a verificare se sono stati utilizzati tutti gli strumenti di sensibilizzazione¹², messi a loro disposizione, per facilitare il contatto con le famiglie. È così possibile valutare il grado di acquisizione dei contenuti dell'indagine, la capacità metodologica di somministrare l'intervista, le strategie organizzative per la pianificazione del lavoro, l'approccio con le famiglie e le tecniche per gestire i rifiuti.

Inoltre per garantire la qualità dell'interazione tra intervistato e intervistatore, nel corso dell'intervento formativo viene posta grande attenzione al modo con cui approcciarsi con le famiglie, al fine di padroneggiare al meglio le tecniche per gestire i rifiuti. In particolare, viene presentata una lista di alcuni dei principali motivi per cui le famiglie potrebbero rifiutare, accompagnata da suggerimenti e atteggiamenti pratici per arginarli. Data l'importanza di contenere le mancate risposte totali per garantire la qualità delle informazioni raccolte, a ciascuna tipologia di rifiuto segue un'ampia discussione in cui i partecipanti sono parte attiva, riportando la loro esperienza sul campo e ampliando, in questo modo, le strategie di approccio con la famiglia più adeguate ed efficaci a massimizzare la disponibilità dei rispondenti e l'attendibilità delle loro risposte.

Nel Prospetto 2.4 vengono schematizzate le principali tipologie di motivi che potrebbero indurre le famiglie a rifiutare e i relativi comportamenti suggeriti.

Al fine di valorizzare le esperienze positive e le buone prassi realizzate dagli intervistatori sul campo, è stata proposta, inoltre, un'unità didattica esperienziale attraverso il coinvolgimento di un intervistatore professionista, in grado di esporre in maniera affidabile le diverse esperienze maturate nel corso del suo lavoro e le accortezze impiegate per motivare le famiglie alla collaborazione.

La presenza dell'intervistatore esperto in aula presenta il vantaggio di stabilire una comunicazione tra pari, che risulta essere credibile e ben accolta da tutti; allo stesso tempo, è importante porre attenzione che non siano veicolati messaggi difformi dalle indicazioni metodologiche impartite dall'Istat.

L'intervistatore esperto garantisce ai nuovi rilevatori:

- il passaggio di informazioni, relative al lavoro e al contesto, normalmente inaccessibili;
- il trasferimento di abilità complesse, come il saper gestire momenti critici (ad esempio quando si contatta la famiglia al citofono ricordando il numero verde o la lettera del presidente);
- i richiami metodologici essenziali, come l'astenersi durante l'intervista da qualsiasi commento o dall'effettuare altre attività.

In questo senso, il confronto diretto con un rilevatore esperto riscuote un notevole successo tra i partecipanti, poiché attiva e mantiene desta l'attenzione per

¹² L'Istat da sempre è attivo nell'incentivare la partecipazione delle famiglie alle indagini supportando il rilevatore con due strumenti: la lettera alle famiglie e il numero verde.

tutta la durata del modulo, rendendoli consapevoli dell'importanza del ruolo di intervistatore, che si riveste all'interno del processo d'indagine.

Prospetto 2.4 – Motivi di rifiuto e comportamenti suggeriti

MOTIVI	ESEMPI	COMPORTEMENTI SUGGERITI
<i>Rifiuti che non dipendono dalla volontà del rispondente</i>		
Disturbo, problemi familiari.	Impossibilità a rispondere per: preparazione dei pasti, cura di malati, presenza di un operaio in casa, ecc.	Si consiglia di fissare un appuntamento compatibile con le esigenze della famiglia. È importante stabilire un contatto con il rispondente, un patto comunicativo.
Malattia	La famiglia non è in buone condizioni di salute.	Nel caso di malattia lieve, si può prendere un appuntamento compatibile con le esigenze del rispondente; se si tratta, invece, di una invalidità o di una malattia più grave, si può procedere con un'intervista <i>proxi</i> ad un altro componente della famiglia o altro familiare.
<i>Rifiuti causati dal timore di perdere tempo</i>		
Assenza di interesse, mancanza di tempo	L'intervistato non ha tempo da perdere.	In questi casi è possibile far leva sul senso di appartenenza ad un gruppo sociale significativo, composto da persone molto impegnate, senza il cui apporto una parte di popolazione rimarrebbe sconosciuta. È consigliabile non mentire sulla durata dell'intervista, semmai è possibile ovviare alla lunghezza del questionario sottolineando la semplicità delle risposte e la scorrevolezza delle domande.
	Il questionario è troppo lungo.	
<i>Rifiuti causati da diffidenza o ostilità</i>		
Sfiducia, opinione negativa sulle indagini.	L'intervistato è contrario alle interviste. Ha paura che l'intervista non sia fatta da un'istituzione legittima, ma da privati. Ha sfiducia nello Stato.	In questi casi è possibile rilanciare sull'importanza delle indagini realizzate dall'Istat e sull'utilità per le politiche economiche e sociali del Paese. È importante non polemizzare, rassicurare sulla riservatezza delle informazioni rilasciate. Le domande riguardano problematiche di rilevanza sociale utili per approntare politiche per il benessere collettivo.
Diffidenza, paura.	Timore che i propri dati possano essere utilizzati per scopi commerciali.	Rassicurare la famiglia ricordando tutti gli strumenti di tutela e protezione dei dati personali impiegati dall'Istat, nonché di sensibilizzazione sull'indagine, come il numero verde, la lettera a firma del presidente, ecc.

I debriefing

Nell'ambito della contaminazione tra le diverse fasi d'indagine, di cui si è detto nel paragrafo introduttivo (cfr. par. 2.3.3.1), le riunioni di *debriefing* rappresentano il punto di intersezione privilegiato tra la formazione, il monitoraggio e la valutazione del disegno di ricerca. Il confronto diretto che tali riunioni consentono di svolgere con i rilevatori favorisce il raggiungimento di almeno tre obiettivi: ricevere dalla rete un *feedback* approfondito sull'andamento della rilevazione, valutare l'adeguatezza del questionario e la comprensibilità delle domande, aggiornare la formazione dei rilevatori esperti.

L'esperienza condotta nell'ambito dell'Indagine sulle spese delle famiglie dimostra come, al di là delle modalità di confronto di volta in volta sperimentate con i rilevatori, i *debriefing* consentono di arricchire significativamente la conoscenza dei processi di indagine e di affinare la preparazione degli intervistatori, individuando i punti critici non emersi in sede di formazione iniziale.

Lo svolgimento dei *debriefing* è stato reso possibile, fino ad ora, sia attraverso incontri *face-to-face* con sottogruppi di rilevatori sia attraverso la realizzazione di attività

formative e di valutazione dell'indagine "a distanza". Nella prima modalità organizzativa, si favorisce il confronto diretto e approfondito con la rete; nella seconda, invece, si offre ai rilevatori un'occasione di riflessione autonoma e meditata sull'andamento della rilevazione, oltre a ottenere una sensibile riduzione dei costi dell'operazione.

Gli input raccolti dalla rete, nelle occasioni di *debriefing*, sono molteplici e di diversa natura. Essi si affiancano alle informazioni fornite dal sistema degli indicatori e ne danno conto, approfondendo le dinamiche che avvengono sul campo e individuando le cause delle eventuali anomalie denunciate dal monitoraggio¹³. Gli intervistatori hanno il polso della situazione nei rispettivi territori e sono in grado di suggerire prassi adeguate per invogliare i rispondenti a collaborare. Allo stesso tempo, registrano dal vivo le reazioni delle famiglie al questionario, segnalando le criticità nella conduzione delle interviste e le domande che risultano meno chiare ai rispondenti. Più in generale, la rete è un osservatorio privilegiato di tutte le dinamiche, che scaturiscono nella fase di lavoro sul campo, ponendosi come crocevia tra la società, i comuni e le famiglie. Da questa posizione è quindi in grado sia di descrivere le possibili inefficienze dei processi sia di dare suggerimenti, che possono essere adottati dall'Istituto per migliorare la rilevazione o rendere più fluida la somministrazione. La scheda impiegata nell'ultima occasione di *debriefing* (Prospetto 2.5) – gestita "a distanza" – dimostra la ricchezza di contenuti e la varietà di aspetti che è possibile affrontare con la rete.

Se opportunamente valorizzate, le informazioni che si raccolgono in queste circostanze possono fornire indicazioni preziose per migliorare il disegno di indagine. Di recente, ad esempio, è stato riprogettato l'opuscolo informativo che viene lasciato in omaggio alle famiglie intervistate, per illustrare i principali risultati dell'indagine. Le continue segnalazioni della rete, infatti, che riportavano l'esigenza di far comprendere in modo più semplice e intuitivo il flusso dei dati, dai rispondenti agli utilizzatori finali, ha portato all'introduzione di nuovi contenuti e alla realizzazione di un prodotto più facilmente accessibile e accattivante.

In alcuni casi, il riscontro ottenuto nel confronto con la rete consente di apportare soluzioni tempestive anche in corso di rilevazione, scongiurando il rischio di propagare possibili cause di errori non campionari.

I report, redatti abitualmente al termine dei *debriefing*, riassumono le principali risultanze emerse dal feed back con gli intervistatori ed indicano, per ciascuna di queste, i margini di modifica e di miglioramento ipotizzabili nell'economia generale dell'indagine. In occasione del *debriefing* svolto nel 2014, ad esempio, che ha registrato il coinvolgimento di ben 182 rilevatori, sono emerse delle indicazioni preziose sulle domande del questionario più impegnative, che rendevano l'intervista meno fluida. Le indicazioni ricevute hanno suggerito sia una diversa modulazione degli interventi formativi, intensificando i momenti esercitativi su alcuni argomenti, sia una formulazione più asciutta e semplificata di alcune batterie di domande. Le valutazioni sul processo d'indagine, invece, hanno evidenziato sia una soddisfacente collaborazione da parte dei comuni (che in una scala da 1 a 10 riscuotono un punteggio medio di disponibilità a collaborare pari a 7,7), sia una disponibilità crescente delle famiglie nel corso delle varie fasi dell'intervista (il valore medio passa infatti da 6,2, registrato

¹³ Ad esempio, su indicazione dei rilevatori, sono stati opportunamente prolungati i periodi di contatto con le famiglie in occasione delle festività e dei periodi di ferie, per ridurre l'incremento delle cadute per "fuori periodo" segnalate dal sistema degli indicatori.

in occasione della prima visita, a 6,4 durante la visita intermedia, fino a raggiungere 7,2 nell'intervista finale).

Prospetto 2.5 - Estratto della scheda inviata alla rete in occasione del debriefing 2014

Riferendoti alla tua recente esperienza di indagine (Spese 2013), ti chiediamo di esprimere una VALUTAZIONE dei seguenti aspetti, assegnando un punteggio da 1 a 10		SCRIVI EVENTUALI SUGGERIMENTI	
Disponibilità dei comuni a collaborare			
Disponibilità delle famiglie a partecipare all'indagine			
Fluidità nella somministrazione del questionario			
Utilità degli help in linea			
Modalità di registrazione del diario (software, regole, ...)			
Tempistica e modalità di trasmissione di diari e nominativi			
Tempistica e modalità di invio dei cartacei alla Società			
Funzionamento del software			
	Prima visita (intervista iniziale e consegna diario)	Visita intermedia (controllo diario)	Terza visita (intervista finale e recupero diario)
Considerando le tre fasi di rilevazione, come valuteresti, in una scala da 1 a 10, la DISPONIBILITA' delle famiglie a collaborare nella prima, seconda e terza visita?			
Considerando la versione 2013 dei questionari, ci sono delle domande in particolare che rendono l'intervista meno fluida e/o più problematica? Se sì, ti invitiamo a indicare il numero del/dei quesito/i e il motivo			
	Indica il numero del/dei quesito/i	Indica il motivo	
QUESTIONARIO PER INTERVISTA INIZIALE			
Scheda generale e codifica professioni (Sez. SG e 6)			
Caratteristiche abitazione e altre abit. di proprietà (Sez.2)			
Mezzi di trasporto e comunicazione (Sez. 3)			
QUESTIONARIO PER INTERVISTA FINALE			
Abitudini di spesa (Sez. A)			
Spese per l'abitazione in cui vive la famiglia e spese per altre abitazioni (Sez. B e C)			
Spese ultimi 12 mesi - elettrodomestici, telefonia e altri beni e servizi (Sez. D)			
Spese ultimi 3 mesi - arredamento, collaboratori domestici e altro personale, viaggi organizzati e servizi di pernottamento (Sez. E)			
Spese ultimo mese - mense aziendali, acquisti a rate (Sez. F)			
Focus Viaggi e vacanze (Sez. da RV a M)			

Il ritorno alla rete dei report conclusivi, corredati delle strategie di miglioramento del processo d'indagine ritenute condivisibili e praticabili, alimenta la circolarità della comunicazione e attesta la corresponsabilità di tutti i protagonisti della rilevazione nel miglioramento della qualità dei dati raccolti.

2.3.3.3 Il monitoraggio della qualità dell'indagine

Lottica delle attività di monitoraggio della rilevazione

Con l'obiettivo di tenere sotto controllo il processo e di perseguire un miglioramento della qualità, in ogni indagine statistica viene predisposta una serie di strumenti e di azioni. Strumenti ed azioni, diversi a seconda della tecnica, della metodologia e delle finalità che si vogliono perseguire e che costituiscono tutti insieme il monitoraggio di un'indagine.

Per l'Indagine sulle spese, finché ci si è avvalsi dei questionari cartacei e della collaborazione dei Comuni campione (2013), il monitoraggio era organizzato prevedendo una serie di attività a posteriori, in ciascuna delle fasi di lavorazione successive al rientro all'Istat dei modelli cartacei: controllo quantitativo dei modelli rientrati, revisione manuale qualitativa, controlli a campione sul *data entry*. Successivamente, se necessari, venivano predisposti interventi con l'obiettivo di correggere quelle attività che si discostavano da quanto programmato. Tali interventi risultavano scarsamente tempestivi, per cui il lasso di tempo che intercorreva tra l'individuazione di una eventuale anomalia o criticità e la conclusione dell'indagine rendeva inefficace, se non impossibile, qualsiasi contromisura nell'ambito di un ciclo mensile di rilevazione.

Con il passaggio dell'indagine a una tecnica *Computer Assisted* (CAPI), molti dei limiti delle precedenti attività di monitoraggio sono stati superati, mentre altre criticità sono emerse. È importante evidenziare che un impulso positivo al nuovo sistema di monitoraggio è stato ottenuto anche grazie all'introduzione di un approccio trasversale. La singola indagine non è stata più considerata come un processo autonomo e indipendente e uno degli obiettivi da raggiungere da parte dell'Istituto ha riguardato la costruzione di un sistema di monitoraggio complesso. Infatti, alla maggiore coesione tra le fasi di una stessa indagine, si è aggiunta una più stretta relazione tra le diverse indagini sulle famiglie condotte dall'Istat.

La presenza, contestualmente alla somministrazione dell'intervista, di un questionario elettronico implementato sulla base delle regole di compatibilità e coerenza ha consentito di ridurre sostanzialmente la possibilità di raccogliere informazioni incoerenti e/o incompatibili e di rilevare, sempre più, il dato corretto direttamente dal rispondente. Grazie poi alla trasmissione, pressoché contestuale all'intervista, non solo dei dati raccolti con i questionari, ma anche di una notevole quantità di informazioni aggiuntive, il monitoraggio si è evoluto da supporto alla rilevazione a sua parte integrante.

L'ausilio del personal computer ha reso possibile costruire un vero e proprio sistema di indicatori in grado di monitorare costantemente le evoluzioni della situazione sul campo, misurando altresì gli eventuali scostamenti rispetto a quanto programmato. L'aver poi a disposizione, in maniera integrata, informazioni sulle attività di tutte le fasi del processo, ha reso possibile tener conto delle loro interdipendenze, evitando così che azioni a supporto di talune attività potessero impattare negativamente sulle altre.

Di certo, l'aver accentrato ed armonizzato tutte le attività rientranti nel monitoraggio ha rappresentato un importante miglioramento per l'intero processo; pur tuttavia il passaggio a tecnica *Computer Assisted*, avvenuto in concomitanza con il trasferimento della raccolta dati a soggetti esterni privati, ha fatto emergere nuove criticità che hanno impattato, limitandone il potenziale, anche sul nuovo sistema di monitoraggio.

Il sistema di monitoraggio implementato per l'indagine in oggetto si è concretizzato nella costruzione di una serie di indicatori diretti (calcolati sul file degli esiti dei contatti con le famiglie campione) ed indiretti (calcolati sul file dei dati di indagine) e nell'attività di controllo, attraverso ricontatti telefonici alle famiglie intervistate, al fine di verificare la veridicità delle informazioni rilevate.

Gli indicatori di monitoraggio diretti

Gli indicatori diretti vengono implementati grazie all'attivazione di uno specifico set di regole, sottese alla scheda contatti (cfr. par. 2.3.2.1), grazie alle quali vengono assegnati gli esiti (provvisori o definitivi) degli appuntamenti e gli stati delle interviste (da contattare, tentativi in corso, fissato appuntamento, intervista iniziata). Gli standard utilizzati per la definizione degli esiti seguono le raccomandazioni¹⁴ già ampiamente sperimentate per altre indagini Istat sulle famiglie con tecnica *Computer Assisted* (Istat, 2006b), con l'introduzione di opportune modifiche tese a rendere conto delle specificità metodologiche dell'indagine.

La società elabora e fornisce all'Istituto i seguenti prodotti:

1. rapporti giornalieri sull'andamento dell'attività di rilevazione;
2. file degli esiti definitivi e dei tentativi di contatto.

La reportistica giornaliera viene messa a disposizione dei referenti Istat su un portale web con accesso dedicato. L'Istituto fornisce alla società le specifiche per la costruzione degli indicatori e per la realizzazione delle tavole statistiche da caricare sul portale. A tal proposito, il set completo di indicatori viene generato per ciascun ciclo mensile di rilevazione e replicato con aggregazioni trimestrali e annuali. Inoltre, le tavole sono costruite per l'intero territorio nazionale e a livello regionale; esse contemplano, come ulteriori variabili di incrocio, l'appartenenza al primo o al secondo gruppo del mese di riferimento e, in alcuni casi, la posizione delle famiglie nella quartina. Infine, è sempre presente il dettaglio a livello di singolo rilevatore, per un'analisi puntuale delle performance individuali. Nel corso dell'attività sul campo, i rilevatori sono invitati ad effettuare le trasmissioni tutti i giorni, determinando così il progressivo popolamento del portale. Esso, oltre ad essere il principale strumento di controllo della rilevazione in atto, svolge anche una funzione di *repository* delle tavole degli indicatori riferiti ai diversi cicli mensili di rilevazione, le quali possono continuare ad essere consultate ed esportate per tutto l'anno di indagine e anche oltre.

I file degli esiti definitivi e dei tentativi di contatto, invece, vengono consegnati al termine di ogni ciclo mensile di rilevazione. Entrambi questi file vengono utilizzati per diversi scopi, tra cui la verifica dell'esattezza degli indicatori di monitoraggio pubblicati sul portale o la costruzione di ulteriori indicatori. Le elaborazioni sul file degli esiti definitivi permettono anche di mettere in relazione i diversi tipi di caduta con i contesti territoriali e le caratteristiche familiari, al fine di individuare le principali criticità e di pianificare strategie di miglioramento. I file dei tentativi permettono, invece, di ricostruire la storia dei contatti con una famiglia e consentono di controllare, fra l'altro, il corretto funzionamento dei contattori per la gestione dei contatti.

¹⁴ Nella progettazione degli indicatori di monitoraggio nelle indagini Istat si fa riferimento agli standard nazionali (al sistema Sidi, strumento generalizzato per il controllo della qualità delle indagini Istat) ed internazionali (Aapor, ovvero American association for public opinion research).

Nel loro insieme, le informazioni di monitoraggio forniscono una misura di alcuni rilevanti errori, quali gli errori di lista e quelli di mancata risposta totale e del ruolo svolto in proposito dalla società e dalla rete dei rilevatori. In quest'ottica, ai consueti indicatori di qualità, cui ci si riferisce normalmente per i confronti nazionali ed internazionali¹⁵, sono stati affiancati ulteriori indicatori specifici, con riferimento sia al campione base sia a quello delle famiglie sostitutive (quartine), descritti nel seguito.

Complessivamente, vengono prodotte due classi di indicatori diretti:

- gli indicatori di performance della rete di rilevazione, attraverso l'analisi dell'utilizzo delle quartine;
- gli indicatori sugli esiti familiari che, analizzando le cadute per motivo, danno conto delle difficoltà incontrate dai rilevatori nella gestione delle famiglie.

Indicatori di performance della rete

- *Il tasso di assegnazione*: è il rapporto percentuale tra il numero totale di quartine assegnate ai rilevatori e le quartine teoriche. La società, infatti, distribuisce i carichi di lavoro alla rete e deve garantire la copertura del territorio. Tuttavia una parte delle quartine rimane mensilmente non assegnata, a causa di eventi strutturali o congiunturali connessi con l'organizzazione della rete di rilevazione, generando caduta campionaria. La mancata assegnazione per ogni comune di rilevazione e per ogni mese può essere parziale o totale con evidenti differenze di impatto in termini di errore campionario.
- *Il tasso di pigrizia*: una volta assegnati i carichi di lavoro, i rilevatori iniziano la loro attività. Le quartine non toccate (per le quali cioè il rilevatore non effettua alcuna azione) rapportate a quelle assegnate forniscono il tasso di pigrizia, che rappresenta un indicatore della difficoltà dei rilevatori a smaltire i nominativi nella propria agenda.
- *Il tasso di completezza*: il rapporto percentuale tra le quartine con intervista completa¹⁶ e le quartine assegnate esprime la distanza che c'è tra il campione effettivo e quello potenzialmente lavorabile. Il tasso di completezza viene calcolato anche per singola fase di intervista (interviste iniziali, visite intermedie e interviste finali), rapportando le quartine con intervista completa – Quartine con INI completa, Quartine con VI completa e Quartine con IFIN completa - alle sole quartine che hanno portato a termine la fase immediatamente precedente.
- *La percentuale di quartine bucate*: ad integrazione dei tassi di completezza, è utile osservare la mole di quartine bucate, in cui cioè tutte e quattro le famiglie cadono. Anche questo indicatore viene calcolato per ciascuna fase di intervista, rapportando le quartine bucate in intervista iniziale (Fase 1), in visita intermedia (Fase 2) e in intervista finale (Fase 3), al totale di quartine toccate. In questo modo, si cerca di apprezzare la capacità dei rilevatori di ottenere la collaborazione delle famiglie dall'inizio alla fine del periodo di raccolta dati, depurando il denominatore dalle quartine che non sono riusciti a lavorare, presumibilmente per eccessivo carico di lavoro.
- *Il tasso di fedeltà*: una componente della distorsione campionaria è riconducibile alla sostituzione delle famiglie base con le altre famiglie della quartina. Queste ultime, infatti, seppure selezionate in base a criteri di similitudine con la famiglia base (in termini di ampiezza familiare e zona di residenza), possono differire

¹⁵ Esiste una gamma consolidata di indicatori di qualità standard, tra cui il tasso di risposta, il tasso di cooperazione e il tasso di rifiuto, per i quali si rimanda alla letteratura (Aapor, 2000).

¹⁶ Si ricorda che una quartina può generare una e una sola intervista completa.

anche sostanzialmente dalle famiglie che vanno a sostituire, per caratteristiche anch'esse correlate al livello e alla scelta di allocazione della spesa. Il particolare disegno di rilevazione obbliga costantemente i rilevatori a scegliere tra il perseverare nei tentativi di contatto con le famiglie base, rischiando di non avere più i margini di tempo utili ad effettuare le sostituzioni e causando, quindi, un incremento della mancata risposta, oppure procedere alla sostituzione, determinando però un abbassamento della qualità del dato. Il ricorso alle sostituzioni deve, dunque, essere tenuto sotto controllo attraverso il tasso di fedeltà, che rapporta le quartine in cui l'intervista completa è stata realizzata con la famiglia base al totale di quartine con intervista completa.

A scopo illustrativo, nella Tavola 2.2 si riportano i principali indicatori di performance calcolati sulle quartine¹⁷, per l'anno di indagine 2014.

Tavola 2.2 – Principali indicatori di performance della rete – Anno 2014

INDICATORI	DESCRIZIONE	TASSI
Tasso di Assegnazione	(Quartine assegnate/ Quartine teoriche)*100	81,5
Tasso di Pigrizia	(Quartine non toccate/Quartine assegnate)*100	7,4
Tasso di completezza 1	(Quartine con INI completa/ Quartine assegnate)*100	77,4
Tasso di completezza 2	(Quartine con VI completa/ Quartine con INI completa)*100	99,0
Tasso di completezza 3	(Quartine con IFIN completa/ Quartine con VI completa)*100	97,3
% Quartine bucate 1	(Quartine bucate in Fase 1/Quartine toccate)*100	9,5
% Quartine bucate 2	(Quartine bucate in Fase 2/Quartine toccate)*100	9,8
% Quartine bucate 3	(Quartine bucate in Fase 3/Quartine toccate)*100	9,8
Tasso di Fedeltà	(Quartine complete BASE/ Quartine complete TUTTE)*100	54,6

Indicatori sugli esiti familiari

- *Le interviste complete* (EDFS¹⁸ = 01): gli esiti definitivi individuano, innanzitutto, le famiglie che hanno concluso tutte e tre le fasi di rilevazione, ovvero quelle con intervista completa. Esse possono essere analizzate in funzione della posizione d'ordine occupata, nell'ambito delle quartine di appartenenza; ciò consente di esaminare in maggior dettaglio il ricorso agli elenchi sostitutivi da parte dei rilevatori, evidenziando eventuali difformità sul territorio.
- *Le cadute per interruzione definitiva* (EDFS = 02): le famiglie che inizialmente accettano di partecipare alla rilevazione, ma che successivamente interrompono la loro cooperazione, vengono esitate con appositi codici. Le 'Interruzioni definitive' propriamente dette sono quelle che si verificano durante l'effettuazione di un'intervista. Di queste, è possibile conoscere anche lo specifico quesito, in corrispondenza del quale viene registrato il rifiuto¹⁹ a continuare. Sulla base di tali informazioni, sono possibili riflessioni di carattere metodologico, sulla

¹⁷ Elaborazioni dai file degli esiti mensili.

¹⁸ EDFs, è l'acronimo di Esito Definitivo Familiare Sintetico.

¹⁹ Tra i motivi che portano la famiglia ad interrompere vi sono perlopiù impedimenti oggettivi (lutti, malattie, assenza nei giorni previsti per le visite.), ma anche stanchezza, mancanza di tempo o diffidenza. Le interruzioni definitive sono dunque assimilabili a dei rifiuti. Ecco perché il tasso di rifiuto solitamente include anche le interruzioni definitive.

costruzione dei questionari per intervista, mettendone in luce eventuali punti critici (formulazione, lunghezza, sequenza dei quesiti, etc.). In questo aggregato, perciò, rientrano tutte le cadute, che avvengono nel corso di un'interazione faccia a faccia con l'intervistatore mentre, cioè, si stanno somministrando dei quesiti; ciò può accadere durante l'intervista iniziale, durante il controllo del diario in visita intermedia, oppure mentre si svolge l'intervista finale. Tuttavia, nel caso di questa indagine, la famiglia può interrompere la rilevazione anche tra una fase e la successiva. In particolare, dopo aver completato l'intervista iniziale e dopo aver portato a termine il controllo del diario in visita intermedia. Per tale ragione, sono stati previsti altri due codici, rispettivamente 'Intervista iniziale completa' e 'Visita intermedia completa', in grado di identificare tali situazioni. In questo modo, è sempre possibile imputare tutte le cadute in itinere ad una specifica fase.

- *Le cadute per rifiuto* (EDFS = 03): tra i motivi di caduta più consistenti, vi sono i rifiuti iniziali. La famiglia contattata dichiara direttamente al rilevatore la sua volontà a non partecipare; in questi casi, si procede automaticamente con una sostituzione, qualora vi siano i margini di tempo utili all'effettuazione dell'intervista iniziale e qualora si abbiano ancora a disposizione nella quartina dei nominativi di riserva²⁰. In generale, la metodologia dell'indagine prevede che la caduta di un nominativo possa essere attestata, solo effettuando tutti i tentativi di contatto a domicilio previsti (nella fattispecie almeno tre). Perciò, anche il rifiuto deve essere registrato in occasione di una visita a domicilio. Tra i rifiuti al domicilio, si distinguono quelli manifestati prima ('Rifiuto a domicilio pre-individuazione PR') e quelli dopo ('Rifiuto a domicilio post-individuazione PR') l'individuazione della Persona di Riferimento. Invece, un rifiuto ottenuto al telefono costringe comunque il rilevatore ad andare di persona presso l'indirizzo. L'entità dei rifiuti viene osservata anche per valutare le capacità assertive e comunicative dei rilevatori, nella gestione dei primi contatti con le famiglie.
- *Le cadute per contatore esaurito* (EDFS = 04 e 05): tra i motivi di caduta previsti, ve ne sono alcuni specifici della tecnica CAPI, ovvero quelli inerenti i contatori, che si incrementano al momento della registrazione sul software CAPI di un nuovo tentativo di contatto (telefonico o a domicilio). Si tratta delle cadute per contatore esaurito (CE), assenti nell'elenco dei motivi di mancata risposta previsti dal modello BF/5 per il monitoraggio della rilevazione con tecnica PAPI. Poiché come noto il contatto con la famiglia consente anche l'accertamento della sua eleggibilità, gli esiti per CE si dividono in 'Contatore esaurito con contatto' e 'Contatore esaurito senza contatto'. Nel primo caso, essi rendono conto della difficoltà incontrata dai rilevatori, nel concordare un appuntamento per iniziare l'indagine (famiglia eleggibile), nel secondo caso, dei problemi nel reperire le famiglie in casa (eleggibilità incerta). L'esaurimento dei contatori pone il rilevatore nella condizione di scegliere se sostituire la famiglia, oppure se insistere con ulteriori tentativi, sulla base di valutazioni fatte sul posto (indicazioni dei vicini, altri segni di presenza). In generale, gli intervistatori vengono invitati ad effettuare i vari tentativi, in orari e giorni diversi; ciò può essere verificato a posteriori, ricostruendo gli eventi per le famiglie cadute per questo motivo. Inoltre, quando i contatori si saturano, nonostante ci sia stato almeno un contatto con la famiglia, si potrebbe essere in presenza di un rifiuto camuffato.

²⁰ Ovviamente, queste condizioni valgono per tutte le cadute, che si verificano in fase iniziale.

- *Le cadute per assenza prolungata* (EDFS = 06): le assenze prolungate rappresentano, insieme ai rifiuti iniziali, le cadute più considerevoli. Il codice di 'Assenza prolungata' viene assegnato quando la famiglia eleggibile non è presente presso l'abitazione negli intervalli di tempo previsti. Tale codice, registrato nei giorni utili per l'intervista iniziale, fa scattare automaticamente la possibilità di sostituzione. Se da un lato nei periodi dell'anno dedicati alle vacanze le assenze prolungate sono un fisiologico ostacolo, dall'altro potrebbero includere casi diversi, come ad esempio le seconde case.
- *Le cadute per la presenza di malato grave* (EDFS = 07): indicano le situazioni in cui lo stato di salute di uno o più membri della famiglia rende non praticabile la partecipazione alla rilevazione. Anche in questo caso, si configura la sostituzione automatica della famiglia.
- *Le cadute per irreperibilità* (EDFS = 08): le famiglie esitate come irreperibili sono quelle che non si trovano all'indirizzo che risulta dagli elenchi anagrafici e di cui non si riescono ad ottenere notizie circa la reale dimora. L'eleggibilità rimane incerta.
- *Le cadute fuori target* (EDFS = 09): le famiglie che risultano fuori target (non eleggibili) forniscono una misura del grado di sovra-copertura della lista di estrazione. Questo aggregato include le famiglie decedute, quelle che vivono all'estero, le persone che vivono permanentemente in un'istituzione, le attività commerciali e simili.
- *Le cadute per fine periodo* (EDFS = 10): le famiglie che, infine, cadono per 'Fine periodo' rappresentano un problema connaturato con la metodologia dell'indagine, che vincola a ben delimitate finestre temporali lo svolgimento delle tre fasi d'intervista. Come per le quartine non toccate, anche per le famiglie cadute per fine periodo, si può ipotizzare un problema di carichi di lavoro inadeguati. I nominativi che, pur assegnati, non vengono toccati sono esitati come 'FP per intervista iniziale: nuovo'. Gli altri esiti analitici specificano, ciascuno, se la caduta per fine periodo è intervenuta nel corso dei tentativi per effettuare l'intervista iniziale o quella finale; inoltre, è possibile desumere se, al momento della registrazione dell'esito, siano stati fatti più o meno tentativi, rispetto a quelli previsti. Il codice EDFA 'FP: manca codifica 9999', invece, si concretizza quando il rilevatore non procede alla codifica della professione entro l'intervallo destinato al completamento dell'intervista iniziale.

Nella Tavola 2.3 si riportano gli esiti familiari definitivi sintetici (EDFS) e analitici (EDFA) e i relativi valori²¹, registrati nell'anno d'indagine 2014.

21 Elaborazioni dai file degli Esiti mensili. I codici degli esiti familiari analitici non sono progressivi, in quanto i codici numerici vengono assegnati agli esiti in modo univoco tra le varie indagini *Computer Assisted*, per facilitare le analisi dei confronti.

2. Il disegno della nuova indagine

Tavola 2.3 – Principali esiti definitivi familiari – Anno 2014

EDFS	EDFA	ESITI DEFINITIVI FAMILIARI	v.a.	%
01	001	Complete	17.239	39,89
02		Interruzione definitiva	721	1,67
	002	- Interruzione definitiva	64	0,15
	101	- Intervista iniziale completa	171	0,40
	201	- Visita intermedia completa	486	1,12
03		Rifiuto iniziale	10.932	25,30
	005	- a domicilio pre-individuazione PR	6.012	13,91
	006	- a domicilio post-individuazione PR	4.920	11,39
04		Contatore esaurito con contatto	2	0,00
	011	- per eccesso di tent. con almeno un contatto per 1° appunt./interv.iniz.	-	-
	013	- per eccesso di appunt. dopo 1° appunt./inizio interv. Iniz.	2	0,00
	014	- per eccesso di tent. dopo 1° appunt./interv.iniz. (escluso caso precedente)	-	-
05		Contatore esaurito senza contatto	392	0,91
	007	- per 1° appunt./interv.iniz. per tutti non risponde/utente non raggiu./nessuno in casa	348	0,81
	008	- per 1° appunt./interv.iniz. per tutti occup.	-	-
	009	- per 1° appunt./interv.iniz. per tutti segret./fax	-	-
	010	- per 1° appunt./interv.iniz. per combinazione di soli non risp./non raggiu./occup./segret./fax/nessuno	2	0,00
	012	- per eccesso di altri tent. senza cont. per 1° appunt./interv.iniz.	42	0,10
06	015	Assenza prolungata	8.048	18,62
07	023	Malato grave	21	0,05
08		Famiglia irreperibile	1.676	3,88
	017	- famiglia vive nel comune ma non si sa dove	47	0,11
	018	- famiglia vive in altro comune italiano	1.367	3,16
	020	-irreperibilità per nome/indirizzo/ sconosciuto	262	0,61
09		Fuori target	672	1,56
	016	- seconda casa	-	-
	019	- famiglia vive all'estero	64	0,15
	021	- famiglia deceduta	458	1,06
	022	- famiglia vive permanentemente in istituzione/convivenza	38	0,09
	036	- negozio, impresa, azienda, ufficio	25	0,06
	037	- insieme di individui (ISF vive altrove)	87	0,20
10		Fine periodo	3.510	8,12
	024	- f.p. per INI: nuovo (nessuna azione)	2.880	6,66
	025	- f.p. per INI: meno tent. di quelli richiesti	404	0,93
	026	- f.p. per INI: più tent. di quelli richiesti, tutti non risponde/non raggiu./nessuno	-	-
	027	- f.p. per INI: più tent. di quelli richiesti, tutti occupati	15	0,03
	028	- f.p. per INI: più tent. di quelli richiesti, tutti segreteria/fax	-	-
	029	- f.p. per INI: più tent. di quelli richiesti, combinazione di non risp./non raggiu./occup./segret.	129	0,30
	030	- f.p. per INI: più tent. con almeno un contatto	10	0,02
	031	- f.p. per INI: più tent. senza un contatto	-	-
	032	- f.p. per INI: iniziati tent. senza contatto Comune	48	0,11
	033	- f.p. per INI: preso il 1° appunt.	13	0,03
	034	- f.p. per INI: intervista iniziata	3	0,01
	035	- Fine periodo: manca codifica 99999	4	0,01
	038	- f.p. per IFIN: intervista non iniziata	4	0,01
	039	- f.p. per IFIN: intervista iniziata	-	-
Totale esiti definitivi			43.213	100,00

Appare dunque chiaro come il complesso degli esiti definitivi familiari, codificati grazie alla tecnica *Computer Assisted*, consenta un quadro informativo molto più dettagliato di quanto era possibile con il solo supporto del modello cartaceo BF/5. Nel Prospetto 2.6, è disponibile la corrispondenza tra le due categorie di esiti di caduta.



Prospetto 2.6 – Corrispondenza dei motivi di caduta con tecnica PAPI e CAPI

MOTIVI DI CADUTA PAPI (modello BF/5)	MOTIVI DI CADUTA CAPI (EDFA) (questionario elettronico Scheda contatti)
Rifiuto iniziale	Rifiuto iniziale (005, 006)
Irreperibilità	Famiglia irreperibile (017, 020)
Decesso	Famiglia deceduta (021)
Momentanea assenza	Assenza prolungata (015)
Domicilio di fatto in altro comune	La famiglia vive in un altro comune italiano (018)
Emigrazione estero	La famiglia vive all'estero (019)
Rilevazione incompleta	Interruzione definitiva (002, 101, 201)
Altri motivi	Malato grave, Fine periodo, Contatore esaurito (023, 024 - 035, 038, 039, 007 - 009, 010 - 014)

Infine, il monitoraggio serve anche a definire il contesto d'intervista, considerando principalmente due indicatori: il primo riguarda la ricezione della lettera da parte della famiglia; il secondo la durata delle singole interviste. In questo modo, si può valutare l'impatto delle strategie di sensibilizzazione adottate e il disturbo statistico arrecato alle famiglie coinvolte.

Gli indicatori di monitoraggio indiretti

Per completare il monitoraggio del lavoro della rete di rilevazione, al set di indicatori basati sulle informazioni desumibili dalla scheda contatti deve affiancarsi quello che si fonda sui dati di indagine. Si tratta, in questo caso, di indicatori indiretti relativi all'operato della ditta e a quello dei rilevatori. Tali indicatori consentono di controllare i record familiari ricevuti in termini sia quantitativi, verificando la discrepanza tra campione mensile teorico ed effettivo, sia qualitativi, individuando record di qualità non accettabile.

Inoltre, poiché l'investimento formativo è un aspetto assolutamente centrale della nuova indagine CAPI, come già ricordato nei paragrafi precedenti, gli indicatori hanno anche l'obiettivo di indirizzare azioni preventive, confluendo in nuovi input da trasmettere nelle successive tornate formative.

Indicatori per il controllo quantitativo

- *Percentuale di record registrati per comune*: per ciascun ciclo mensile di rilevazione, viene calcolato il rapporto percentuale tra numero di record pervenuti per comune e numero di interviste attese; tale indicatore fornisce una misura preliminare²² del tasso di copertura campionaria.
- *Percentuale di record con diario cartaceo non pervenuto*: per ciascun ciclo mensile di rilevazione, viene calcolato il rapporto tra numero di record registrati per i quali non risulta pervenuto il diario cartaceo (e che pertanto sono eliminati dal file dei dati) e totale dei record registrati.

Indicatori elementari per il controllo qualitativo

- *Diario vuoto*: il diario non è stato compilato in alcuna delle sue parti.
- *Numero di spese riportate nel diario*: contatore del numero totale di voci di spesa riportate nel diario. Se il numero di spese registrate è troppo alto o, al contrario, eccessivamente contenuto, la rilevazione non è stata condotta correttamente.

²² Il tasso di copertura definitivo è invece calcolato al termine delle procedure di controllo e correzione dei dati.

2. Il disegno della nuova indagine

- *Intervista finale vuota*: contatore del numero totale di voci di spesa riportate nell'intervista finale. I record con contatore pari a zero vengono considerati con intervista finale vuota.
- *Numero di spese riportate nell'intervista finale*: contatore del numero totale di voci di spesa riportate nell'intervista finale. Si selezionano le interviste con contatore troppo basso (in particolare quelle in cui sono assenti le spese per abitazione e utenze) o troppo elevato.
- *Numero di giorni del diario con almeno una spesa*: contatore del numero di giorni in cui nel diario è riportata almeno una voce di spesa.
- *Percentuale di importi di spesa arrotondati nel diario*: rapporto percentuale tra il numero totale di voci di spesa riportate nel diario per le quali gli importi risultano arrotondati (privi dell'indicazione dei centesimi di euro) e il numero totale di spese riportate nel modello.
- *Frequenza di acquisto di zucchero, olio e sale*: contatore del numero di spese per zucchero, olio e sale. Una frequenza elevata è indicatore di errore in quanto la confezione minima costituisce in generale una scorta utile per un periodo superiore ai 14 giorni di compilazione del diario.
- *Frequenza di autoconsumi elevati*: contatore del numero di autoconsumi che superano i 50 chilogrammi o litri, i 10 bicchieri o i 20 cucchiaini. Una frequenza elevata è indice di mancata comprensione del concetto stesso di autoconsumo o di non accurata compilazione del diario da parte della famiglia.
- *Frequenza di autoconsumi espressi in una doppia unità di misura*: contatore del numero di voci autoconsumate per le quali la quantità è stata espressa, nell'ambito di un singolo giorno, con due diverse unità di misura. Benché in alcuni casi la doppia indicazione possa essere accettabile (se la famiglia riporta due quantità diverse) nella maggior parte dei casi ciò è indice di duplicazione dei quantitativi consumati.

I record affetti da anomalie gravi in assenza di note (della famiglia o dell'intervistatore) che le giustifichino, in tutto o in parte, sono annullati ed eliminati dal file dei dati. Diversamente, qualunque decisione in merito è preceduta da un'attenta analisi dell'insieme degli indicatori e del contenuto di tali note (cfr. par. 2.4.2.1).

Gli indicatori elementari descritti vengono utilizzati per la costruzione di indicatori sintetici allo scopo di monitorare le performance mensili della rete di rilevazione:

- *Percentuale di record familiari annullati automaticamente per rilevatore*: per ciascun rilevatore, si calcola il rapporto percentuale tra i record eliminati dal file dei dati per anomalie gravi non altrimenti giustificate e il totale dei record registrati.
- *Percentuale di record familiari annullati previa verifica per rilevatore*: per ciascun rilevatore, si calcola il rapporto percentuale tra i record gravemente anomali eliminati dal file dopo averne verificato la qualità e le note e il totale dei record registrati.
- *Numero di diari vuoti per rilevatore*: per ciascun rilevatore, si calcola il numero di diari vuoti. L'elenco dei rilevatori con diari vuoti è aggiornato ogni mese e tutti coloro che presentano una percentuale elevata di diari vuoti in più cicli mensili di rilevazione sono richiamati ed eventualmente sostituiti.

Qualora le anomalie si ripetano più volte per uno stesso rilevatore, nell'arco di un singolo mese o in mesi successivi, si procede ad una segnalazione al supervisore di rete e ad un ritorno formativo sul rilevatore stesso.

I controlli telefonici

La valutazione della qualità di un'indagine passa anche attraverso una verifica attenta e puntuale della rispondenza dei comportamenti dei rilevatori alle indicazioni impartite in sede di formazione. Le interviste telefoniche di *recall*, sottoposte a un sottogruppo di famiglie del campione, è la modalità prescelta dall'Indagine sulle spese per accertare la corretta conduzione delle interviste da parte della rete. Tale attività è rimasta invariata nel passaggio dalla vecchia alla nuova indagine, ma è stata migliorata in termini di tempestività e, conseguentemente, di efficacia. Il periodo entro cui vengono completati i controlli è attualmente pari a trenta giorni, a partire da quando si conclude il ciclo mensile di rilevazione, mentre nella vecchia indagine erano necessari sessanta giorni. Ciò consente di individuare per tempo le performance anomale meritevoli di essere evidenziate.

L'obiettivo di questa attività di ritorno sulle famiglie è duplice: da una parte, si intende prevenire la diffusione di comportamenti non in linea con le scelte metodologiche adottate per l'indagine, dall'altra, si vuole testimoniare ai rilevatori l'attenzione dell'Istituto al corretto svolgimento della rilevazione e al conseguimento di dati di qualità.

Da un punto di vista operativo, si è scelto di procedere suddividendo i controlli telefonici in due sottogruppi: nel primo, rientrano le interviste svolte nel mese di riferimento da un insieme di rilevatori estratti casualmente; nel secondo, invece, le interviste eseguite dai rilevatori che attraverso il sistema di monitoraggio dell'indagine hanno mostrato comportamenti non conformi alle attese. Tra i criteri adottati per individuare questi ultimi si segnalano: il conseguimento di tassi di completezza anomali; lo svolgimento di un numero insostenibile di interviste nello stesso giorno; la registrazione di tempi di somministrazione del questionario eccessivamente brevi. Alcuni controlli inoltre, come si è detto, derivano dalle analisi sulla qualità dei dati effettuati dai responsabili tematici dell'indagine, che possono intravedere la presenza di valori anomali in alcune variabili, la ricorrenza di diari non compilati o altre evidenze di impiego inadeguato degli strumenti di rilevazione.

Per facilitare l'approccio con le famiglie – sottoposte a un nuovo 'disturbo statistico' da parte dell'Istituto – e standardizzare la raccolta delle informazioni, sono stati predisposti i seguenti materiali ad uso dei responsabili dei *recall* telefonici:

- manuale di istruzioni sulle procedure del contatto;
- canovaccio di comunicazione da seguire nelle telefonate;
- tracciato record e relativa maschera di inserimento per la registrazione informatica dei dati.

Ogni mese, al termine dei *recall* telefonici eseguiti sulle interviste estratte per il periodo in questione, i dati vengono analizzati e riassunti in un'apposita reportistica. I report sono quindi utilizzati sia per valutare l'accettazione della fornitura, sia per segnalare alla società le criticità eventualmente riscontrate a cui porre rimedio. Nel 2014 sono stati sottoposti a controlli telefonici circa quindici rilevatori al mese. La lista delle famiglie da contattare è stata costruita estraendo quelle lavorate dai rilevatori segnalati dal totale delle famiglie completate nel mese di riferimento. Tale lista contiene mediamente il dieci per cento delle famiglie con intervista completa, tuttavia solo per una parte di esse si dispone del numero di telefono. Complessivamente pertanto, sempre nel 2014, sono state effettuate poco meno di seicento chiamate, delle quali

più della metà sono stati contatti utili alle verifiche²³. Gli esiti dei controlli hanno portato ad evidenziare alcune anomalie, che si sono attestate mediamente intorno al tre per cento delle chiamate di controllo utili.

2.4 L'acquisizione e il trattamento dei dati

Arricchendosi di nuovi contenuti informativi, l'indagine necessita di un sistema di *data entry* e di controllo/correzione dei dati più ampio e complesso rispetto a quello utilizzato per l'indagine PAPI (D'Aurizio, Grassi, Macchia, Manzari, Pannuzi, 1999). L'utilizzo di un tecnica CAPI, attraverso l'implementazione di un piano di controllo delle compatibilità tra variabili, ha consentito di anticipare alla fase di acquisizione dell'informazione una parte considerevole dei controlli, ottenendo dati grezzi meno affetti da errore. Il piano delle compatibilità del questionario e il piano di controllo e correzione dei dati risultano integrati in un unico sistema, garantendo una maggiore qualità dell'informazione rilasciata.

Nonostante l'anticipazione dei controlli alla fase di acquisizione, il piano di controllo e correzione rimane complesso e la procedura si articola in diverse fasi:

- Acquisizione controllata delle informazioni
 - i. Definizione dei percorsi del questionario elettronico
 - ii. Definizione del piano di compatibilità
 - iii. Acquisizione controllata del diario in modalità CADI
- Controlli mensili sui record familiari
 - iv. Controlli quantitativi
 - v. Controlli qualitativi
 - vi. Individuazione record familiari con anomalie gravi
 - vii. Controllo e correzione delle variabili socio-demografiche individuali per la ricostruzione delle relazioni familiari (Procedura famiglie)
 - viii. Controlli e correzioni deterministiche
- Controlli trimestrali sui record familiari
 - ix. Controlli e correzioni probabilistiche delle variabili qualitative
 - x. Controlli e correzioni probabilistiche delle variabili quantitative: spese, reddito e superficie dell'abitazione
- Controlli annuali sui record familiari e predisposizione dei file di microdati per la diffusione
 - xi. Controllo dell'effetto complessivo delle correzioni apportate in base al piano delle compatibilità
 - xii. Predisposizione dei file di microdati: ARMIDA, File per il SISTAN, File per la ricerca, File mlcro.STAT.

23 Il numero delle telefonate eseguite tiene conto delle regole di contatto stabilite dall'Istituto. Queste prevedono di interrompere le telefonate per lo stesso rilevatore se, dopo tre interviste, non emerge alcuna segnalazione di comportamento anomalo da parte delle famiglie.

2.4.1 L'acquisizione controllata delle informazioni

Nell'indagine PAPI tutte le informazioni, sia quelle rilevate tramite l'intervista sia quelle rilevate tramite il diario, venivano acquisite in modalità CADI (Computer Assisted Data Input), attraverso il software Blaise che controllava l'acquisizione segnalando la presenza di errori (D'Aurizio, Grassi, Macchia, Manzari, Pannuzi, 1999). Poiché il dato veniva acquisito in una fase successiva a quella di rilevazione, il controllo degli errori aveva unicamente lo scopo di inibire eventuali errori di digitazione. Per tale motivo tutti i tipi di errore, ad eccezione dei percorsi filtrati del questionario, venivano segnalati all'operatore attraverso messaggi di *warning* e non inibivano il proseguimento dell'acquisizione.

L'utilizzo di un questionario elettronico per la somministrazione delle due interviste, iniziale e finale, ha consentito di effettuare come già detto, controlli già in fase di rilevazione dei dati, dando la possibilità di prevenire gli errori causati dal rispondente o dal rilevatore, e di sanare eventuali incongruenze assieme al rispondente (Istat, 2006b). L'acquisizione in modalità CADI si è quindi ridotta alla registrazione dei diari, con notevole contrazione dei tempi e dei carichi di acquisizione e con indubbi guadagni sia in termini di tempestività che di qualità.

Il piano di acquisizione dei dati rilevati (tramite intervista, diario e scheda contatti) è stato implementato dalla società esterna incaricata di gestire l'attività di rilevazione con il *software* Converso.

2.4.1.1 Il piano delle compatibilità e i percorsi del questionario elettronico

Al fine di sfruttare al meglio le potenzialità del questionario elettronico e minimizzare i possibili errori in fase di rilevazione/acquisizione è necessario definire regole che controllino:

- errori di rotta in base ai percorsi filtrati del questionario
- errori di *range* dovuti all'inserimento di un dato esterno al campo di definizione di una variabile
- errori di coerenza tra variabili.

Per definire un piano di compatibilità (Istat, 2004b) occorre innanzitutto distinguere tra:

- regole formali, che derivano dalla struttura del questionario, cioè direttamente dalle norme di compilazione e dai percorsi interni (salti);
- regole sostanziali, che derivano da considerazioni di tipo statistico o da conoscenze specifiche del fenomeno oggetto di rilevazione.

Di fronte a un'informazione incoerente, due sono le possibilità (D'Aurizio et al., 1999; Istat, 2006b):

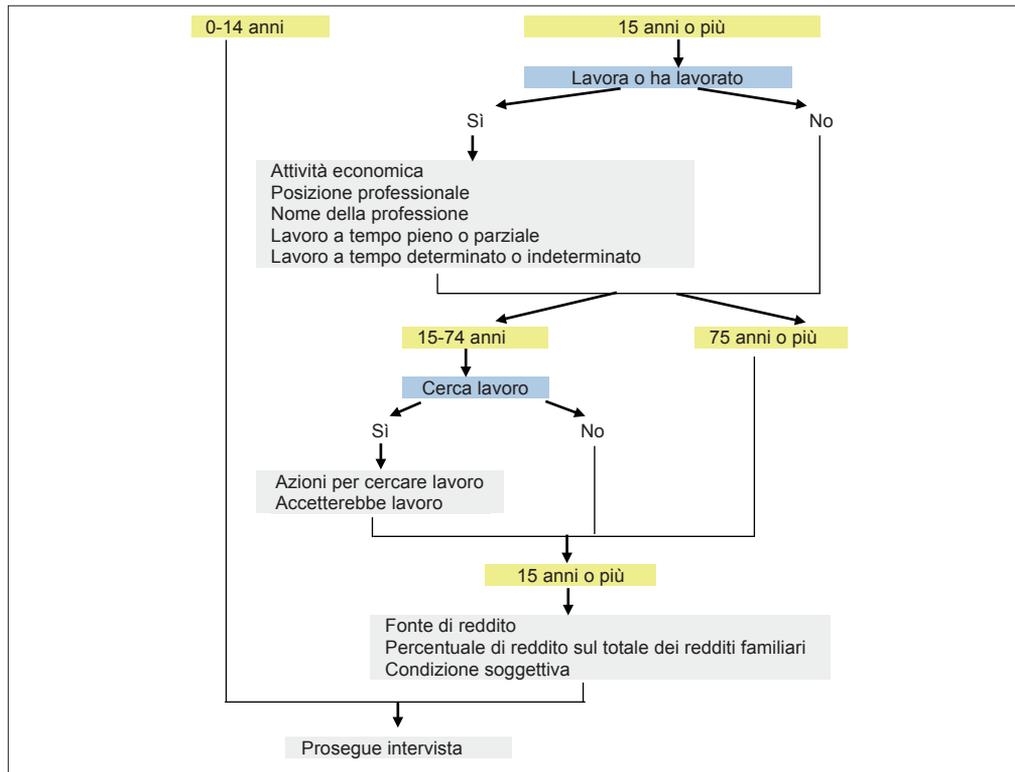
- inibire la registrazione del dato errato, in tal caso l'intervista non può proseguire finché non viene corretto il dato errato;
- permettere la registrazione del dato errato, previa segnalazione dell'errore.

I controlli del primo tipo vengono definiti *hard*, quelli del secondo tipo *soft*.

Nell'Indagine PAPI sui consumi, in cui la fase di acquisizione era successiva a quella di rilevazione, le sole regole che venivano trattate in modalità *hard* erano quelle di rotta e quelle di coerenza tra i codici di aggancio dei diversi questionari.

Nella nuova indagine, l'opportunità di effettuare controlli immediati e sanare le incongruenze in presenza dell'intervistato, ha reso possibile l'utilizzo di un maggior numero di regole *hard*.

Figura 2.13 – Percorso filtrato sulla condizione lavorativa



Le regole formali sono, in generale, piuttosto semplici in quanto come già detto (cfr. par. 1.3.6.1) la rilevazione della spesa non si presta ad essere filtrata. I percorsi filtrati riguardano solamente alcune categorie di spesa: i) la manutenzione ordinaria dell'abitazione ii) l'istruzione iii) le altre abitazioni di proprietà e iv) collaboratori conviventi; più numerosi sono invece quelli presenti nella sezione dedicata alla rilevazione delle caratteristiche individuali, in particolare della condizione lavorativa (Figura 2.13).

Le regole sostanziali inserite nel questionario elettronico riguardano essenzialmente i quesiti della scheda generale. Sono stati definiti due intervalli di accettazione, uno più stringente che attiva controlli di tipo *hard* quando l'occorrenza, per legge o consuetudine, risulta assolutamente inaccettabile e uno più lasco che attiva controlli di tipo *soft* quando l'occorrenza è altamente improbabile (Prospetto 2.7). Ad esempio, nel controllo tra il titolo di studio più elevato conseguito e l'età, per la licenza media è prevista una regola *hard* nel caso l'età sia inferiore a 11 anni e una regola *soft* nel caso l'età sia pari a 12 anni.

Qualora i controlli riguardino quesiti collocati in sezioni diverse o addirittura in questionari diversi, vista la difficoltà a ripercorrere l'intero flusso del questionario, sono stati previsti quesiti di controllo che consentono l'inserimento dell'informazione corretta.

Nel questionario elettronico, infine, per ogni variabile di spesa è stato previsto un controllo *soft* che, in base a intervalli di accettazione (aggiornati periodicamente

sulla base delle distribuzioni empiriche della spesa), segnala l'inserimento di valori anomali; ciò ha lo scopo di minimizzare gli errori di digitazione.

Prospetto 2.7 – Principali regole del piano di compatibilità per la Scheda Generale

Hard	Soft
Relazione familiare	
Coniugati di sesso maschile con meno di 16 anni Coniugati di sesso femminile con meno di 14 anni Differenza di età tra genitore e figlio inferiore a 10 anni Differenza di età tra padre e figlio superiore a 75 anni Differenza di età tra madre e figlio superiore a 60 anni	Coniugati di sesso maschile con meno di 14 anni Coniugati di sesso femminile con meno di 12 anni Conviventi di sesso maschile con meno di 16 anni Conviventi di sesso femminile con meno di 14 anni Differenza di età tra genitore e figlio inferiore a 14 anni Differenza di età tra padre e figlio superiore a 60 anni Differenza di età tra madre e figlio superiore a 50 anni Differenza di età tra genitore e figlio nato da precedente matrimonio inferiore a 14 anni Differenza di età tra padre e figlio nato da precedente matrimonio superiore a 60 anni Differenza di età tra madre e figlio nato da precedente matrimonio superiore a 50 anni Differenza di età tra i due coniugi/conviventi superiore a 30 anni
Numero di componenti con stato civile coniugato non è un multiplo di 2 Coniuge della PR con stato civile non coniugato Convivente delle PR con stato civile coniugato Coniugato o separato di fatto o separato legalmente di sesso maschile con meno di 14 anni Coniugato o separato di fatto o separato legalmente o vedovo di sesso femminile con meno di 12 anni Divorziato di sesso maschile con meno di 17 anni Divorziato di sesso femminile con meno di 15 anni Anno di matrimonio tra PR e il suo coniuge diversi Anno di matrimonio precedente all'anno di nascita	(Coniugato, separato di fatto o separato legalmente) di sesso maschile con meno di 16 anni (Coniugato, separato di fatto, separato legalmente o vedovo) di sesso femminile con meno di 14 anni Divorziato di sesso maschile con meno di 19 anni Divorziato di sesso femminile con meno di 12 anni
Titolo di studio	
Licenza elementare a meno di 9 anni Licenza media a meno di 12 anni Diploma che permette l'accesso all'università a meno di 14 anni Diploma accademia a meno di 16 anni Diploma universitario a meno di 18 anni Diploma di laurea o laurea di 1° livello a meno di 19 anni Laurea di 1° livello a meno di 19 anni Diploma di laurea a meno di 20 anni Laurea del vecchio ordinamento a meno di 21 anni Laurea specialistica a meno di 22 anni Titolo post-laurea o dottorato a meno di 21 anni	Licenza elementare a meno di 10 anni Licenza media a meno di 13 anni Diploma che permette l'accesso all'università a meno di 16 anni Diploma accademia a meno di 19 anni Diploma universitario a meno di 20 anni Diploma di laurea o laurea di 1° livello a meno di 22 anni Laurea di 1° livello a meno di 22 anni Diploma di laurea a meno di 23 anni Laurea del vecchio ordinamento a meno di 24 anni Laurea specialistica a meno di 24 anni Titolo post-laurea o dottorato a meno di 24 anni
Condizione lavorativa	
Apprendista con più di 34 anni	Dirigente o libero professionista ma non ha conseguito la laurea Apprendista con età compresa tra 30 e 34 anni Dirigente, quadro, imprenditore o libero professionista con età compresa tra 15 e 17 anni Lavoratore autonomo e lavora nel settore "amministrazione pubblica e difesa; assicurazione sociale obbligatoria" o "Organizzazioni ed organismi extraterritoriali" Imprenditore o libero professionista e lavora nel settore "Attività di famiglie e convivenze come datori di lavoro per personale domestico; produzione di beni e servizi indifferenziati per uso proprio da parte di famiglie e convive"

2.4.1.2 L'acquisizione CADI del diario

Con l'indagine PAPI sui consumi, i questionari cartacei venivano inviati mensilmente all'Istat dagli uffici comunali di statistica. Personale esperto verificava la corretta esecuzione di tutte le operazioni di raccolta dei dati e, in caso di lacune e anomalie, provvedeva a contattare telefonicamente il rilevatore o la famiglia stessa per sanare l'incongruenza riscontrata (Istat, 2000). Una volta revisionati, i questionari venivano registrati su supporto informatico da una società esterna, mediante il software Blaise predisposto dall'Istat al fine di ridurre gli errori di digitazione.

Le fasi di rilevazione e acquisizione risultavano pertanto nettamente distinte; con la nuova indagine tale distinzione permane solamente per il diario, la cui acquisizione è a cura del rilevatore che diventa responsabile dell'intero processo di acquisizione dell'informazione (Barcherini, Falorni, Grassi, 2014).

Il controllo della compilazione del diario viene effettuato dal rilevatore assieme alla famiglia in due occasioni: la prima a metà circa del periodo di compilazione del diario e la seconda al ritiro dello stesso. Il rilevatore è supportato da un questionario elettronico che, mediante percorsi filtrati (in base alle diverse problematiche riscontrate), fornisce indicazioni circa le soluzioni più adeguate e consente di tracciare gli eventuali errori di compilazione. Incongruenze ed errori possono essere sanati assieme alla famiglia e ciò ovviamente comporta un abbattimento dei controlli e degli interventi a posteriori, riguardo sia l'annullamento di eventuali record, sia la correzione delle spese errate.

Anche per l'acquisizione dei dati del diario il *software* è Converso, lo stesso utilizzato per le due interviste, e riproduce a video la pagina del diario. Il rilevatore, una volta "certificato" il dato assieme alla famiglia, deve acquisire tutte le informazioni contenute nel diario cartaceo (Figura 2.14).

Una questione cruciale in una indagine che si avvale di strumenti di rilevazione diversi (nella fattispecie diari cartacei e questionari CAPI) è, come già detto, l'aggancio delle informazioni rilevate con i diversi strumenti (cfr. par. 2.3.2.2). Al fine di contenere gli errori di digitazione, la maschera di acquisizione che il rilevatore utilizza per il diario è già parzialmente compilata dal sistema e le informazioni da inserire si limitano al codice quartina, al numero progressivo generale e al codice elenco.

Figura 2.14 – Un esempio della maschera di acquisizione

	GIORNO 1	GIORNO 2	GIORNO 3	GIORNO 4	GIORNO 5	GIORNO 6	GIORNO 7
PANE Pane, inclusi pan carrè, pan biosche, piadine	2.45		3.08				
GRISSINI E CRACKER Orsini, cracker, gallette, friselle e altri scottati secchi del pane, incluso pangrattato							
PASTA Pasta fresca o secca; inclusi pastine per bambini, pasta all'uovo, ripiena o farcita, gnocchi			1.00				
RISO Riso e risotti pronti in busta, inclusi spaghetti di riso, creme di riso per bambini			3.89		70.00		
ORZO, FARRO E ALTRI CEREALI IN CHICCHI Orzo, farro, mais, cous-cous e altri cereali in chicchi, anche aggiunto di altri ingredienti							

I controlli previsti durante la fase di acquisizione dei dati sono solamente di tipo *soft* al fine di evitare i possibili errori di digitazione da parte del rilevatore; come per il questionario CAPI, sono previsti, per ogni voce di spesa, intervalli di accettazione per segnalare eventuali valori anomali e sono anche previsti controlli specifici per le diverse sezioni del questionario (Prospetto 2.8) soprattutto in corrispondenza dei punti critici del questionario.

Prospetto 2.8 – Controlli di coerenza nelle maschere di acquisizione CADI

SEZIONE DEL QUESTIONARIO E TIPO DI VARIABILE	CONTROLLO INSERITO
Sezione 1 - Spese per bar, pasti fuori, da portar via	
Spesa	Range di accettazione
Quanto pagato con buoni pasto	Spesa pagata con buono pasto superiore alla spesa totale informazione mancante
Numero di persone a cui si riferisce la spesa	informazione mancante
Sezione 2 e 3 - Spese	
Spesa	Range di accettazione
Sezione 4 – Beni prodotti dalla famiglia	
Quantità	Informazione duplicata espressa con entrambe le unità di misura
	Range di accettazione
Sezione 5 – Luoghi di acquisto	
Numero di luoghi (possibili 2)	Barrati più di due luoghi Informazione mancante (nessun luogo barrato)
Domande finali	
Domande	Rispetto dei filtri del questionario

La violazione di tali controlli comporta un messaggio di *alert* per il rilevatore, che potrà procedere a correggere gli eventuali errori di digitazione o a mantenere i valori fuori-*range* nel caso siano effettivamente riportati sul diario cartaceo.

Il numero di diari da acquisire, nell'arco del mese, per ciascun rilevatore è piuttosto contenuto (circa 9 diari al mese con una media di 140 informazioni digitate per questionario), ciò riduce gli errori di digitazione rispetto a quanto avveniva con l'utilizzo di personale preposto alla sola acquisizione dei dati.

2.4.2 Il trattamento dei dati

La fase di trattamento dei dati è suddivisa, come nell'indagine PAPI sui consumi, in tre sotto-fasi: i) mensile, ii) trimestrale, iii) annuale (Prospetto 2.9).

I trattamenti mensili sono di tipo deterministico, mentre al file trimestrale vengono applicate tutte le procedure di trattamento basate su metodi di correzione di tipo probabilistico, con particolare riferimento ai metodi da donatore; la procedura prevede software e metodi di imputazione distinti per variabili qualitative e quantitative e include anche la stima del valore monetario delle quantità autoconsumate.

A livello annuale, il file risultante dall'accorpamento dei quattro trimestri viene validato nel suo complesso e reso disponibile.

La strategia di controllo e correzione cerca di salvaguardare il più possibile l'informazione rilevata, partendo dal presupposto che l'insieme dei controlli inseriti in fase di acquisizione fornisca dati grezzi di elevata qualità. Si è inoltre cercato, nei limiti imposti dalle caratteristiche e dalle peculiarità dell'Indagine sulle spese, di armonizzare le procedure di controllo e correzione con quanto previsto dalle altre indagini dell'Istituto.

Prospetto 2.9 – Gli strumenti utilizzati nel trattamento dei dati

TIPO DI VARIABILE	STRUMENTO	
	PAPI	CAPI
	Controlli mensili	
Controlli quantitativi	Manuali	Manuali
Controlli qualitativi	Manuali*	Indicatori di performance e correzioni deterministiche
Controllo di relazione di parentela, stato civile, anno di matrimonio	(Tra i controlli trimestrali)	Procedura famiglie
	Controlli trimestrali	
Variabili qualitative		
Controllo di relazione di parentela, stato civile, anno di matrimonio	Scia	(Tra i controlli mensili)
Altre variabili sulle caratteristiche dei componenti	Scia	Scia
Caratteristiche dell'abitazione e possesso di mezzi di trasporto e comunicazione	Scia	Scia
Abitudini di spesa e luoghi di acquisto	Scia	Scia
Mese di effettuazione della spesa	Non rilevate	Scia
Variabili quantitative		
Superficie dell'abitazione	Deterministico	IVEaware
Reddito	RIDA	IVEaware
Definizione limiti di spesa per individuazione dei dati anomali	Deterministico	Banff
Spese	RIDA	Banff
	Controlli annuali	
Controllo delle regole formali	Deterministico	Deterministico
Controllo delle regole sostanziali	Deterministico	Deterministico

* Effettuati da revisori esperti dell'Istat che, seguendo criteri standard, controllavano tutti i dati riportati e, in caso di informazioni mancanti o incongruenti, ricontattavano la famiglia o il rilevatore per procedere alla eventuale correzione o ad annullamento del questionario.

2.4.2.1 La fase di trattamento mensile

Il primo controllo dei dati avviene attraverso l'utilizzo degli indicatori elementari indiretti (cfr. par. 2.3.3.3) che permette di individuare i record anomali, distinguendo quelli da eliminare da quelli che, in base ad alcune caratteristiche, necessitano di un approfondimento di analisi.

Le soglie per l'individuazione dei record anomali vengono definite attraverso le distribuzioni mensili dei principali indicatori di performance, sono pertanto considerati anomali i record:

- con diari in cui il numero di spese riportate è superiore al 99° percentile;
- con diari in cui il numero di spese riportate è inferiore al 1° percentile;
- con diari in cui il numero di spese arrotondate supera il 95° percentile;
- con numero di spese nell'intervista finale inferiore al 1° percentile.

Sono inoltre considerati anomali tutti i record che hanno spese con informazione incompleta (a seguito della mancata indicazione del periodo di effettuazione della spesa, dell'importo o del luogo di acquisto) superiore all'80%.

Una volta individuati, tra i record anomali vengono eliminati in modo automatico tutti quelli:

- con codici identificativi mancanti (che generalmente corrispondono alla mancata compilazione del diario da parte della famiglia);
- senza alcuna nota esplicitiva che giustifichi l'anomalia.

La presenza di note rende indispensabile una valutazione *ad hoc*, associata anche agli altri indicatori di performance. Spiegazioni plausibili possono infatti indurre ad accettare un diario o un'intervista finale anche se completamente vuoti.

Nel caso di diari vuoti, il confronto con il cartaceo permette anche di recuperare i casi di mancata digitazione.

Complessivamente, nel 2013, i record annullati sono stati il 2,2% nell'indagine CAPI e il 3,1% nella PAPI, a conferma che le innovazioni introdotte hanno permesso di aumentare la qualità dell'informazione raccolta.

Mensilmente vengono corrette, per gli individui appartenenti alla famiglia di fatto, le variabili relative alla relazione di parentela, allo stato civile al momento dell'intervista e allo stato civile precedente al matrimonio. La correzione delle variabili avviene attraverso l'ormai consolidata "Procedura Famiglie" che, in base a regole complesse di compatibilità, segnala al rilevatore la presenza di errori, proponendo le possibili correzioni; l'operatore valuta quindi se e come correggere, anche in base alle ulteriori informazioni disponibili. In tal senso, l'utilizzo delle LAC per la selezione del campione rappresenta una preziosa fonte di informazioni: la possibilità di visionare il contenuto degli stati di famiglia rende assolutamente residuale l'intervento soggettivo dell'operatore.

Al fine di applicare correttamente tale procedura, la relazione di parentela viene rilevata in base a 17 modalità (non più 6 come nell'indagine PAPI), viene inoltre rilevato l'anno di matrimonio per i partner coniugati e, in armonia con altre indagini dell'Istituto, sono state inserite tre variabili che consentono di ricostruire i nuclei familiari: per ciascun componente si rilevano la presenza in famiglia del padre, della madre e del partner (coniuge o convivente). Nel 2013, la percentuale di correzioni sulla relazione di parentela scende dall'1,33% della PAPI allo 0,04% della CAPI e, sullo stato civile, dall'1,94% allo 0,33%.

Il trattamento mensile di tipo deterministico riguarda l'attribuzione degli importi di spesa, riportati nei campi aperti "altro specificare", alla voce di riferimento e la corretta collocazione degli importi indicati nelle note.

Sempre di tipo deterministico sono i controlli relativi alle variabili qualitative sulle caratteristiche dell'abitazione e il possesso dei mezzi di trasporto che predispongono (attraverso l'inserimento di *flag*), in caso di errore, la correzione probabilistica che verrà effettuata nella fase di trattamento trimestrale. Queste variabili pur non potendo presentare valori mancanti (la presenza di controlli *hard* impedisce infatti di proseguire la somministrazione del questionario), fatta eccezione per l'anno di costruzione dell'edificio in cui vive la famiglia, possono essere errate e generare incongruenze, come nel caso in cui il numero di stanze sia incompatibile con la superficie dell'abitazione.

2.4.2.2 La fase di trattamento trimestrale: le variabili qualitative

Le variabili qualitative rilevate dall'indagine CAPI sono aumentate considerevolmente (da 52 a 3.109). L'ampliamento dei contenuti informativi ha richiesto, infatti, l'introduzione di numerose variabili individuali, in particolare sulle caratteristiche socio-economiche, di alcune variabili relative alle caratteristiche dell'abitazione, ma soprattutto delle variabili qualitative riguardanti i mesi di effettuazione delle spese (per ogni voce con periodo di riferimento superiore al mese) e il luogo di acquisto (Tavola 2.4) (per ogni spesa con periodo di riferimento pari a 3 mesi sono previste 4 variabili qualitative, 13 nel caso di spese con periodo di riferimento di 12 mesi).

Tavola 2.4 – Numero di variabili qualitative rilevate con la vecchia e con la nuova indagine

	PAPI	CAPI
Caratteristiche dell'abitazione e possesso mezzi di trasporto e comunicazione	47	56
Luogo di acquisto	5	115
<i>rilevati con la finale</i>	5	65
<i>rilevati con il diario</i>	-	50
Mese di effettuazione della spesa	-	2.938
Totale	52	3.109

Come per la PAPI, per la correzione delle variabili qualitative è stato utilizzato il software SCIA (Sistema di controllo e imputazione automatica) che consente il controllo e la correzione probabilistica delle mancate risposte parziali e delle incoerenze logico/formali utilizzando la metodologia proposta da Fellegi e Holt. L'uso del questionario elettronico ha quasi annullato l'impatto del piano di correzione (nel 2013, oltre il 50% delle famiglie intervistate con la PAPI presentavano un'incoerenza nelle variabili individuali, mentre con la CAPI tale percentuale è scesa allo 0,5%): la percentuale di errori scende dal 6,59% allo 0,05% nel caso della condizione professionale, dal 30,93% allo 0,01% per la posizione nella professione e dal 24,91% allo 0,1% per l'attività economica. Anche per il tipo di combustibile prevalentemente utilizzato per il riscaldamento dell'abitazione, la presenza di errori si limita allo 0,43% dei casi (contro il 6,51% della PAPI) e allo 0,38% dei casi per il tipo di contratto di affitto (contro l'1,02% della PAPI).

Anche per le variabili relative al mese e al luogo di acquisto, la qualità dei dati rilevati nell'indagine CAPI è molto più elevata: nel 2013 i record con almeno una variabile qualitativa da correggere sono stati il 25%, contro l'oltre 50% della PAPI.

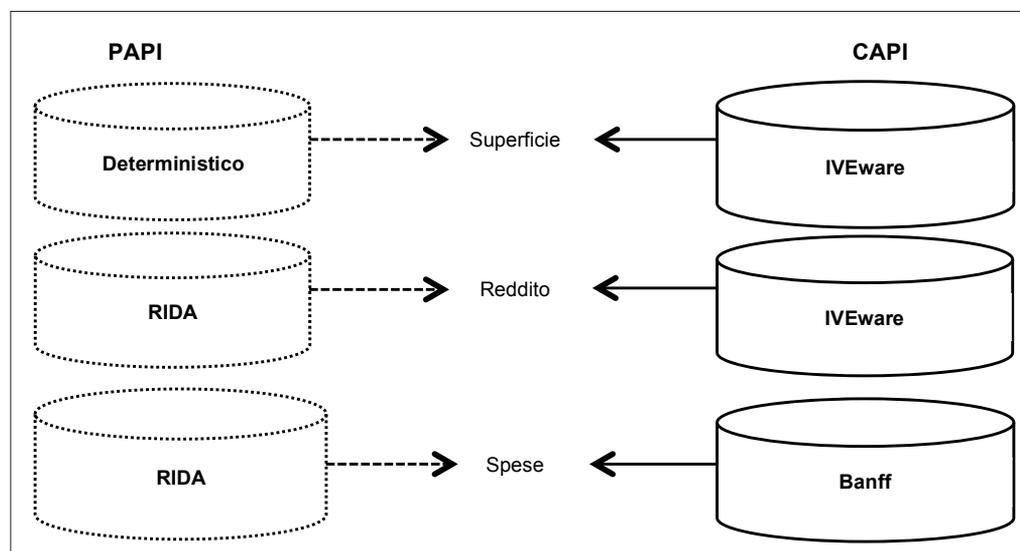
La percentuale di errore sui luoghi di acquisto è del tutto trascurabile e meno di un terzo delle variabili relative al mese di effettuazione della spesa sono affette da errore (la percentuale di record incongruenti è sempre inferiore all'1%); l'unica eccezione è rappresentata dal gas da rete: il 4% delle famiglie intervistate non è stato in grado di indicare la periodicità del pagamento.

2.4.2.3 La fase di trattamento trimestrale: le variabili quantitative

Le variabili quantitative presenti nell'indagine sono costituite, oltre che dagli importi di spesa, anche dalla superficie dell'abitazione in cui la famiglia vive e dal reddito netto medio mensile familiare.

Il trattamento di queste ultime precede quello delle spese visto che, come per l'indagine PAPI, il reddito è utilizzato come variabile ausiliaria (*matching*) per l'imputazione delle spese (Figura 2.15).

Figura 2.15 – Il flusso per la correzione delle variabili qualitative



Per il controllo e la correzione della superficie e del reddito, in accordo con altre indagini sulle famiglie, si è scelto di utilizzare il modulo *Impute* del software *IVEware*²⁴. Tale software effettua l'imputazione, singola o multipla, dei valori mancanti da modello, usando il metodo delle regressioni sequenziali (Raughunathan, *et al.* 2001). Nell'indagine PAPI, la superficie veniva corretta in modo deterministico, in base al numero di stanze; il reddito veniva invece trattato, similmente alle variabili di spesa, attraverso il metodo RIDA (*Ricostruzione Automatica dei Dati*) che effettua l'imputazione mediante l'individuazione del donatore a minima distanza (D'Aurizio *et al.* 1999).

Attraverso *IVEware*, è possibile sia definire il modello di regressione appropriato alla natura della variabile, sia specificare le covariate rilevanti. La tecnica consiste nell'imputazione multivariata dei valori mancanti, attraverso la generazione di modelli di regressione multipla univariata (approccio "variable by variable"). Ciascun modello confronta la distribuzione originaria con quella che si genera, passo dopo passo, con le imputazioni. Sia per la superficie che per il reddito, si è utilizzato il modello di regressione lineare normale; per rispettare la normalità, i valori delle variabili sono stati preventivamente trasformati attraverso il logaritmo e, dopo l'imputazione, sono stati ricondotti alla scala originaria.

Per la superficie dell'abitazione, come variabili esplicative sono state utilizzate la condizione professionale della persona di riferimento, il tipo di abitazione, il numero di stanze e il tipo di cucina²⁵, variabili qualitative corrette nel passo precedente. Nel 2013, la percentuale di valori mancanti per la superficie dell'abitazione è stata pari all'1,6%, valore pressoché identico a quello che si osservava nella PAPI (1,7%).

Nell'indagine PAPI, il quesito sul reddito veniva posto chiedendo alle famiglie di indicare la classe di riferimento; nella indagine CAPI, invece, si chiede di indicare il valore puntuale e, solamente nel caso in cui la famiglia non voglia o non sappia fornire l'informazione, è previsto un quesito di rilancio (cfr. par. 1.3.4.2). Nell'indagine CAPI, il valore puntuale viene imputato generando un valore casuale intorno

24 IVEware è un *freeware* (<http://www.isr.umich.edu/src/smp/ive/>) sviluppato dal Survey Research Center, Institute for Social Research dell'University of Michigan.

25 Cucina separata, angolo cottura, cucinino.

all'importo approssimativo indicato dall'intervistato nel quesito di rilancio. Nel caso la famiglia non indichi nemmeno l'importo approssimativo, l'imputazione viene fatta tenendo conto delle seguenti variabili esplicative: il titolo di studio, la posizione nella professione e il sesso della persona di riferimento, oltre al numero dei componenti della famiglia, la regione e la ripartizione di residenza. In accordo con quanto fatto da altre indagini dell'Istituto, il *range* di accettazione per il reddito è compreso tra i 150 euro e i 20.000 euro.

Gli effetti delle imputazioni e correzioni vengono valutati analizzando i principali parametri della distribuzione ottenuta (media, mediana, primo e terzo quartile, minimo, massimo, *standard error*) (Tavola 2.5).

Nel 2013, la distribuzione dei dati corretti non differisce in maniera rilevante da quella dei dati grezzi: la media è pressoché identica, pari a circa 1.900 euro, mentre la mediana aumenta, passando da 1.500 a 1.600 euro. Il valore minimo, come imposto dalle regole di correzione, è pari a 150 euro, ma è accettato anche il valore zero, nel caso la famiglia dichiari di non percepire alcun reddito.

Tavola 2.5 – Principali parametri della distribuzione del reddito prima e dopo la correzione (valori in euro)

Reddito	Dati grezzi	Dati corretti
Minimo	100	150
Massimo	20.000	20.000
Media	1.913	1.944
Mediana	1.500	1.600
Deviazione standard	1.574	1.584

Per il trattamento delle variabili di spesa, come già detto, la differenza sostanziale rispetto all'indagine PAPI è rappresentata dal fatto che, in passato, gli zeri di spesa potevano essere individuati solamente per poche voci di spesa, quelle concettualmente legate ad altre variabili: la spesa per affitto e il fitto figurativo (legato al titolo di godimento dell'abitazione), la spesa per le principali utenze domestiche (legata alla presenza dei servizi nell'abitazione), il pagamento dell'assicurazione (legato all'acquisto o possesso di mezzi di trasporto) e le spese per alcuni beni e servizi con carattere di eccezionalità (ad esempio onorari a professionisti, traslochi, assicurazioni sulla vita, assicurazioni per malattie e così via), per le quali era previsto uno specifico filtro. Con il passaggio alla nuova indagine CAPI, il filtro è stato inserito per tutte le spese rilevate con l'intervista diretta.

Per la correzione e l'imputazione degli importi di spesa, si è deciso di non utilizzare *IVEware*, ma di utilizzare metodi probabilistici da donatore. La spesa per singola voce, infatti, configurandosi spesso come un fenomeno raro, non si presta ad essere ben rappresentata tramite modello. La scelta²⁶ è pertanto ricaduta sul software Banff²⁷, che si basa sulla metodologia di Fellegi e Holt e che consente di individuare i valori anomali per ogni spesa in base alla loro distribuzione. Tale metodologia è stata preferita a quella utilizzata nell'indagine PAPI, che richiedeva la definizione a priori del *range* di accettazione, definito avvalendosi delle informazioni allora disponibili (il livello dei prezzi minimi e massimi rilevati dall'indagine sui prezzi al consumo; i limiti

²⁶ A tale proposito si ringrazia Marco Di Zio per i preziosi suggerimenti.

²⁷ Statistics Canada. 2008, Banff, functional description of the Banff system for edit and imputation, Versione 2.03, Banff Support Team.

inferiori e superiori delle distribuzioni campionarie delle spese) e massimizzando la probabilità di escludere i soli *outliers* tramite la tecnica D-Maso²⁸.

La logica sottostante il software Banff segue una struttura a moduli: a) localizzazione degli errori, b) identificazione dei valori anomali e c) imputazione/correzione.

I valori anomali (identificati tramite la specifica procedura Outlier) sono distinti in *outliers* (da imputare) e non *outliers* (da mantenere ma che non possono essere utilizzati come donatori). I limiti che individuano i valori *outliers* sono definiti in funzione di alcuni parametri impostati dal ricercatore, personalizzati per ciascuna voce di spesa; la dinamicità nel calcolo di tali limiti permette di tener conto degli eventuali cambiamenti nelle abitudini di spesa e della stagionalità delle voci da correggere trimestralmente.

Il serbatoio dei possibili donatori corrisponde ai valori compresi nell'intervallo interquartile; tale scelta garantisce la solidità dei valori imputati, appartenendo questi alla parte centrale della distribuzione. Il metodo adottato per l'imputazione e la correzione dei dati è quello del donatore di minima distanza (*Nearest-Neighbor Donor*), selezionato dal un serbatoio di unità prive di valori mancanti o anomali.

Nel 2013, l'indagine PAPI presentava il 50,6% dei record con valori di spesa corretti; il 33,5% necessitava di una correzione o imputazione, mentre il 15,9% di due o più correzioni/imputazioni. Nell'indagine CAPI, il 61,5% dei record era corretto, il 20,7% dei record ha subito una sola correzione/imputazione, mentre il 17,8% due o più correzioni/imputazioni.

2.4.2.4 La fase di trattamento trimestrale: la valutazione monetaria delle quantità autoconsumate

La spesa per consumi, per definizione, include anche il valore monetario dei beni e dei servizi autoconsumati. Le famiglie che nel corso del 2013 hanno effettuato autoconsumi (dall'indagine CAPI) rappresentano circa il 12% del totale; le voci oggetto di autoconsumo rilevate tramite diario sono 45, ma in media solo 5 beni sono consumati dalle famiglie.

Al fine di poter includere tali consumi nella spesa totale, si è reso necessario stimarne il valore monetario, applicando alle quantità consumate, quelle indicate dalle famiglie, il prezzo medio mensilmente stimato dal servizio Prezzi dell'Istat. L'Indagine sui prezzi al consumo rileva con un dettaglio molto più spinto rispetto a quello dell'Indagine sulle spese, di conseguenza alle 45 voci oggetto di autoconsumo sono stati associati circa 240 prodotti alimentari. Ogni mese, nei circa 80 comuni capoluogo di provincia che partecipano all'indagine sui prezzi, per i 240 prodotti in oggetto, vengono rilevate 107 mila quotazioni di prezzo; tali quotazioni sono state aggregate sia a livello merceologico, per passare dai singoli prodotti alle voci spesa, sia a livello territoriale, per passare dai prezzi medi provinciali a quelli regionali e poi ripartizionali (la stima ripartizionale risulta più robusta di quella regionale perché basata su un maggior numero di comuni che partecipano all'indagine).

Nello specifico, è stato calcolato il prezzo medio provinciale di ciascun prodotto come media geometrica delle quotazioni di prezzo rilevate (per i prodotti a rilevazione bimensile viene prima effettuata una media semplice delle due quotazioni rilevate nel mese per ciascuna referenza); tale prezzo medio è stato poi aggregato a livello

28 D'Aurizio, Grassi, Macchia, Manzari, Pannuzi, op. cit. 1999.

territoriale, attraverso una media aritmetica ponderata con un sistema di pesi basato sulla popolazione residente. Infine, i prezzi medi di ciascuna voce dell'Indagine sulle spese sono stati ottenuti come media ponderata (il sistema di pesi è stimato utilizzando i dati sui consumi finali delle famiglie della Contabilità Nazionale) dei prezzi medi dei prodotti afferenti a ciascuna voce. Per i generi alimentari autoconsumati che presentano una forte connotazione stagionale, la rilevazione dei prezzi viene effettuata solo in alcuni mesi dell'anno; per i periodi in cui il prezzo non è disponibile (non necessariamente il consumo è contestuale alla raccolta, si pensi ad esempio ai prodotti conservati) si utilizzano i valori medi mensili più elevati ottenuti per i quindici mesi precedenti.

I prezzi mensili così elaborati vengono applicati alle quantità autoconsumate rilevate tramite diario; in tal modo si stima il valore monetario degli autoconsumi che può essere direttamente sommato alla spesa sostenuta per gli acquisti.

2.4.2.5 La fase di trattamento annuale

Alla fine di ogni anno, prima della diffusione del dato, vengono eseguiti alcuni controlli di coerenza generale, si verifica ancora una volta, in base al piano delle compatibilità, l'effetto complessivo delle correzioni apportate nei singoli trimestri.

L'assenza di valori anomali nella spesa totale familiare contestualmente al rispetto delle regole formali e sostanziali rendono rilasciabile il dato.

Vengono quindi calcolati i coefficienti di riporto all'universo per i dati annuali tenendo conto degli aggiornamenti anagrafici sulla popolazione residente (cfr. paragrafo 2.5.2) e si procede alla predisposizione dei principali file di Microdati per la diffusione (file per ARMIDA e File di Microdati per la Ricerca), dei file per il datawarehouse dell'Istituto (I.Stat) e della nota per i media con il commento dei principali risultati.

2.5 La costruzione delle stime

2.5.1 La stima della spesa

L'indagine rileva le spese in maniera continua nell'intero arco di un anno, con tecnica mista e periodi di riferimento molto diversi. Il calcolo della spesa totale richiede dunque una standardizzazione temporale del dato rilevato, ovvero un riferimento temporale comune, in modo che tutte le spese possano essere sommate tra loro.

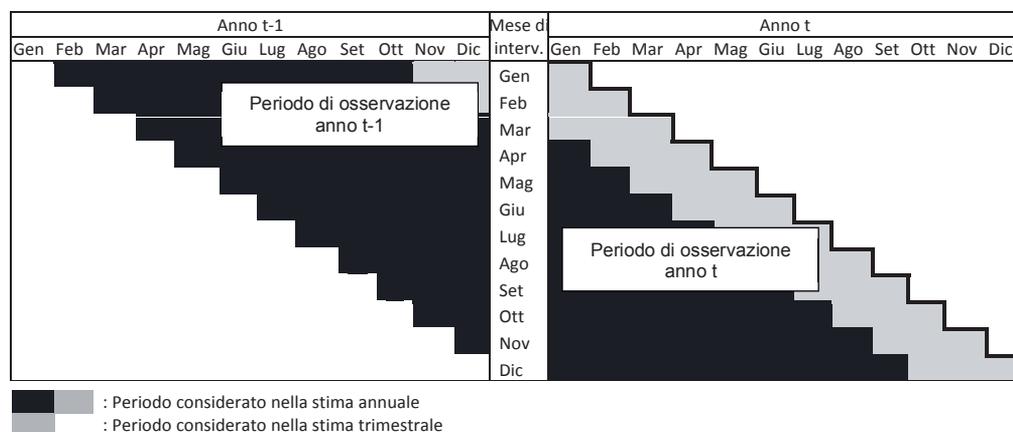
Con l'indagine PAPI tutte le spese venivano riportate al mese utilizzando opportuni stimatori, in modo da calcolare la spesa mensile complessiva per consumi. Se il periodo di riferimento della rilevazione era superiore al mese, si ipotizzava l'uniforme distribuzione nei mesi che componevano il periodo e la spesa rilevata veniva quindi divisa per il numero di mesi compresi nel periodo di rilevazione (ad esempio, le spese rilevate con riferimento a tre mesi venivano divise per tre, mentre quelle rilevate con riferimento a dodici mesi per dodici); per le spese rilevate con periodi di riferimento inferiori al mese (diario a 7 giorni), lo stimatore teneva conto di due elementi fondamentali che caratterizzano, in maniera diversa, i comportamenti di acquisto: la confezione minima e la scorta. Nell'ipotesi che il comportamento delle famiglie vari in funzione del bene acquistato e della settimana di rilevazione, il modello statistico di

riporto al mese teneva conto della proporzione di famiglie che acquistano il bene nei sette giorni di rilevazione (Crescenzi, Ceccarelli, 1999).

Con la nuova indagine CAPI, l'ampliamento dei periodi di riferimento e la possibilità di calendarizzare gli eventi di spesa consentono di effettuare ipotesi più precise circa l'effettivo comportamento di spesa delle famiglie. In particolare, un diario a 14 giorni approssima in modo sufficientemente corretto il comportamento mensile di spesa, giustificando l'utilizzo di un coefficiente pari a 2,14 (ottenuto come rapporto tra il numero medio di giorni nel mese e il numero di giorni di rilevazione). Per le spese rilevate con periodo di riferimento superiore al mese, l'informazione sul/i mese/i di effettuazione ha permesso di differenziare le stime trimestrali da quelle annuali.

L'ampliamento dei periodi di riferimento aumenta la numerosità campionaria, rendendo le stime più robuste, ma diminuisce la rappresentatività congiunturale delle spese, con conseguente attutimento della stagionalità. Si è quindi stabilito di utilizzare due sistemi di stima diversi, al fine di soddisfare, in un caso (R1), la rappresentatività trimestrale di un insieme ristretto di aggregati (richiesto dalla Contabilità Nazionale) e, nell'altro (R2), la rappresentatività annuale di un più ampio dettaglio di spesa (Figura 2.16).

Figura 2.16 – Stima della spesa trimestrale e annuale (R1 e R2) per le spese rilevate con periodo di riferimento di 12 mesi



Stima di tipo R1 (trimestrale):

- per le spese rilevate con periodo di riferimento di 14 giorni (diario), la spesa mensile di ogni famiglia si ottiene moltiplicando per 2,14 la spesa rilevata nei 14 giorni;
- per le spese rilevate con riferimento all'ultimo mese, la spesa mensile è pari a quella rilevata;
- per le spese rilevate con riferimento agli ultimi 3 mesi, la spesa mensile si ottiene dividendo per tre quella rilevata;
- per le spese rilevate con riferimento agli ultimi 12 mesi, si considerano solo quelle sostenute negli ultimi tre mesi e la spesa mensile si ottiene dividendole per tre.

Stima di tipo **R2** (annuale):

La spesa mensile si ottiene:

- per le spese rilevate con periodo di riferimento di 14 giorni (diario), la spesa mensile di ogni famiglia si ottiene moltiplicando per 2,14 la spesa rilevata nei 14 giorni;
- per le spese rilevate con riferimento all'ultimo mese, la spesa mensile è pari a quella rilevata;
- per le spese rilevate con riferimento agli ultimi 3 mesi, la spesa mensile si ottiene dividendo per tre quella rilevata;
- per le spese rilevate con riferimento agli ultimi 12 mesi, la spesa mensile si ottiene dividendo per dodici quella rilevata.

I due metodi forniscono stime diverse, in altre parole la media mensile fornita alla Contabilità Nazionale ottenuta attraverso R1 non è direttamente riconducibile alla media mensile ottenuta annualmente attraverso R2 (Prospetto 2.10). La differenza riguarda esclusivamente le voci rilevate con periodo di riferimento a 12 mesi: nel calcolo delle stime annuali vengono considerate tutte le spese osservate, mentre nelle stime trimestrali si considerano solo quelle effettuate nei tre mesi precedenti l'intervista.

Prospetto 2.10 – Stima della spesa media mensile per tipo di metodo applicato

	Stima di tipo R1	Stima di tipo R2
Stima della spesa media mensile per la voce i rilevata con periodo di riferimento r (diario a 14 giorni)	$Sm^i = \frac{\sum_{j=1}^p S_{j,r}^i \cdot (2,14) \cdot (w_j)}{\sum_{j=1}^p (w_j)}$	$Sm^i = \frac{\sum_{j=1}^p S_{j,r}^i \cdot (2,14) \cdot (w_j)}{\sum_{i=1}^p (w_j)}$
Stima della spesa media mensile per la voce i rilevata con periodo di riferimento r (mensile o trimestrale)	$Sm^i = \frac{\sum_{j=1}^p (S_{j,r}^i / k) \cdot (w_j)}{\sum_{j=1}^p (w_j)}$ <i>dove $k=1$ se periodo mensile; 3 se trimestrale</i>	$Sm^i = \frac{\sum_{j=1}^p (S_{j,r}^i / k) \cdot (w_j)}{\sum_{j=1}^p (w_j)}$ <i>dove $k=1$ se periodo mensile; 3 se trimestrale</i>
Stima della spesa media mensile per la voce i rilevata con periodo di riferimento r (12 mesi precedenti l'intervista)	$Sm^i = \frac{\sum_{j=1}^p S_{j,r}^i \cdot (w_j)}{\sum_{j=1}^p (w_j)} / 3$ <i>dove $r=1-3$ (i soli 3 mesi precedenti l'intervista)</i>	$Sm^i = \frac{\sum_{j=1}^p S_{j,r}^i \cdot (w_j)}{\sum_{j=1}^p (w_j)} / 12$ <i>dove $r=1-12$ (i 12 mesi precedenti l'intervista)</i>
<i>Sm^i è la spesa media mensile per il bene i</i>		
<i>$S_{j,r}^i$ è la spesa della famiglia j per il bene i nel periodo r</i>		
<i>w_j è il peso di riporto all'universo della famiglia j</i>		
<i>p è il numero di famiglie campione</i>		

2.5.2 Il sistema di ponderazione: scelta dei vincoli e trattamento degli outliers

Dato il carattere campionario dell'indagine, per poter ottenere stime per l'intera popolazione di riferimento (le famiglie residenti) è necessario utilizzare opportuni coefficienti di riporto all'universo. Il valore di tale coefficiente indica il numero di unità nella popolazione rappresentate dall'unità campionaria. L'indagine produce stime trimestrali (con metodo R1) e stime annuali (con metodo R2).

Differentemente dall'indagine PAPI, i coefficienti di riporto all'universo per i dati annuali non si ottengono come semplice media aritmetica dei pesi trimestrali, ma tengono conto degli aggiornamenti anagrafici sulla popolazione residente.



Il calcolo dei coefficienti di riporto all'universo, siano essi trimestrali o annuali, utilizza una procedura generalizzata basata sull'uso di una famiglia di stimatori noti in letteratura come *calibration estimator* (stimatori di ponderazione vincolata). La metodologia alla base di tali stimatori consente la determinazione di un unico coefficiente di riporto all'universo in grado di produrre stime coerenti a totali noti, desunti da fonti esterne, che siano correlati alle principali variabili oggetto di indagine.

La famiglia di stimatori di ponderazione vincolata coincide asintoticamente con lo stimatore di regressione generalizzato e, per campioni sufficientemente grandi, ha approssimativamente le stesse proprietà, ovvero si tratta di stimatori corretti, consistenti e con la stessa varianza campionaria²⁹.

La strategia adottata per la costruzione dei coefficienti di riporto all'universo è simile a quella utilizzata dalle varie indagini campionarie sulle famiglie dell'Istituto e si sviluppa in tre fasi:

- la determinazione della probabilità di inclusione di ogni unità statistica e del relativo peso diretto, pari all'inverso della probabilità di inclusione;
- calcolo dei coefficienti di correzione per mancata risposta totale;
- determinazione dei coefficienti di riporto all'universo finali vincolati ai totali noti desunti da fonti esterne all'indagine.

2.5.2.1 La probabilità di inclusione e il peso diretto

Il principio su cui è basato ogni metodo di stima campionaria è che le unità appartenenti al campione rappresentino anche le unità della popolazione non incluse nel campione stesso. Per questo motivo ad ogni unità campionaria è possibile attribuire un peso, il coefficiente di riporto all'universo, che indica quante unità della popolazione essa rappresenta.

Senza perdere di generalità, definiamo la seguente simbologia:

- U popolazione di riferimento oggetto di indagine;
- y_k valore della variabile Y assunto dalla k -esima osservazione della popolazione;
- y_j valore della variabile Y assunto dalla j -esima osservazione della popolazione;
- j probabilità, assegnata dal disegno di campionamento, che l'unità j -esima sia inclusa nel campione S ;

Il totale di una generica variabile Y , calcolato sull'intera popolazione, assume la seguente forma:

$$Y = \sum_{k \in U} y_k \quad (1)$$

Il disegno di campionamento assegna le probabilità di inclusione ad ogni unità del campione in modo tale che

²⁹ La metodologia è illustrata da Deville, J.C. e Särndal, C.E. in *Calibration Estimation in Survey Sampling*, Journal of the American Statistical Association, Vol. 87, n.418, 1992.

$$\hat{Y} = \sum_{j \in S} y_j \frac{1}{\pi_j} \quad (2)$$

sia uno stimatore corretto della (1).

Nel disegno di campionamento dell'Indagine sulle spese, la probabilità di inclusione di un generico individuo è data dalla probabilità di estrazione del suo comune di residenza (direttamente proporzionale all'ampiezza demografica dei comuni all'interno dello strato) e dalla probabilità di estrazione della sua famiglia tra le famiglie eleggibili del comune. Per la natura del disegno campionario, la probabilità di inclusione, e dunque il peso base, è uguale per tutte le famiglie appartenenti allo stesso comune.

Per una generica famiglia eleggibile j , nel comune i dello strato h , il peso diretto d_{hij} , inverso della probabilità di inclusione π_{hij} , assume la seguente forma:

$$d_{hij} = \frac{1}{\pi_{hij}} = \frac{P_h}{P_{hi}} \frac{M_{hi}}{m_{hi}} \quad (3)$$

dove:

- h denota l'indice di strato;
- i è l'indice di comune;
- j denota l'indice della famiglia;
- P_h indica il totale della popolazione residente nello strato h ;
- P_{hi} indica il totale della popolazione residente nel comune i dello strato h ;
- M_{hi} indica il totale di famiglie eleggibili nel comune i dello strato h ;
- m_{hi} indica il numero di famiglie campione nel comune i dello strato h .

I pesi base vengono di volta in volta calcolati sui dati anagrafici più aggiornati per popolazione e famiglie residenti.

2.5.2.2 La correzione per mancata risposta

Nel corso della fase di raccolta delle informazioni, come sempre accade nelle indagini, alcune unità campionarie non partecipano (per irreperibilità, rifiuto, interruzione definitiva o altri motivi). Tale fenomeno, noto come mancata risposta totale, implica che la numerosità campionaria teorica e quella effettivamente raggiunta siano diverse. Nella fase di calcolo dei coefficienti di riporto all'universo si tiene conto di tale fenomeno e viene introdotto un correttore per mancata risposta totale: sotto l'ipotesi che in determinati gruppi (ottenuti come aggregazioni di strati appartenenti alla stessa regione e tipologia comunale) il comportamento dei rispondenti sia simile a quello dei non rispondenti, il correttore assume la forma dell'inverso del tasso di risposta (δ_g):

$$\frac{1}{\delta_g} = \frac{m_g}{m_g^r} \quad (4)$$

in cui m_g^r rappresenta il numero di famiglie rispondenti nel determinato gruppo g .

Il coefficiente di riporto all'universo, corretto per mancata risposta, per una generica famiglia eleggibile j , nel comune i dello strato h , del gruppo g (dato come intersezione tra regione e tipologia comunale), risulta pertanto essere:

$$k_{ghij} = d_{hij} \frac{1}{\delta_g} = \frac{P_h}{P_{hi}} \frac{M_{hi}}{m_{hi}} \frac{m_g}{m_g^r} \quad (5)$$

2.5.2.3 La calibrazione a fonti esterne

Per il calcolo dei coefficienti di riporto all'universo finali si adottano gli stimatori di ponderazione vincolata (*calibration estimator*). La metodologia si basa sull'utilizzo di informazioni ausiliarie, sintetizzate in totali noti, che, essendo correlate con le variabili principali oggetto di indagine, hanno la funzione di aumentare l'accuratezza delle stime. I pesi finali si ottengono risolvendo un problema di minimo vincolato, in cui la funzione da minimizzare è una funzione di distanza tra i pesi diretti corretti per la mancata risposta (\mathbf{k}) e i pesi finali (\mathbf{w}) delle famiglie del campione rispondente (); i vincoli sono proprio le condizioni di uguaglianza tra le stime campionarie delle variabili ausiliarie e i rispettivi totali noti desunti da fonti esterne all'indagine³⁰. Formalmente:

$$\begin{cases} \text{Min} \left\{ \sum_{j \in S^r} \text{dist}(k_j, w_j) \right. \\ \left. \sum_{j \in S^r} x_j * w_j = t \right. \end{cases} \quad (5)$$

dove t è il vettore dei totali noti e x_j è il vettore delle variabili ausiliarie osservate sulla j -esima unità campionaria appartenente al campione rispondente (S^r). La funzione di distanza utilizzata è la logaritmica troncata.

I totali noti, introdotti come vincoli nel calcolo dei pesi finali, consentono di migliorare l'accuratezza delle stime: quanto più le variabili ausiliarie considerate sono correlate con le variabili oggetto d'indagine, tanto più si riduce la distorsione delle stime. Nello specifico dell'Indagine sulle spese delle famiglie, le stime campionarie sono state vincolate ai seguenti totali noti:

1. Popolazione residente per ripartizione, sesso e classi di età (0-14, 15-29, 30-44, 45-59, 60-74, 75 e più);
2. Popolazione residente per regione;
3. Famiglie residenti per regione;
4. Popolazione residente per ripartizione e tipologia di comune (Comuni metropolitan, comuni periferia dei comuni metropolitan e comune oltre 50.000 abitanti, altri comuni);

³⁰ La calibrazione è una calibrazione integrata, vale a dire che i coefficienti di riporto all'universo sono stati determinati in modo da poter essere utilizzati, indifferentemente, per la costruzione delle stime relative alle famiglie e agli individui. Tutti gli individui di una stessa famiglia hanno il medesimo coefficiente di riporto della famiglia stessa.

5. Famiglie residenti per ripartizione e tipologia di comune (Comuni metropolitani, comuni della cintura metropolitana e comuni con più di 50.000 abitanti, altri comuni);
6. Popolazione straniera residente per ripartizione e sesso;
7. Popolazione di 15 anni e più per ripartizione, condizione e posizione nella professione (Dirigenti, quadri e impiegati; operai e assimilati; imprenditori e liberi professionisti; lavoratori in proprio e altri autonomi; disoccupati; inattivi);
8. Popolazione per ripartizione e mese;
9. Famiglie per ripartizione e mese.

I vincoli ai punti 1, 2 e 3 sono gli stessi utilizzati nella vecchia indagine sui consumi delle famiglie (al netto delle classi 30-44 anni e 45-59 anni che prima erano unite in un'unica classe d'età).

I vincoli ai punti 4 e 5 sono stati inseriti per rispettare il nuovo disegno di campionamento e considerare i differenti livelli e stili di spesa associati alle diverse tipologie di comune di residenza.

Il vincolo al punto 6 è stato inserito per tener conto del peso che, nel corso del tempo, ha assunto la popolazione straniera sul totale della popolazione residente, e del fatto che questa sottopopolazione ha stili e livelli di consumo generalmente differenti³¹.

Il vincolo al punto 7 è stato introdotto in considerazione del fatto che la condizione e la posizione professionale impattano sui livelli di reddito, e quindi di spesa, della popolazione. Questo totale è desunto dalla distribuzione proveniente dall'indagine Istat sulle forze di lavoro.

I vincoli 8 e 9 permettono di ripartire uniformemente la popolazione e le famiglie nei diversi mesi (tre mesi nel trimestre o dodici nell'anno), al fine di tenere conto del carattere fortemente stagionale di alcune tipologie di spesa.

I vincoli imposti sono gli stessi sia per le stime trimestrali che per le stime annuali. La differenza è che per le stime annuali le informazioni sui totali noti sono più aggiornate. Questo vale sia per le informazioni anagrafiche su popolazione e famiglie, ma anche per i totali provenienti dall'Indagine sulle forze di lavoro. Per motivi di disponibilità temporale, infatti, nelle stime trimestrali i vincoli provenienti dall'Indagine sulle forze di lavoro sono sfasati di un trimestre. Ad esempio, per il secondo trimestre di un determinato anno dell'Indagine sulle spese si fa riferimento al primo trimestre dello stesso anno dell'Indagine sulle forze di lavoro. Questo non vale per le stime annuali, per le quali si utilizza come totale noto la media d'anno dell'Indagine sulle forze di lavoro.

Come già nella vecchia Indagine sui consumi delle famiglie, si è adottata una procedura di individuazione delle osservazioni anomale, vale a dire osservazioni con livelli e comportamenti di spesa (calcolata con metodo R2) che, seppur veri, sono molto lontani o dissimili dal collettivo socio-economico al quale appartengono. Tali unità campionarie rappresentano quindi un numero ristretto di famiglie dell'universo ed è quindi opportuno che abbiano associato un coefficiente di riporto abbastanza basso per non influenzare la stima dal punto di vista sia della distorsione che della

³¹ Senza vincolo, la popolazione straniera risulterebbe ampiamente sottostimata. Per diversi motivi, infatti, le famiglie con stranieri hanno generalmente tassi di risposta più bassi rispetto alle famiglie di soli italiani.

varianza. I gruppi socio-economici all'interno dei quali individuare comportamenti di spesa anomali sono stati definiti dall'incrocio tra ripartizione di appartenenza, tipologia di comune e numerosità familiare; all'interno di ogni gruppo, le famiglie anomale vengono individuate tramite una *cluster analysis* che considera sia il livello complessivo di spesa che la sua composizione tra i diversi capitoli. Alle famiglie individuate come anomale viene attribuito, nella fase precedente alla calibrazione finale, un peso corretto per la mancata risposta più basso rispetto a quello che avrebbero avuto tramite la procedura generalizzata. Questa procedura, simile a quanto già in uso per l'indagine PAPI³², permette di mantenere nel campione le osservazioni anomale e rende possibile l'uso degli stimatori classici. Il numero di osservazioni anomale è comunque limitato, in media meno di 50 famiglie a trimestre.

Rispetto all'indagine PAPI, il disegno di campionamento più fine, l'introduzione di nuovi vincoli di popolazione e la minore numerosità campionaria effettiva avrebbero comportato un campo di variazione dei coefficienti di riporto più ampio. Per non avere famiglie con coefficienti troppo elevati è stata quindi adottata una procedura di calibrazione ricorsiva, individuando le famiglie con valori dei coefficienti esterni a un intervallo di accettazione (inferiori al quinto percentile o superiori al 95esimo) e imponendo a queste valori dei coefficienti pari ai limiti del suddetto intervallo. La procedura di calibrazione viene nuovamente eseguita in modo da rispettare i vincoli imposti; in tale maniera, i valori estremi dei coefficienti di riporto sono sostanzialmente in linea con quanto accadeva nella vecchia indagine.

2.5.3 Errori di campionamento

Per valutare la precisione delle stime si ricorre al calcolo dell'errore di campionamento³³. In particolare, indicando con $\hat{V}ar({}_d\hat{Y})$ la varianza della stima riferita al dominio d , la stima dell'errore di campionamento assoluto di ${}_d\hat{Y}$ si può ottenere mediante la seguente espressione:

$$\hat{\sigma}({}_d\hat{Y}) = \sqrt{\hat{V}ar({}_d\hat{Y})} \quad (1)$$

La stima dell'errore di campionamento relativo di ${}_d\hat{Y}$ è definita dall'espressione:

$$\hat{\varepsilon}({}_d\hat{Y}) = \frac{\sqrt{\hat{V}ar({}_d\hat{Y})}}{{}_d\hat{Y}} \quad (2)$$

La stima della varianza, $\hat{V}ar({}_d\hat{Y})$, viene calcolata come somma della stima della varianza dei singoli strati appartenenti al dominio d . In simboli:

$$\hat{V}ar({}_d\hat{Y}) = \sum_{h=1}^{H_d} \hat{V}ar(\hat{Y}_h) = \sum_{h=1}^{H_d} M_h^2 \frac{M_h - m_h}{m_h M_h} \sum_{j=1}^{m_h} \frac{(\hat{Y}_{hj} - \hat{Y}_h)^2}{m_h - 1} \quad (3)$$

32 D'Aurizio L. et al. 1991. Acquisizione e trattamento dei dati nella nuova indagine sui consumi. Roma: Istat. (Documenti Istat, n. 21)

33 Il calcolo degli errori campionari è stato effettuato da Rosalba Bravi mediante il software Genesees.

dove

$$\hat{Y}_{hj} = Y_{hj} W_{hj} \quad (4)$$

e

$$\hat{Y}_h = \frac{1}{m_h} \sum_{j=1}^{m_h} \hat{Y}_{hj} \quad (5)$$

Gli errori campionari espressi nelle formule (1) e (2) consentono di valutare il grado di precisione delle stime. L'errore assoluto permette di costruire l'intervallo di confidenza, che, con una certa probabilità p , contiene il parametro d'interesse. Con riferimento ad una generica stima tale intervallo assume la seguente forma:

$$\Pr\left(\hat{Y} - k\hat{\sigma}(\hat{Y}) \leq Y \leq \hat{Y} + k\hat{\sigma}(\hat{Y})\right) = p \quad (6)$$

Nella (6) il valore di k dipende dal valore fissato per la probabilità p ; ad esempio, per $p=0,95$ si ha $k=1,96$.

Nella Tavola 2.6 sono riportati gli errori campionari relativi, in percentuale, per l'anno 2014.

Tavola 2.6 – Errore relativo della spesa media mensile familiare per capitolo di spesa - Anno 2014 (in percentuale)

CAPITOLI DI SPESA	Errore relativo
	(%)
Pane e cereali	0,67
Carni	0,75
Pesci e prodotti ittici	1,17
Latte, formaggi e uova	0,69
Oli e grassi	1,49
Frutta	0,88
Vegetali	0,87
Zucchero, confetture, miele, cioccolato e dolciumi	1,17
Prodotti alimentari n.a.c.	2,34
Caffè, tè e cacao	1,28
Acque minerali, bevande analcoliche, succhi di frutta e verdura	1,10
Prodotti alimentari e bevande analcoliche	0,62
Bevande alcoliche e tabacchi	1,40
Abbigliamento e calzature	1,85
Abitazione, acqua, elettricità e altri combustibili	0,58
Manutenzione straordinaria	6,60
Fitti figurativi	0,79
Mobili, articoli e servizi per la casa	1,97
Servizi sanitari e spese per la salute	1,81
Trasporti	1,30
Comunicazioni	0,76
Ricreazione, spettacoli e cultura	1,45
Istruzione	3,50
Servizi ricettivi e di ristorazione	1,52
Altri beni e servizi	1,38
Beni e servizi non alimentari	0,58
Totale	0,52

La stima della spesa media mensile presenta un errore relativo pari allo 0,52%. In generale, gli errori relativi più contenuti si osservano per le spese affrontate più frequentemente dalle famiglie, abitazione (0,58%) e alimentari (0,62%); le spese con errori più elevati si osservano per l'istruzione (3,50%).

Nella Tavola 2.7, l'errore di campionamento della stima della spesa media mensile per i capitoli di spesa viene riportato per ripartizione geografica.

La minore numerosità del campione di famiglie nelle singole ripartizioni rende le stime meno precise rispetto al dato nazionale; in particolare l'errore relativo della spesa media mensile risulta più elevato nelle Isole (1,56%) e più contenuto nel Sud (0,88%) e nel Nord-est (0,98%).

Tavola 2.7 – Errore relativo della spesa media mensile familiare per capitolo di spesa e ripartizione geografica – Anno 2014 (in percentuale)

CAPITOLI DI SPESA	Nord-ovest	Nord-est	Centro	Sud	Isole
Pane e cereali	1,44	1,28	1,44	1,26	2,18
Carni	1,62	1,51	1,76	1,39	2,19
Pesci e prodotti ittici	2,60	2,83	2,67	1,91	3,36
Latte, formaggi e uova	1,43	1,25	1,51	1,34	2,80
Oli e grassi	3,01	2,75	3,23	2,57	6,41
Frutta	1,88	1,76	2,05	1,55	2,76
Vegetali	1,75	1,60	2,06	1,55	3,38
Zucchero, confetture, miele, cioccolato e dolci	2,36	2,32	2,80	2,14	3,88
Prodotti alimentari n.a.c.	3,47	3,62	7,72	3,63	11,61
Caffè, tè e cacao	2,92	2,42	3,28	2,04	3,10
Acque minerali, bevande analcoliche, succhi di frutta e verdura	2,27	2,35	2,53	1,99	3,18
Prodotti alimentari e bevande analcoliche	1,28	1,19	1,46	1,16	2,28
Bevande alcoliche e tabacchi	2,82	3,02	3,21	2,69	4,26
Abbigliamento e calzature	3,94	3,10	3,80	2,46	7,35
Abitazione, acqua, elettricità e altri combustibili	1,22	1,20	1,23	1,01	1,61
Manutenzione straordinaria	12,94	11,68	14,39	14,42	23,35
Fitti figurativi	1,82	1,40	1,59	1,50	1,97
Mobili, articoli e servizi per la casa	4,04	3,63	5,33	3,19	4,58
Servizi sanitari e spese per la salute	4,21	3,37	3,24	2,68	4,03
Trasporti	2,70	2,44	2,67	2,27	4,08
Comunicazioni	1,57	1,52	1,72	1,40	2,73
Ricreazione, spettacoli e cultura	2,70	2,76	3,53	2,49	5,11
Istruzione	7,58	5,21	8,14	5,74	11,07
Servizi ricettivi e di ristorazione	2,86	2,70	3,38	3,18	5,63
Altri beni e servizi	3,36	2,21	2,55	2,15	2,78
Beni e servizi non alimentari	1,23	1,06	1,33	0,97	1,59
Totale	1,10	0,98	1,21	0,88	1,56

A livello regionale (Tavola 2.8), gli errori più contenuti della spesa media mensile familiare si osservano in Campania (1,52%), in Veneto (1,61%) e in Lombardia (1,64%),

2. Il disegno della nuova indagine

111

mentre nella provincia di Trento (3,51%) e in quella di Bolzano (5,40%), in Umbria (3,08), in Molise (3,55%), in Basilicata (3,84%) e in Calabria (3,04) l'errore relativo è superiore al 3%.

Tavola 2.8 – Errore relativo della spesa media mensile familiare per regione, prodotti alimentari e bevande analcoliche, beni e servizi non alimentari, totale - Anno 2014 (in percentuale)

REGIONE	Prodotti alimentari e bevande analcoliche	Beni e servizi non alimentari	Totale
Piemonte	2,24	1,94	2,28
Valle d'Aosta/Vallée d'Aoste	3,18	2,48	2,67
Lombardia	1,77	1,49	1,64
Trentino-Alto Adige	4,33	2,85	2,96
<i>Trento</i>	3,16	3,15	3,51
<i>Bolzano</i>	7,37	5,25	5,40
Veneto	1,53	1,46	1,61
Friuli-Venezia Giulia	3,10	2,63	2,82
Liguria	2,67	2,42	2,65
Emilia-Romagna	2,20	1,69	1,83
Toscana	2,33	2,26	2,44
Umbria	3,29	2,55	3,08
Marche	2,53	2,43	2,67
Lazio	2,38	1,85	2,02
Abruzzo	2,87	2,18	2,40
Molise	4,53	3,54	3,55
Campania	2,14	1,39	1,52
Puglia	2,01	1,75	1,93
Basilicata	3,68	3,57	3,84
Calabria	2,54	2,64	3,04
Sicilia	2,87	1,94	1,97
Sardegna	2,78	2,40	2,70

2.6 La riconciliazione tra indagine PAPI e indagine CAPI

2.6.1 Il raccordo tra le voci di spesa

L'ampliamento dei contenuti informativi e la nuova classificazione delle voci di spesa hanno richiesto l'implementazione di tavole di corrispondenza tra indagine PAPI e indagine CAPI a fini comparativi e di ricostruzione di serie storica.

La classificazione delle voci adottata nella indagine CAPI si compone di 486 voci, a fronte delle 263³⁴ rilevate con la PAPI.

34 A tale numero si aggiungono 12 voci con cui venivano rilevate le spese per manutenzione straordinaria dell'abitazione principale e delle secondarie che non sono incluse nella Coicop.



La nuova classificazione, attraverso opportune aggregazioni, riesce a soddisfare le esigenze informative dei principali utenti – da Eurostat, al servizio Prezzi, alla Contabilità Nazionale – che fanno riferimento alla Coicop più recente. D'altra parte, molti utenti esterni non di rado necessitano di far riferimento a segmenti specifici di spesa, richiedendo quindi la massima disaggregazione possibile.

Le tavole di corrispondenza elaborate definiscono i raccordi tra le due indagini e quelli tra ciascuna indagine e la classificazione Coicop 2013:

Tavola di corrispondenza di Tipo 1	PAPI – CAPI
Tavola di corrispondenza di Tipo 2	CAPI – Coicop 2013
Tavola di corrispondenza di Tipo 3	PAPI – Coicop 2013

La tavola di corrispondenza di Tipo 1 (Tavola 2.9) contiene 265 aggregati, tra i quali 218 sono comuni ad entrambe le indagini. In particolare, 107 si ottengono dalla corrispondenza biunivoca di singole voci (una voce CAPI con una voce PAPI), 115 dalla corrispondenza di un diverso numero di voci (garantendo comunque la confrontabilità) e 43 si riferiscono a beni o servizi rilevati con la CAPI che nella PAPI non erano espressamente indicati (ad es. alcuni nuovi prodotti tecnologici).

Tavola 2.9 – Sintesi della tavola di corrispondenza di Tipo 1 – Anno 2013

Tipo di abbinamento	Numero di aggregati
1 voce nella CAPI – 1 voce nella PAPI	107
1 voce nella CAPI – Più voci nella PAPI	16
Più voci nella CAPI – 1 voce nella PAPI	75
Più voci nella CAPI – Più voci nella PAPI	24
1 voce nella CAPI – 0 voci nella PAPI	43
Totale	265

La tavola di corrispondenza di Tipo 2, a partire dai dati dell'indagine CAPI ricostruisce la classificazione Coicop 2013 fino a cinque *digit*; tuttavia 11 categorie Coicop, per diversi motivi, non sono state rilevate con il dettaglio richiesto (Prospetto 2.11). Si tratta delle voci: “test di gravidanza e contraccettivi di barriera” (come singola voce avrebbe rappresentato un quesito sensibile ed è quindi stata inserita tra le spese per “prodotti farmaceutici”), “trasporto per vie d'acqua” (non si è distinto tra vie d'acqua interne e marittime), “servizi pubblici di assicurazione connessi con la salute” (non esistenti in Italia), “birra” (in quest'ultimo caso la Coicop prevede quattro voci in base al contenuto alcolico – Birre lager, altre birre alcoliche, birre a basso contenuto alcolico o non alcoliche, altre bevande a base di birra – e la risposta ai singoli quesiti avrebbe richiesto alla famiglia uno sforzo veramente eccessivo). Lo stesso motivo ha portato a non separare i costi sostenuti per i tessuti per la casa e per gli effetti personali da quelli sostenuti per eventuali riparazioni (spesso la spesa è contestuale). Infine, non si rilevano le spese per narcotici e quelle per prostituzione, per motivi legislativi o di riprovazione sociale, e quelle per *FISIM* (*Financial Intermediation Services Indirectly Measured*- Servizi Finanziari di Intermediazione Misurati Indirettamente), che non rientrano nel computo della spesa per consumi dal lato HBS. In definitiva, dunque, si ottiene una classificazione di 292 categorie, 114 classi, 45 gruppi e 12 divisioni (Tavola 2.10).

2. Il disegno della nuova indagine

Prospetto 2.11 – Voci non rilevate al 5 digit con l'indagine sulle spese delle famiglie – Anno 2013

Coicop 2013	Indagine sulle spese delle famiglie 2013
02.1.3.2 Other alcoholic beer	Incluso nel codice 02.1.3.1 (Lager beer)
02.1.3.3 Low and non-alcoholic beer	Incluso nel codice 02.1.3.1 (Lager beer)
02.1.3.4 Beer-based drinks	Incluso nel codice 02.1.3.1 (Lager beer)
02.3.0.0 Narcotics	Non rilevato
05.2.0.4 Repair of household textiles	and curtains)
06.1.2.1 Pregnancy tests and mechanical contraceptive devices	Incluso nel codice 06.1.2.9 (Other medical products n.e.c.)
07.3.4.2 Passenger transport by inland waterway	Incluso nel codice 07.3.4.1 (Passenger transport by sea)
12.2.0.0 Prostitution	Non rilevato
12.3.2.3 Repair of other personal effects	Incluso nel codice 12.3.2.9 (Other personal effects n.e.c.)
12.5.3.1 Public insurance connected with health	Non rilevato
12.6.1.0 FISIM	Non rilevato

Tavola 2.10 – Sintesi della tavola di corrispondenza di Tipo 2. Anno 2013

Divisione	Coicop 2013			Indagine CAPI 2013			Voci rilevate
	Gruppi	Classi	Categorie	Gruppi	Classi	Categorie	
01 Prodotti alimentari e bevande analcoliche	2	11	61	2	11	60	92
02 Bevande alcoliche e tabacchi	3	5	14	2 (a)	4 (a)	10(b)	10
03 Abbigliamento e calzature	2	6	12	2	6	12	17
04 Abitazione, acqua, elettricità e altri combustibili	5	15	25	5	15	25	74
05 Mobili, articoli e servizi per la casa	6	12	40	6	12	40	54
06 Servizi sanitari e spese per la salute	3	7	14	3	7	13 (c)	15
07 Trasporti	3	14	28	3	14	27 (d)	41
08 Comunicazioni	3	3	11	3	3	11	20
09 Ricreazione, spettacoli e cultura	6	21	53	6	21	53	82
10 Istruzione	5	5	6	5	5	6	20
11 Servizi ricettivi e di ristorazione	2	3	6	2	3	6	19
12 Altri beni e servizi	7	15	33	6(e)	13 (f)	29 (f)	42
Totale	47	117	303	45	114	292	486

(a) Esclusi i narcotici

(b) Esclusi i narcotici e unificate le 4 voci della birra

(c) Unificate 2 voci di articoli sanitari

(d) Unificate le 2 voci del trasporto su acqua

(e) Esclusa la prostituzione

(f) Escluse la prostituzione, l'assicurazione pubblica connessa con la salute e la riparazione di altri effetti personali inclusa nell'acquisto.

La tavola di corrispondenza di Tipo 3 è decisamente diversa. Le categorie ricostruibili a 5 *digit* sono solo 155 su 303. Le altre categorie, infatti, non possono essere ricostruite in quanto la struttura del questionario PAPI non seguiva la più recente classificazione europea: 28 categorie non erano oggetto di rilevazione e 117 venivano rilevate in maniera aggregata (una sola voce PAPI si riferiva a 2 o più codici Coicop). Di conseguenza 9 delle 114 classi³⁵ (3 digit) non possono essere identificate e la ricostruzione totale si limita quindi a 45 gruppi (3 digit) (Tavola 2.11).

35 Prodotti per la riparazione e la manutenzione della casa, Raccolta rifiuti, Trasporto multimodale passeggeri, Riparazione e manutenzione dei beni durevoli per ricreazione al coperto, Articoli sportivi, per campeggio e attività ricreative all'aperto, Servizi veterinari e altri servizi per animali domestici, Materiale stampato vario, Istruzione per ordine e grado scolastico, Altri servizi finanziari n.a.c.

Tavola 2.11– Sintesi della tavola di corrispondenza di TIPO 3 - Anno 2013

Divisione		Coicop 2013	Indagine PAPI 2013	
		Gruppi	Gruppi	Voci rilevate
01	Prodotti alimentari e bevande analcoliche	2	2	59
02	Bevande alcoliche e tabacchi	3	2	4
03	Abbigliamento e calzature	2	2	15
04	Abitazione, acqua, elettricità e altri combustibili	5	5	30
05	Mobili, articoli e servizi per la casa	6	6	37
06	Servizi sanitari e spese per la salute	3	3	14
07	Trasporti	3	3	23
08	Comunicazioni	3	3	9
09	Ricreazione, spettacoli e cultura	6	6	38
10	Istruzione	5	5	3
11	Servizi ricettivi e di ristorazione	2	2	9
12	Altri beni e servizi	7	6	22
Totale		47	45	263

L'utilizzo di una classificazione molto fine in una indagine campionaria non consente di assicurare significatività statistica alle stime, soprattutto quando le variabili di classificazione prevedono un numero elevato di modalità (Tavola 2.12). In tal senso, le categorie più problematiche sono, oltre ai Servizi di posa di moquettes e linoleum, le spese per grandi articoli e mezzi di trasporto, per lo sport e per lo svago (veicoli a trazione animale e altri veicoli; autocaravan, caravan e rimorchi; aeroplani, velivoli ultraleggeri, alianti e mongolfiere; imbarcazioni, motori fuoribordo ed equipaggiamento per imbarcazioni; cavalli e pony e equipaggiamento collegato; grandi articoli per gioco e sport) che, anche a livello nazionale, presentano una numerosità campionaria (numero di famiglie che hanno effettuato la spesa) decisamente ridotta. Ovviamente il problema della rappresentatività si estende all'aumentare del dettaglio di analisi: se a livello nazionale le categorie con problemi di robustezza sono solo 7, quando si considera la regione il numero sale a 181, per arrivare a problemi a livello di divisione (spese per istruzione) quando si considera la tipologia familiare. Di ciò è dunque opportuno tenere conto in fase di analisi e di diffusione dell'informazione stabilendo di volta in volta l'opportuno grado di disaggregazione rilasciabile.

Tavola 2.12 – Numero di variabili con numerosità campionaria inferiore a 20 per variabile di classificazione Coicop – Anno 2013

Variabile di classificazione	Divisioni	Gruppi	Classi	Categorie
Italia	0	0	1	7
Ripartizione (3 modalità)	0	0	8	41
<i>di cui per tutte le modalità</i>	0	0	1	13
Ripartizione (5 modalità)	0	0	14	85
<i>di cui per tutte le modalità</i>	0	0	4	21
Ampiezza familiare (5 modalità)	0	3	11	76
<i>di cui per tutte le modalità</i>	0	1	3	20
Quinto di spesa equivalente (5 modalità)	0	2	11	71
<i>di cui per tutte le modalità</i>	0	1	2	20
Tipologia familiare (12 modalità)	1	11	47	180
<i>di cui per tutte le modalità</i>	0	0	1	7
Regione (21 modalità)	0	13	49	181
<i>di cui per tutte le modalità</i>	0	1	6	47
Totale voci	12	45	114	292

2.6.2 Il raccordo dei periodi di riferimento

Con la tavola di corrispondenza di Tipo 1 è stato creato un raccordo definitorio tra le voci della indagine PAPI e quelle dell'indagine CAPI. Tuttavia ciò non è sufficiente a far "parlare" le due indagini, in quanto è stato necessario considerare anche le altre variazioni introdotte nel disegno. In particolare, con la nuova indagine CAPI i periodi di riferimento risultano generalmente più ampi che nella PAPI; fanno eccezione 2 voci che transitano dall'intervista finale al diario (spese per confezionare e riadattare abiti nuovi e usati, riparazione di capi di abbigliamento) e 36 voci per le quali il periodo di riferimento rimane inalterato (Tavola 2.13).

È stato dunque necessario creare un'ulteriore tavola di corrispondenza (di Tipo 4) per raccordare gli aggregati anche rispetto al periodo di riferimento; a partire dalla tavola di corrispondenza di Tipo 1, le voci che compongono ciascun aggregato, sia della PAPI sia della CAPI, sono state riportate al periodo di riferimento (R3) minimo (in genere quello utilizzato nell'indagine PAPI), in modo da ottenere stime della spesa media mensile direttamente confrontabili.

Tavola 2.13 – Classificazione degli aggregati in base ai periodi di riferimento – Anno 2013

Ampiezza del periodo di riferimento e tecnica di rilevazione	Numero di aggregati
Alimentari rilevati tramite diario (Da 7 a 14 giorni)	54
Non alimentari rilevati tramite diario (Da 7 a 14 giorni)	30
Non alimentari rilevati indagine diretta	107
<i>Stesso periodo di riferimento</i>	36
<i>Da ultimo mese a ultimi tre mesi</i>	20
<i>Da ultimo mese a ultimi dodici mesi</i>	11
<i>Da ultimi tre mesi a ultimi dodici mesi</i>	38
<i>Da ultimi tre mesi a ultimo pagamento</i>	2
Non alimentari rilevati con tecnica diversa nelle due indagini	27
Aggregati non comuni alle due indagini	47
Totale	265

Per ogni voce rilevata nella CAPI con periodo di riferimento r , si osservano solo gli eventi accaduti nel periodo di riferimento utilizzato nella PAPI (\tilde{r}); la spesa associata a tali eventi si ottiene dividendo la spesa totale sostenuta per il numero di eventi nel periodo CAPI (k) e moltiplicandola per il numero di eventi nel periodo PAPI (q). Ad esempio, la spesa per riscaldamento centralizzato nella CAPI viene rilevata con riferimento agli ultimi 12 mesi, mentre nella PAPI si chiedeva rispetto agli ultimi 3 mesi. Si supponga che, nella CAPI, la spesa sostenuta per il riscaldamento centralizzato sia stata effettuata per 10 mesi, di cui 2 nell'ultimo trimestre; il riporto al periodo PAPI si ottiene dividendo la spesa totale per il numero di eventi spesa nel periodo r (nell'esempio il valore è pari a 10) e moltiplicandola per il numero di eventi spesa in \tilde{r} (nell'esempio il valore è pari a 2). La spesa media mensile, che può essere direttamente confrontata con quella ottenuta dalla PAPI, si ottiene dividendo l'ammontare così ottenuto per \tilde{r} (nel caso specifico il valore ottenuto viene diviso per 3). Formalmente:

$$Sm^i = \frac{\sum_{j=1}^p \left(\frac{s_{j,r}^i}{k}\right) \cdot q \cdot w_j}{\sum_{j=1}^p w_j} \cdot \tilde{r}$$

2.6.3 Un quadro di sintesi

Il Prospetto 2.12 sintetizza le principali differenze tra l'indagine PAPI e l'indagine CAPI, ampiamente discusse nei paragrafi precedenti.

Prospetto 2.12 – Sintesi delle differenze tra indagine PAPI e indagine CAPI

	Indagine sui consumi delle famiglie 1997-2013	Indagine sulle spese delle famiglie 2012-2014
Schema di campionamento	Casuale a due stadi (comuni-famiglie) con stratificazione delle unità di primo stadio	Casuale a due stadi (comuni-famiglie) con stratificazione delle unità di primo stadio
Domini di studio territoriali	Ripartizione, Regione	Ripartizione, Tipologia del comune ³⁶ , Regione
Domini di studio temporali	Trimestre	Trimestre
Numerosità campionaria teorica	Circa 28 mila famiglie residenti in circa 480 comuni italiani, con sostituzione delle unità di secondo stadio	Circa 28 mila famiglie residenti in 502 comuni italiani, con sostituzione delle unità di secondo stadio
Periodicità della rilevazione	Continua, mensile	Continua, mensile
Unità di rilevazione	Famiglia di fatto definita come l'insieme di persone coabitanti e legate da vincoli di matrimonio, parentela, affinità, adozione, tutela o affettivi.	Famiglia di fatto definita come l'insieme delle persone coabitanti, legate da vincoli di matrimonio o parentela, affinità, adozione, tutela o affettivi e che partecipano alla spesa familiare e/o condividono il reddito familiare.
Persona di riferimento	Intestatario del foglio di famiglia anagrafico	Intestatario del foglio di famiglia anagrafico (è possibile individuare anche il principale percettore di reddito)
Rispondente	Tutti i componenti della famiglia; se uno dei componenti non può partecipare alla rilevazione, sono ammesse risposte proxy	Tutti i componenti della famiglia; se uno dei componenti non può partecipare alla rilevazione, sono ammesse risposte proxy
Tecnica di rilevazione	Mista in modalità PAPI (diario autocompilato e una intervista diretta)	Mista in modalità CAPI (diario cartaceo autocompilato e due interviste dirette)
Modalità di svolgimento della rilevazione	Per ciascuna famiglia selezionata sono previste due visite a domicilio da parte del rilevatore: <i>prima visita</i> - il rilevatore consegna il diario alla famiglia; <i>seconda visita</i> - nei primi giorni del mese successivo a quello di compilazione del diario, si effettua la visita finale per ritirare il diario e per svolgere un'intervista finale.	Per ciascuna famiglia selezionata sono previste tre visite a domicilio da parte del rilevatore: <i>prima visita</i> - il rilevatore effettua un'intervista iniziale e consegna il diario alla famiglia; <i>seconda visita</i> - a metà del periodo di compilazione del diario, viene effettuata una visita per verificare come procede la compilazione; <i>terza visita</i> - nei primi giorni del mese successivo a quello di compilazione del diario, si effettua la visita finale per ritirare il diario e per svolgere una seconda intervista (intervista finale).
Rete di rilevazione	Costituita dai rilevatori, selezionati dagli uffici di statistica dei Comuni, formati e coordinati dall'Istat, tramite i propri Uffici Regionali, con la supervisione del Servizio Condizioni economiche delle Famiglie	Costituita dai rilevatori, selezionati e coordinati dalla società incaricata e formati dal Servizio Condizioni economiche delle Famiglie dell'Istat
Variabile rilevata	Spesa per consumi delle famiglie intesa come l'insieme di tutte le spese sostenute dalle famiglie per acquistare beni e servizi destinati al diretto soddisfacimento dei bisogni dei propri componenti (inclusi gli autoconsumi, il fitto figurativo e i regali).	Spesa per consumi delle famiglie intesa come l'insieme di tutte le spese sostenute dalle famiglie per acquistare beni e servizi destinati al diretto soddisfacimento dei bisogni dei propri componenti (inclusi gli autoconsumi, il fitto figurativo e i regali), rilevata contestualmente al mese di effettuazione della spesa e al luogo di acquisto.
Classificazione delle voci di spesa	COICOP 1993	COICOP 2013

36 Area metropolitana, Periferia dell'area metropolitana e Comuni con 50.000 abitanti e più, Comuni con meno di 50.000 abitanti.

Prospetto 2.12 (segue) – Sintesi delle differenze tra indagine PAPI e indagine CAPI

	Indagine sui consumi delle famiglie 1997-2013	Indagine sulle spese delle famiglie 2012-2014
Periodo di riferimento delle variabili	7 giorni Alimentari, autoconsumi, servizi di ristorazione e mense, beni e servizi non alimentari correnti (110 voci)	14 giorni Alimentari, autoconsumi, servizi di ristorazione, beni e servizi non alimentari correnti (170 voci)
	Ultimo mese Abbigliamento, arredamenti, articoli per la casa, servizi domestici, sanità, vacanze, articoli sportivi e sport, cultura, istruzione, trasporti (90 voci)	Ultimo mese Abbigliamento, mense, articoli sportivi, articoli per la casa e piccoli elettrodomestici, manutenzione veicoli (54 voci)
	Ultimi 3 mesi Combustibili e utenze, manutenzioni abitazioni, beni durevoli, mezzi di trasporto e di comunicazione (61 voci)	Ultimi 3 mesi Arredamenti, servizi domestici, sanità, vacanze e servizi di trasporto (68 voci)
	Ultimi 12 mesi Spese eccezionali e assicurazione veicoli (2 voci)	Ultimi 12 mesi Combustibili e utenze, manutenzioni abitazioni, mezzi di trasporto e di comunicazione, istruzione, assicurazione veicoli, sport (194 voci)
Rilevazione delle Core variables Eurostat	Parziale	Completa
Rilevazione degli autoconsumi	Quantità consumata giornalmente e autovalutazione del suo valore monetario rilevati con diario separato consegnato alle sole famiglie che dichiarano di effettuare autoconsumo	Quantità consumata giornalmente rilevata in una sezione a parte del diario utilizzato per le spese. Valutazione monetaria a posteriori in base ai prezzi di mercato forniti dal servizio Prezzi dell'Istat.
Tecnica di correzione dei dati	Deterministica e probabilistica (SCIA-RIDA)	Deterministica e probabilistica (Procedura famiglie-SCIA-BANFF- IVEware)
Metodo di stima della spesa	R2	R1, R2 e R3

3. LA RICOSTRUZIONE DELLE STIME IN SERIE STORICA¹

Contestualmente alla pubblicazione dei dati relativi al 2014, provenienti dalla nuova indagine CAPI sulla spesa delle famiglie, viene diffusa la ricostruzione in serie storica dei principali aggregati di spesa e della povertà (relativa e assoluta). L'insieme dei dati ricostruiti è un sistema informativo articolato che copre un arco temporale di 17 anni, dal 1997 al 2013.

Prima di esaminare gli elementi fondanti della ricostruzione e considerare nel dettaglio il lavoro svolto, è necessario soffermarsi sul termine “ricostruzione”, argomento centrale di questo capitolo.

A livello teorico, si parla di ricostruzione in senso stretto (EUROSTAT, 2007) solo quando il fenomeno viene misurato con lo stesso strumento (la stessa indagine), ma con codifiche o con scale di misura modificate: classificazioni di riferimento, dettaglio delle voci elementari, ecc.

In realtà, l'esigenza di una ricostruzione sorge anche quando il fenomeno viene osservato con strumenti statistici simili ma non identici, cioè quando le innovazioni introdotte riguardano una molteplicità di fattori (come il disegno di campionamento, il questionario, il riferimento temporale di osservazione, la tecnica e così via). È questo il caso dell'indagine sulle spese che, come dettagliatamente illustrato nei capitoli precedenti, è stata completamente ridisegnata.

In presenza di un'indagine nuova bisogna seguire particolari procedure per ricollegare i risultati con quelli ottenuti attraverso l'indagine precedente. La ricostruzione ideale dovrebbe basarsi non solo sull'effettuazione in parallelo delle due indagini (quella vecchia e quella nuova) per un periodo lungo², ma dovrebbe prevedere anche la sovrapposizione di una parte del campione.

È noto e ben documentato che questo tipo di approccio è quasi impraticabile (Zhang e Von Sanden, 2008) essendo “ad alta intensità di risorse” e anche gli istituti nazionali di statistica, incaricati di effettuare rilevazioni continuative, non hanno mai (o quasi mai) risorse sufficienti per effettuare *parallel run* per periodi temporali così lunghi.

A livello internazionale, la letteratura sulla ricostruzione (*back casting*) dei dati d'indagine si sofferma, in prevalenza, sulle soluzioni adottate nel caso di rilevazioni continuative “specializzate” come, ad esempio, quelle sull'occupazione (le forze di lavoro) o sul fatturato degli esercizi commerciali (le vendite della piccola e grande distribuzione); le esperienze di ricostruzione proposte e i coefficienti di raccordo presentati (*conversion coefficients*) si fondano sulla sovrapposizione di indagini per periodi di tempo relativamente brevi.

L'indagine sulla spesa delle famiglie non è del tutto assimilabile alle indagini “specializzate”, essendo per sua natura generalista (fa riferimento a tutte le tipologie di spesa) e presentando un alto livello di complessità (il campione deve essere ampio per garantire significatività a livello territoriale, per le differenti tipologie di comune e per le diverse caratteristiche socio-demografiche delle famiglie).

¹ Il capitolo è stato redatto da F. Di Leo ad eccezione del paragrafo 3.2.2 redatto da A. Cutillo.

² Nel caso di un'indagine trimestrale si parla di 12 trimestri consecutivi.

Data la portata dell'operazione di ricostruzione, si è cercato di attingere all'esperienza di altri paesi; tuttavia nel panorama europeo non esiste un modello d'indagine *standard*, così come non è possibile far riferimento a *best practices* condivise.

L'indagine sulle spese non è infatti condotta con le stesse modalità e tempistica in tutti i paesi europei (EUROSTAT, 2003).

In Francia (INSEE, 2011), ad esempio, la periodicità è quinquennale; in Germania (Statistisches Bundesamt, 2013) all'ampia indagine quinquennale su reddito, consumo e condizioni di vita (EVS) se ne affianca una continuativa (LWR³) basata su un campione ridotto di circa 8.000 famiglie (circa il 12% di quello completo). In Spagna esiste un'indagine continuativa (annuale) dal 1997 e tale esperienza è, probabilmente, la più simile a quella italiana.

Nel 2006, l'indagine spagnola è stata ristrutturata per tenere conto delle rinnovate esigenze degli utilizzatori finali e delle nuove classificazioni e definizioni internazionali (INE, 2007). Con il passaggio al nuovo strumento d'indagine si è verificata una interruzione della serie storica⁴ e i dati del periodo 2006-2013 non sono direttamente confrontabili con quelli del periodo 1997-2005.

Anche l'esperienza italiana, come già detto, si contraddistingue per un'interruzione tra il 1996 e il 1997, quando è stata introdotta una nuova indagine PAPI che andava a sostituire la precedente sui Bilanci di Famiglia – BF⁵.

Con il passaggio da PAPI a CAPI si è voluto evitare una nuova interruzione ed è stata condotta una doppia rilevazione per un periodo di tempo sufficientemente lungo, mettendo in campo risorse tecniche ed economiche straordinarie.

Le caratteristiche delle due indagini hanno richiesto una riflessione accurata prima di effettuare qualsiasi azione di raccordo; l'analisi dei dati di base è stata infatti propedeutica all'individuazione degli aggregati da ricostruire.

L'assenza di un approccio *standard* cui fare riferimento ha comportato lo sviluppo di una metodologia che, pur partendo da esperienze già condotte, di consolidata tradizione⁶ e già apprezzate a livello internazionale⁷, ha assunto una propria originalità tanto da poter rappresentare un punto di riferimento significativo anche a livello internazionale.

Nell'effettuare in parallelo le due indagini, l'Istat ha raddoppiato lo sforzo organizzativo: la vecchia PAPI e la nuova CAPI sono state somministrate a due diversi campioni di famiglie per sei trimestri consecutivi, uno in più rispetto a quanto fatto per l'indagine sulle Forze di Lavoro (Istat, 2006), coinvolgendo ben 56.480 famiglie (32.003 per PAPI e 24.477 per CAPI). Un simile *parallel run* ha permesso un'analisi accurata dei dati e il raggiungimento di un livello di raccordo decisamente ampio.

3 Continuous Household Budget Surveys (Laufende Wirtschaftsrechnungen, LWR). Il modulo era presente anche prima del 2005 ma è stato rinnovato in maniera ampia.

4 Si veda quanto riportato nel sito dell'INE (accesso del 12 novembre 2014) <http://www.ine.es/jaxi/menu.do?L=1&type=pcaxis&path=%2Ft25/p458&file=inebase>

5 Cfr. Istat (2000), I consumi delle famiglie 1997, Annuari, n.4, pag. 9.

6 Cfr. Istat (2006), La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione, Metodi e Norme n. 32. Si veda in particolare il capitolo 8: La ricostruzione delle serie storiche dei principali indicatori del mercato del lavoro.

7 In un documento del 2008, ad esempio, Zhang e Von Sanden parlano dell'approccio euristico utilizzato per le forze di lavoro in Italia (Gatto, 2006) come un riferimento significativo e una *good practice* in presenza di *parallel run* per periodi relativamente brevi (5 trimestri nel caso considerato).

Le analisi e le elaborazioni dei dati in parallelo hanno permesso, innanzitutto, il massimo accostamento micro delle due indagini; a livello macro, si è quindi andati oltre il semplice “utilizzo meccanico” di coefficienti di raccordo. Le due indagini sono state infine raccordate tenendo conto di modelli teorici di riferimento: la distribuzione della spesa per consumi, la metodologia di calcolo della povertà relativa e le diverse componenti delle serie storiche trimestrali dell’indagine PAPI⁸.

3.1 Il processo della ricostruzione

Il processo di ricostruzione ha permesso di riconciliare i risultati ottenuti con la nuova metodologia di rilevazione (CAPI) con quelli ottenuti con la precedente (PAPI) e produrre delle serie storiche per il periodo 1997-2013 direttamente confrontabili con i dati diffusi a partire dal 2014.

Tre ordini di motivi rendono questo raccordo particolarmente importante anche per gli utenti:

- ogni anno vengono pubblicate, a diversi livelli di dettaglio, le stime provenienti dall’indagine campionaria sulla spesa delle famiglie, incluse quelle di povertà relativa e assoluta. Senza ricostruzione in serie storica, sarebbe impossibile condurre confronti diretti con gli anni precedenti al 2014;
- una corretta analisi di lungo periodo dei comportamenti di consumo e della povertà presuppone la disponibilità di serie storiche coerenti. Tale coerenza deve essere estesa al dettaglio territoriale, e, nel caso della spesa, alla tipologia familiare (il numero dei componenti, ad esempio) e alla natura della spesa (contenuto merceologico);
- una molteplicità di utenti, sia interni che esterni, utilizzano spesso i dati dell’indagine, come *input* per ulteriori elaborazioni e analisi. Per un utilizzo ottimale dei nuovi dati è necessario disporre di serie storiche che permettano di valutare in maniera prospettica l’evoluzione della condizione sociale ed economica delle famiglie.

Il periodo di sovrapposizione (i trimestri dal III 2012 al IV 2013) ha permesso di effettuare elaborazioni particolarmente complesse in due fasi successive (una micro e una macro), propedeutiche al vero e proprio lavoro di raccordo così come proposto dalla letteratura sul *back-casting* (Eurostat, 2007).

In generale, il lavoro micro è da considerarsi sempre necessario per una buona ricostruzione e, in questo caso, ha comportato una molteplicità di interventi che hanno interessato diverse fasi della lavorazione. In particolare le due indagini sono state “accostate” il più possibile alla luce delle pluralità di innovazioni apportate. Una volta completata la fase micro, sono stati calcolati i coefficienti macro di raccordo (i cosiddetti *conversion coefficients*) e i dati rettificati con tali coefficienti sono stati messi a confronto con diversi modelli teorici. In particolare l’andamento delle

8 Quando saranno disponibili i dati trimestrali CAPI per un periodo sufficientemente lungo (almeno 4 anni) si potrà valutare anche come la nuova indagine descriva l’andamento stagionale della spesa e le differenze rispetto ai dati PAPI. Ciò potrà rappresentare un ulteriore *step* di ricostruzione, basato sulla valutazione della stagionalità misurata con l’indagine CAPI e una revisione di quella misurata con la PAPI. A margine di tale lavoro si potrà valutare se e in che misura la crisi economica ha modificato il comportamento stagionale di consumo, così come appare dai dati attualmente disponibili.

serie storiche trimestrali e la distribuzione della spesa equivalente, che si utilizza per la misura della povertà relativa, sono state messe a confronto con la distribuzione log-normale⁹.

La ricostruzione della serie storica “per componenti” (a partire dai dati trimestrali) non è stata giudicata, al momento, la tecnica ottimale dato il periodo di sovrapposizione relativamente breve (sei trimestri) e alla luce dei modificati periodi di riferimento e del metodo di riporto al mese delle spese.

3.2 L'approccio micro al trattamento dei dati

L'approccio micro ha avuto una particolare rilevanza nel lavoro di ricostruzione. Oltre ad essere una modalità d'intervento prevista in letteratura, tale approccio salvaguarda al meglio la coerenza strutturale dei dati e permette di cogliere l'evoluzione delle variabili osservate anche in presenza di un *break* introdotto da un'indagine completamente nuova¹⁰. L'approccio micro ha dovuto tener conto preliminarmente di due elementi di differenziazione importanti:

1. la corrispondenza tra le voci di spesa elementari delle due indagini;
2. le modalità di costruzione dei coefficienti di riporto all'universo.

La corrispondenza tra le voci di spesa ha richiesto una serie di valutazioni, in quanto il passaggio alla nuova classificazione Coicop¹¹ è stato fatto con l'introduzione di vari cambiamenti (così come illustrato in precedenza¹²).

Tali cambiamenti non hanno infatti comportato una semplice ricodifica delle voci, cosa che, nella ricostruzione, sarebbe stata facilmente recepita introducendo una tavola di corrispondenza (*concordance table*), così come previsto dal manuale sul *backcasting* (Eurostat, 2007); la nuova indagine ha introdotto modifiche sostanziali nel questionario, che hanno riguardato sia il contenuto merceologico dei prodotti sia l'approccio alla rilevazione della spesa.

3.2.1 Le voci elementari di spesa nelle due indagini

Come già ampiamente illustrato nel par. 2.6.1, la massima corrispondenza merceologica tra le voci di spesa utilizzate nelle due indagini (CAPI \Leftrightarrow PAPI) si ottiene attraverso 265 aggregati derivati da opportune combinazioni delle voci elementari. D'altra parte solo per 107 voci si è in presenza di una corrispondenza biunivoca (1 prodotto PAPI = 1 prodotto CAPI), molto spesso la corrispondenza si ottiene pas-

9 La teorica distribuzione della spesa totale “normalizzata” con una scala di equivalenza può essere rappresentata con i parametri della distribuzione log-normale (Kalecki, 1945); di conseguenza le due indagini, a parità di definizioni e di riporto all'universo, devono tendere (per il periodo di sovrapposizione) a tale distribuzione. La ricostruzione si basa sull'individuazione dei parametri teorici della distribuzione - $\log \chi (\mu, \sigma^2)$ - e sulla traslazione delle due distribuzioni osservate verso la teorica, attraverso appropriati coefficienti di raccordo.

10 Cfr. Eurostat (2007), §2.1 - Methods based on detailed re-working of individual data (micro-approaches).

11 Cfr. Eurostat (2013), Coicop Five-Digit - Structure and Explanatory Notes, documento della Unit B5, del 6 dicembre 2013.

12 Cfr. par. 2.6.1.1 - Il raccordo tra le voci di spesa

sando da un singolo prodotto (PAPI) ad un dettaglio più ampio¹³ (CAPI); altre volte si ottiene attraverso il minimo comune multiplo tra una pluralità di voci CAPI e una pluralità di voci PAPI. Solo per 16 voci, l'indagine CAPI prevede un dettaglio minore¹⁴ rispetto a quello utilizzato nella PAPI.

Oltre al dettaglio delle voci, l'"asimmetria" tra le indagini si realizza a seguito della (a volte) diversa definizione puntuale degli *item*, della diversa collocazione del quesito all'interno del questionario¹⁵ e del diverso percorso logico seguito per arrivare a identificare e valorizzare la spesa¹⁶.

Tali asimmetrie possono modificare le reazioni degli intervistati e rendere la spesa, misurata con le due indagini, leggermente diversa. Ciò risulta più evidente in presenza di spese rare o non frequenti; in generale, infatti, all'aumentare della variabilità della spesa (e al diminuire della frequenza) cresce l'errore delle stime e, di conseguenza, cresce la distorsione associata al calcolo dei coefficienti di raccordo. Il raccordo è tanto più robusto quanto più ridotta è la variabilità della stima¹⁷.

Infine, il raccordo deve garantire la rappresentatività dei diversi comportamenti di spesa che caratterizzano i gruppi socio-economici di popolazione presenti sul territorio nazionale; si è quindi cercato un *trade-off* tra livello di analiticità del raccordo, per quanto riguarda le voci di spesa, e rappresentatività delle differenze socio-economiche nei *pattern* di consumo.

Nel prospetto 3.1 si riporta la classificazione adottata per la ricostruzione e il suo posizionamento all'interno della Coicop-HBS.

13 Si tratta di una circostanza già illustrata nel paragrafo 2.6.1 dedicato al raccordo tra le voci di spesa. Se si fosse verificato solo questo "spacchettamento" ci troveremmo davanti ad una situazione *standard* che avrebbe garantito una ricostruzione ad un livello elevato di dettaglio.

14 A titolo di esempio si può far riferimento al caso dell'acquisto di mobili. Con la nuova indagine è prevista una voce unica mentre in precedenza ne erano rilevate distintamente tre: le spese per mobili da interno, mobili da esterno e per il bagno.

15 In teoria, nel caso dell'acquisto di calzature, il raccordo tra le due indagini è diretto (corrispondenza biunivoca) ma in pratica c'è almeno una differenza definitoria significativa: il trattamento delle calzature per uso sportivo. Nel vecchio questionario era chiesta la spesa dell'ultimo mese per "calzature da uomo (comprese calzature per lo sport, escluse quelle ortopediche)". Nel nuovo questionario (CAPI) è precisato che devono essere scorporati quei prodotti destinati esclusivamente alla pratica sportiva (scarponi da sci, quantoni da boxe, cap da equitazione, eccetera).

16 Considerando le spese per l'affitto dell'abitazione principale (per le abitazioni secondarie vale lo stesso ragionamento) si verifica, ancora una volta, una situazione di corrispondenza biunivoca. L'esame dei questionari ci dice, però, che mentre nell'indagine PAPI c'era una sezione interamente dedicata all'abitazione, che comprendeva sia l'affitto sia altre spese connesse (utenze, spese condominiali, manutenzioni, ecc.) nell'indagine CAPI ci sono due separate sezioni. Le informazioni sull'affitto vengono raccolte nell'intervista iniziale, tutte le altre in quella finale.

17 Si veda al riguardo quanto detto in precedenza sulla significatività delle stime dell'indagine CAPI secondo un dettaglio a quattro e cinque cifre Coicop.

Prospetto 3.1 – Classificazione adottata per la ricostruzione COICOP (REV 2013)

COICOP (Rev 2013)	
	01. Prodotti alimentari e bevande analcoliche
	01.1 Prodotti alimentari
1	01.1.1 Pane e cereali
2	01.1.2 Carne
3	01.1.3 Pesce e frutti di mare
4	01.1.4 Latte, formaggio e uova
5	01.1.5 Oli e grassi
6	01.1.6 Frutta
7	01.1.7 Ortaggi
8	01.1.8 Zucchero, marmellata, miele, cioccolata e dolci
9	01.1.9 Prodotti alimentari n.c.a.
	01.2 Bevande analcoliche
10	01.2.1 Caffè, tè e cacao
11	01.2.2 Acque minerali, bibite analcoliche, succhi di frutta e di verdure
	02. Bevande alcoliche, tabacco e droghe
12	02.1 Bevande alcoliche
13	02.2 Tabacco
	03. Articoli di abbigliamento e calzature
	03.1 Articoli di abbigliamento
14	03.1.1,3 e 4 Articoli di abbigliamento (esclusi indumenti 03.1.2) e relativi servizi
15	03.1.2 Indumenti
16	03.2 Calzature
	04. Abitazione, acqua, elettricità, gas e altri combustibili
17	04.1 Canoni reali di affitto dell'abitazione
18	04.2 Canoni figurativi di affitto dell'abitazione
19	04.3 Manutenzione e riparazioni dell'abitazione
20	04.4 Distribuzione d'acqua e servizi vari connessi con l'abitazione
21	04.5 Elettricità, gas e altri combustibili
	05. Arredamenti, apparecchi per uso domestico e manutenzione corrente dell'abitazione
22	05.1-6 Arredamenti, apparecchi per uso domestico e manutenzione corrente dell'abitazione (escl. 05.6.1)
23	05.6.1 Beni per la casa non durevoli
	06. Sanità
24	06.1-3 Sanità (esclusi prodotti farmaceutici 06.1.1)
25	06.1.1 Prodotti farmaceutici
	07. Trasporti
26	07.1-3 Trasporti (esclusi carburanti e lubrificanti 07.2.2)
27	07.2.2 Carburanti e lubrificanti per mezzi personali di trasporto
28	08. Comunicazioni
	09. Ricreazione e cultura
29	09.1, 2 e 3 Beni e articoli per ricreazione e cultura
30	09.4 e 6 Servizi ricreativi e culturali e viaggi "tutto compreso"
31	09.5 Giornali, libri e articoli di cartoleria
32	10. Istruzione
33	11. Ristoranti e alberghi
	12. Beni e servizi vari
34	12.1 e 3 Beni vari (esclusi parrucchieri e istituti di bellezza 12.1.3)
35	12.2, 4, 6 e 7 Servizi vari (inclusi parrucchieri e istituti di bellezza 12.1.3)
36	12.5 Assicurazione
Totale consumi (somma 01. - 12.)	
37	04.3r Manutenzione e riparazioni straordinaria dell'abitazione

Le voci di spesa (*item*) rilevate con l'indagine PAPI e prese in considerazione per la ricostruzione sono state 279, di queste quasi tutte¹⁸ (277) hanno contribuito alla ricostruzione finale. Per la nuova indagine CAPI, sono stati utilizzati 485 *item*. La ricostruzione ha tenuto conto di tutte le voci inserite nella Coicop-HBS (2013),

18 In questo lavoro non è stata presa in considerazione la spesa per il mutuo per l'abitazione primaria e secondaria, i rimborsi per l'assicurazione sulla casa, il denaro avuto in prestito da parenti e/o amici, le spese per la restituzione di prestiti.

3. La ricostruzione delle stime in serie storica

125

comprese le manutenzioni straordinarie dell'abitazione già presenti nella rilevazione PAPI (Tavola 3.1).

Tavola 3.1 – Numero di voci per macro categoria

	COICOP (Rev 2013)	Numero di voci elementari (item)	
		PAPI	CAPI
1	01.1.1 Pane e cereali	7	11
2	01.1.2 Carne	10	7
3	01.1.3 Pesce e frutti di mare	4	6
4	01.1.4 Latte, formaggio e uova	6	9
5	01.1.5 Oli e grassi	5	5
6	01.1.6 Frutta	9	11
7	01.1.7 Ortaggi	8	19
8	01.1.8 Zucchero, marmellata, miele, cioccolata e dolci	3	6
9	01.1.9 Prodotti alimentari n.c.a.	2	8
10	01.2.1 Caffè, tè e cacao	2	3
11	01.2.2 Acque minerali, bibite analcoliche, succhi di frutta e di verdure	3	4
12	02.1 Bevande alcoliche	3	7
13	02.2 Tabacco	1	3
14	03.1.1,3 e 4 Articoli di abbigliamento (esclusi indumenti 03.1.2) e relativi servizi	6	7
15	03.1.2 Indumenti	5	6
16	03.2 Calzature	4	4
17	04.1 Canoni reali di affitto dell'abitazione	3	5
18	04.2 Canoni figurativi di affitto dell'abitazione	3	6
19	04.3 Manutenzione e riparazioni dell'abitazione	10	21
20	04.4 Distribuzione d'acqua e servizi vari connessi con l'abitazione	6	21
21	04.5 Elettricità, gas e altri combustibili	10	21
22	05.1-6 Arredamenti, apparecchi per uso domestico e manutenzione corrente dell'abitazione (escl. 05.6.1)	31	50
23	05.6.1 Beni per la casa non durevoli	6	4
24	06.1-3 Sanità (esclusi prodotti farmaceutici 06.1.1)	14	14
25	06.1.1 Prodotti farmaceutici	1	1
26	07.1-3 Trasporti (esclusi carburanti e lubrificanti 07.2.2)	20	37
27	07.2.2 Carburanti e lubrificanti per mezzi personali di trasporto	3	4
28	08. Comunicazioni	9	20
29	09.1-3 Beni e articoli per ricreazione e cultura	21	36
30	09.4 e 6 Servizi ricreativi e culturali e viaggi "tutto compreso"	12	31
31	09.5 Giornali, libri e articoli di cartoleria	5	15
32	10. Istruzione	3	20
33	11. Ristoranti e alberghi	9	19
34	12.1 e 3 Beni vari (esclusi parrucchieri e istituti di bellezza 12.1.3)	7	12
35	12.2, 4, 6-7 Servizi vari (inclusi parrucchieri e istituti di bellezza 12.1.3)	9	21
36	12.5 Assicurazione	5	8
37	04.3r Manutenzione e riparazioni straordinaria dell'abitazione	12	3
Totale voci che partecipano alla ricostruzione		277	485

Le singole voci delle due indagini sono state inserite all'interno di un quadro di riferimento coerente con la Coicop (scendendo in alcuni casi ad un dettaglio più am-



pio delle prime due cifre¹⁹). La ricodifica²⁰ ha permesso, inoltre, di implementare un sistema di ricostruzione (un iper-cubo statistico) tenendo conto sia della natura della spesa che della tipologia familiare²¹.

3.2.2 Il sistema di ponderazione e di calibrazione

Il raccordo micro non si è limitato alla verifica della coerenza delle classificazioni con la ricodifica delle voci elementari; un altro elemento particolarmente importante e innovativo è stata la realizzazione di un sistema di ponderazione e di calibrazione PAPI coerente con quello CAPI.

Il sistema di ponderazione e di calibrazione CAPI è, come già illustrato, significativamente diverso da quello adottato per la vecchia indagine. Nell'indagine CAPI, infatti, oltre ad essere cambiato il disegno di campionamento, sono stati modificati i vincoli di calibrazione²² e le popolazioni di riferimento, a seguito delle risultanze del Censimento del 2011 che, come accade usualmente, ha comportato una revisione delle stime anagrafiche²³. I coefficienti di riporto della indagine sui Consumi PAPI, utilizzati in tutte le pubblicazioni precedenti al 2014, fanno riferimento a una popolazione *ad hoc* che, a partire dal Censimento 2001, è stata utilizzata per tutte le indagini campionarie sulle famiglie con serie storica iniziata prima del 2011.

Nell'ottica della ricostruzione, è stata testata la possibilità e il vantaggio di utilizzare, anche per i dati PAPI, il procedimento di calcolo dei coefficienti di riporto all'universo adottato per la CAPI; la ricostruzione effettuata sui dati PAPI ingloba quindi anche l'effetto dovuto alle nuove stime di popolazione. La metodologia di calibrazione utilizzata per ottenere i coefficienti²⁴ è quella di Deville e Särndal (1992).

Per i sei trimestri di sovrapposizione, dal III del 2012 al IV del 2013, l'indagine PAPI è stata quindi ricalibrata sui seguenti totali, ottenuti dalla nuova indagine:

1. Popolazione residente per ripartizione, sesso e classi di età (0-14, 15-29, 30-59, 60-74, 75 e più)
2. Popolazione residente per regione
3. Famiglie residenti per regione
4. Popolazione straniera residente per ripartizione
5. Popolazione di 15 anni e più per ripartizione, condizione e posizione nella professione auto-dichiarate (Dirigenti, quadri e impiegati; operai e assimilati; imprenditori e liberi professionisti; lavoratori in proprio e altri autonomi; disoccupati; inattivi)²⁵.

19 Si possono identificare, ad esempio, le spese ad un livello di dettaglio di tre-quattro cifre Coicop.

20 La ricodifica è stata fatta al massimo livello possibile, ovvero prevedendo un particolare dettaglio per alcune voci. Per le spese alimentari, ad esempio, sono state considerate quattro cifre della classificazione. Nel caso delle spese di trasporto è stato possibile distinguere l'acquisto di combustibili e lubrificanti da tutte le altre spese.

21 La tipologia familiare è stata identificata considerando il dettaglio territoriale e molte caratteristiche socio-economiche della famiglia.

22 Si veda il par. 2.5.2 per le differenze nei vincoli della vecchia e della nuova indagine.

23 Si veda Istat (2012), 15° Censimento generale della popolazione e delle abitazioni, Roma, 19 Dicembre 2012.

24 Si veda al riguardo quanto illustrato nel par. 2.5.2.

25 Si noti che per l'indagine sulle spese, il vincolo sulla condizione e la posizione nella professione sono relative alla ricostruzione oggettiva, secondo le modalità dell'ILO-Eurostat. In questo caso, invece, si utilizza come vincolo noto la condizione auto-dichiarata così come stimato dall'indagine sulle Spese. Nella vecchia

Per i periodi precedenti (dal 1997 al II trimestre 2012), la ri-calibrazione è stata effettuata sui totali noti desumibili dalle serie ufficiali della popolazione, prodotte dal servizio demografico dell'Istituto (la stessa fonte utilizzata per vincolare l'indagine sulle spese); i totali per la condizione e posizione nella professione sono stati ottenuti sulla base di matrici di raccordo trimestrali tra condizione ILO e condizione auto-dichiarata²⁶ stimate per i trimestri di sovrapposizione. A titolo di esempio, nella Tavola 3.2 si riportano i dati riferiti al 3° trimestre 2012 per la ripartizione Nord-Ovest: per gran parte degli intervistati la condizione auto-dichiarata coincide con quella prevista dalla classificazione ILO (le quote sono sempre superiori al 96%). L'unica eccezione è rappresentata dai disoccupati tra i quali solo il 63% risulta disoccupato in entrambe le classificazioni: l'8,2% dei disoccupati ILO si dichiara occupato dipendente, un ulteriore 3,3% lavoratore autonomo e il restante 2,3% inattivo.

Tavola 3.2 – Matrice di confronto tra condizione ILO e auto-dichiarata - III trimestre 2012 (Ripartizione Nord-Ovest)

Condizione ILO	Condizione autodichiarata							Totale (=100%)
	Dirigenti e impiegati	Operai e assimilati	Imprenditori e liberi professionisti	Indipendenti: altri	Disoccupati	Ritirato dal lavoro	Altro inattivo	
Dirigenti e impiegati	99,42	0,01	0,02	0,02	0,10	0,00	0,44	2.751.652
Operai e assimilati	0,04	97,87	0,10	0,10	0,44	0,00	1,45	2.295.391
Imprenditori e liberi professionisti	-	-	100,00	-	-	0,00	-	418.570
Indipendenti: altri	0,06	0,05	0,15	96,79	0,65	0,00	2,31	1.155.038
Disoccupati	4,23	4,01	-	3,26	63,09	0,00	25,42	760.655
Ritirato dal lavoro	0,00	0,00	0,01	0,01	0,04	99,70	0,23	3.602.791
Altro inattivo	0,40	0,31	0,86	0,85	2,60	0,00	94,97	2.692.892
Totale	2.780.428	2.286.206	446.802	1.168.753	571.741	3.592.076	2.830.982	

Anche nei casi peggiori, che in generale si verificano per le regioni del Sud (Tavola 3.3), la coincidenza non scende mai al di sotto del 95%, ma, ancora una volta, con l'unica eccezione dei disoccupati: solo il 43% è disoccupato per entrambe le classificazioni, il 7,1% dei disoccupati ILO si dichiara occupato dipendente, il 6,8% autonomo e ben il 43% inattivo.

26 indagine, infatti, la professione era rilevata solamente seconda la modalità auto-dichiarata, mentre nella nuova viene rilevata in entrambe le modalità, così come nell'indagine sulle Forze di Lavoro. Ovviamente, le distribuzioni secondo le due modalità sono correlate tra loro, e quindi si è preferito utilizzare questo vincolo. Ci si riferisce alla convenzione ILO-EUROSTAT (Fortunato e Verzicco, 2004) comunemente utilizzata nelle indagini campionarie sulle Forze di Lavoro, ovvero "persone che abbiano lavorato almeno un'ora nella settimana di riferimento". La condizione auto-dichiarata viene valutata sulla base delle indicazioni fornite dall'intervistato sulla propria condizione professionale e su quella degli altri componenti della famiglia di 16 anni o più.

Tavola 3.3 – Matrice di confronto tra condizione ILO e auto-dichiarata - III trimestre 2012 (Ripartizione Sud)

Condizione ILO	Condizione autodichiarata							Totale (=100%)
	Dirigenti e impiegati	Operai e assimilati	Imprenditori e liberi professionisti	Indipendenti: altri	Disoccupati	Ritirato dal lavoro	Altro inattivo	
Dirigenti e impiegati	99,45	0,01	0,00	0,02	0,01	0,00	0,51	1.506.245
Operai e assimilati	0,02	99,06	0,00	0,03	0,02	0,00	0,87	1.339.163
Imprenditori e liberi professionisti	-	-	100,00	-	-	0,00	-	292.434
Indipendenti: altri	0,03	0,04	0,00	98,33	0,04	0,00	1,55	767.998
Disoccupati	2,01	5,08	0,12	6,71	43,10	0,00	42,98	1.654.520
Ritirato dal lavoro	0,00	0,00	0,00	0,00	0,00	99,88	0,11	2.383.227
Altro inattivo	0,58	0,75	0,05	0,90	2,83	0,00	94,90	4.044.186
Totale	1.554.933	1.441.393	296.591	903.564	828.192	2.380.468	4.582.633	

Il totale di cui al punto 4 (popolazione straniera) è stato utilizzato solamente per gli anni dal 2009 al 2012, poiché, per gli anni precedenti (1997-2008), l'indagine PAPI non rilevava l'informazione sulla cittadinanza dei componenti la famiglia.

3.3 L'impatto del raccordo micro

L'impatto del raccordo micro descritto nei paragrafi precedenti può essere misurato in maniera puntuale, sia per quanto riguarda il sistema di ponderazione e calibrazione, sia per quanto riguarda il sistema classificatorio²⁷.

La modifica dei coefficienti di riporto all'universo (il sistema di ponderazione e calibrazione) ha ovviamente avuto impatto sulla stima di tutti i 12 capitoli di spesa Coicop (Tavola 3.4) ed ha comportato una riduzione della stima della spesa media mensile pari a -2,6% (circa 60 Euro).

Le differenze più marcate si osservano per le spese legate al tempo libero e alla cultura (Coicop 09.-11.), mentre molto più contenuto è stato l'impatto sulle spese per prodotti alimentari e bevande e su quelle per l'abitazione²⁸.

27 La nuova classificazione, tra l'altro, non presuppone l'adozione di una diversa definizione dell'aggregato "spesa totale" con l'inclusione di nuove voci seppure previste dalla Coicop (ad esempio non viene fatta una stima degli acquisti di narcotici o stupefacenti che compaiono negli schemi allargati della Contabilità Nazionale secondo il SEC 2010).

28 Il nuovo sistema di ponderazione e di calibrazione dei pesi ha permesso di evidenziare meglio i comportamenti di spesa più "bassi" e, dunque, nei casi in cui non esistono differenze sostanziali nei comportamenti d'acquisto delle famiglie l'impatto del nuovo riporto all'universo è contenuto.

Tavola 3.4 – Spesa media mensile per capitolo di spesa e sistema di ponderazione utilizzato - Anno 2013
(valori in euro e differenza percentuale)

COICOP (Rev 2013)	Ponderazione		Diff. % (new/old)
	Old	New	
01. Prodotti alimentari e bevande analcoliche	440,3	437,0	-0,7
02. Bevande alcoliche, tabacco e droghe	40	40,3	-1,2
03. Articoli di abbigliamento e calzature	117	112,5	-4,2
04. Abitazione, acqua, elettricità, gas e altri combustibili	774	760,3	-1,8
05. Arredamenti, apparecchi per uso domestico e manutenzione dell'abitazione	95	92,2	-3,7
06. Sanità	87	85,6	-2,1
07. Trasporti	267	257,0	-3,9
08. Comunicazioni	44	44,0	-0,7
09. Ricreazione e cultura	119	113,7	-4,8
10. Istruzione	19	17,6	-7,4
11. Ristoranti e alberghi	115	107,2	-6,8
12. Beni e servizi vari	190,0	183,6	-3,4
Totale consumi	2311,2	2251,0	-2,6

L'impatto del sistema classificatorio, si può valutare considerando che, per la ricostruzione, come già detto, si è scelto di adottare un livello di dettaglio che, in molti casi, è più ampio rispetto alle prime due cifre della Coicop (Prospetto 3.1). Tuttavia le 37 voci identificate permettono un facile raccordo tra il vecchio sistema di classificazione basato su 19 capitoli di spesa e il nuovo, costituito dai 13 capitoli. Si può dunque far riferimento a una matrice (Tavola 3.5) di dimensioni 19 x 13.

Facendo riferimento alla spesa media mensile rilevata con la PAPI nel 2013, le spese dei 19 capitoli di spesa PAPI sono state riclassificate nei 12+1 capitoli di spesa CAPI (ottenuti dalla classificazione COICOP, con l'aggiunta della manutenzione straordinaria dell'abitazione che, pur non essendo inclusa per definizione nella spesa per consumi, è sempre stata inclusa nell'ammontare di spesa diffuso).

Le differenze riguardano:

- la spesa per alimentari e bevande (la PAPI includeva anche le spese per bevande alcoliche che nella CAPI sono conteggiate insieme a quelle per tabacchi);
- la spesa per combustibili (nella CAPI non viene più distinta da quella per abitazione);
- le spese per la ristorazione e gli alberghi, incluse mense (nella PAPI sono inserite nel capitolo "altri beni e servizi" e nel capitolo istruzione per le spese relative alle mense scolastiche e universitarie, nella CAPI vengono raggruppate nello specifico capitolo);
- le spese di assicurazione (nella PAPI sono attribuite al capitolo di riferimento -ad es. le spese per l'assicurazione dei mezzi di trasporto rientrano tra le spese per trasporti così come le spese per l'assicurazione sull'abitazione rientrano tra le spese per l'abitazione- e nella CAPI raggruppate nel capitolo "Beni e servizi vari").

Tavola 3.5 – Matrice di conversione della spesa: dalla classificazione “Cap19” (PAPI) alla Coicop a 2 cifre (medie mensili 2013)

Classificazione PAPI	COICOP a 2 cifre													
	01. Prodotti alimentari e bevande analcoliche	02. Bevande alcoliche, tabacco e droghe	03. Articoli di abbigliamento e calzature	04. Abitazione, acqua elettrica, gas e altri combustibili	05. Arredamenti, apparecchi per uso domestico e manut. abitaz.	06. Sanità	07. Trasporti	08. Comunicazioni	09. Ricreazione e cultura	10. Istruzione	11. Ristoranti e alberghi	12. Beni e servizi vari	TOTALE	04r. Manutenzione straordinaria dell'abitazione
Alimentari e bevande (01-08)	440,3	20,4											460,7	
09. Tabacchi		20,4											20,4	
10. Abbigliamento			109,1										109,1	
11. Abitazione				635,4								10,1	645,5	47,9
12. Combustibili				137,7									137,7	
13. Mobili			8,3		95,7							5,0	109,0	
14. Sanità						87,4							87,4	
15. Trasporti				1,5			266,2		2,5			65,8	336,0	
16. Comunicazioni								44,3					44,3	
17. Istruzione							0,6		4,9	19,0	2,5		27,0	
18. Tempo libero									94,1				94,1	
19. Beni e servizi vari							0,5		17,9		112,5	109,1	240,0	
Totale consumi	440,3	40,8	117,4	774,6	95,7	87,4	267,3	44,3	119,4	19,0	115,0	190,0	2311,2	47,9

3.3.1 Un ulteriore passaggio micro: il riporto al mese

Un ulteriore effetto sulle stime che può essere puntualmente valutato riguarda i diversi periodi di riferimento delle spese e, di conseguenza, il metodo di calcolo della spesa media mensile.

Tale effetto può essere valutato riportando i dati della nuova indagine allo stesso riferimento temporale della spesa PAPI, attraverso il metodo di stima R3 (cfr. PAR. 2.6.2); tale passaggio, può essere simulato, infatti, grazie alle informazioni presenti nel nuovo questionario. Richiamando quanto già descritto, si prenda ad esempio l'acquisto di frigoriferi e congelatori, che nell'indagine CAPI fa riferimento agli ultimi dodici mesi; agli intervistati è anche chiesto di indicare in quale mese la spesa è stata effettuata. Supponendo che le famiglie effettuino questo tipo di acquisto una sola volta nel corso dell'anno²⁹, le p famiglie campione possono essere suddivise in tre gruppi: quelle che non hanno effettuato l'acquisto nel corso dell'anno $[p_0]$, quelle che lo hanno effettuato nell'ultimo trimestre $[p_1]$ e quelle che lo hanno fatto in uno dei tre trimestri precedenti $[p_{1,2,3}]$. Secondo l'approccio R1, la spesa media mensile per il bene i $[Sm^i]$ si ottiene come rapporto tra l'ammontare della spesa $[S_j^i]$ sostenuta dalla famiglia j sia nell'ultimo trimestre $[S_{j,t_0}^i]$ che nei tre trimestri precedenti $[S_{j,t_{1,2,3}}^i]$ e il totale delle famiglie (per semplicità espositiva non si considera la ponderazione con i pesi di riporto).

²⁹ È possibile, in realtà, che le famiglie effettuino più di un acquisto nel periodo di osservazione CAPI (un anno) mentre è meno probabile che ciò si verifichi nel periodo PAPI (tre mesi). Tuttavia è sempre possibile tenere conto degli acquisti multipli all'interno del sistema di equazioni che è stato formalizzato. Chiaramente il verificarsi di tale circostanza introduce ulteriori condizioni all'interno del sistema e rende ancor più difficile il verificarsi dell'equivalenza tra le spese calcolate secondo approcci diversi.

$$Sm_{R1}^i = \frac{\sum_{j=1}^p S_{j,t_0}^i + \sum_{j=1}^p S_{j,t_{1,2,3}}^i}{p_0 + p_{1,2,3} + p_\phi} \quad [1]$$

Secondo l'approccio R2, la spesa media mensile $[Sm^i]$ si ottiene come rapporto tra l'ammontare di spesa sostenuto dalle famiglie nell'ultimo trimestre $[S_{j,t_0}^i]$ e il campione trimestrale di famiglie che, in questo caso, può essere suddiviso in due gruppi: quelle che hanno effettuato un acquisto nel corso dell'ultimo trimestre $[p_\phi]$ e quelle che non lo hanno fatto $[np_{\phi t_0}]$.

$$Sm_{R2}^i = \frac{\sum_{j=1}^p S_{j,t_0}^i}{p_0 + p_{\phi t_0}} \quad [2]$$

Pertanto sulla base della [1] e della [2] è possibile definire le due condizioni necessarie [3] affinché la spesa media mensile misurata nel primo caso (CAPI) sia coincidente con quella del secondo caso (PAPI): la spesa sostenuta dalle famiglie nei tre trimestri precedenti l'ultimo deve essere uguale al triplo della spesa sostenuta nell'ultimo trimestre e il numero di famiglie che non hanno effettuato la spesa nel corso dei 12 mesi deve essere uguale alla somma di chi non lo ha fatto nell'ultimo trimestre più tre volte il numero di famiglie che lo ha fatto. Formalmente:

$Sm_{R1}^i = Sm_{R2}^i$ se:

$$\begin{cases} \sum_{j=1}^p S_{j,t_{1,2,3}}^i = 3 \sum_{j=1}^p S_{j,t_0}^i \\ np_{\phi t_0} = p_\phi + 3p_\phi \end{cases} \quad [3]$$

Tornando all'esempio precedente, la spesa per l'acquisto dei frigoriferi ottenuta nei tre trimestri precedenti all'ultimo è inferiore a quella osservata nell'ultimo trimestre ($\sum_{j=1}^p S_{j,t_{1,2,3}}^i < 3 \sum_{j=1}^p S_{j,t_0}^i$) e le famiglie che non hanno effettuato acquisti nell'ultimo trimestre sono meno di quelle che non lo hanno fatto nei tre trimestri precedenti ($p_{\phi t_0} < p_\phi + 3p_\phi$).

Questa doppia circostanza fa supporre l'esistenza di un duplice effetto: *forward telescoping*, per quanto riguarda il verificarsi dell'evento di spesa³⁰, e *level expansion* per quanto riguarda il corretto ricordo della spesa sostenuta³¹. In altri termini, quando il periodo di riferimento è relativamente breve, si tende a farvi rientrare anche spese effettuate in periodi antecedenti e si tende a riportare valori di spesa mediamente maggiori. I due effetti combinati producono stime più elevate nel caso queste vengano calcolate con l'approccio R2 (nell'esempio dei frigoriferi la spesa calcolata con R2 è superiore del 10% a quella ottenuta con R1).

Sul totale della spesa, il passaggio da R1 a R2 comporta un innalzamento della stima compreso tra il 2 e il 3%; l'indagine PAPI presentava quindi una distorsione sia nei livelli sia nella stagionalità (decisamente rilevante per alcune tipologie di spesa), a causa dei più corti periodi di riferimento delle spese.

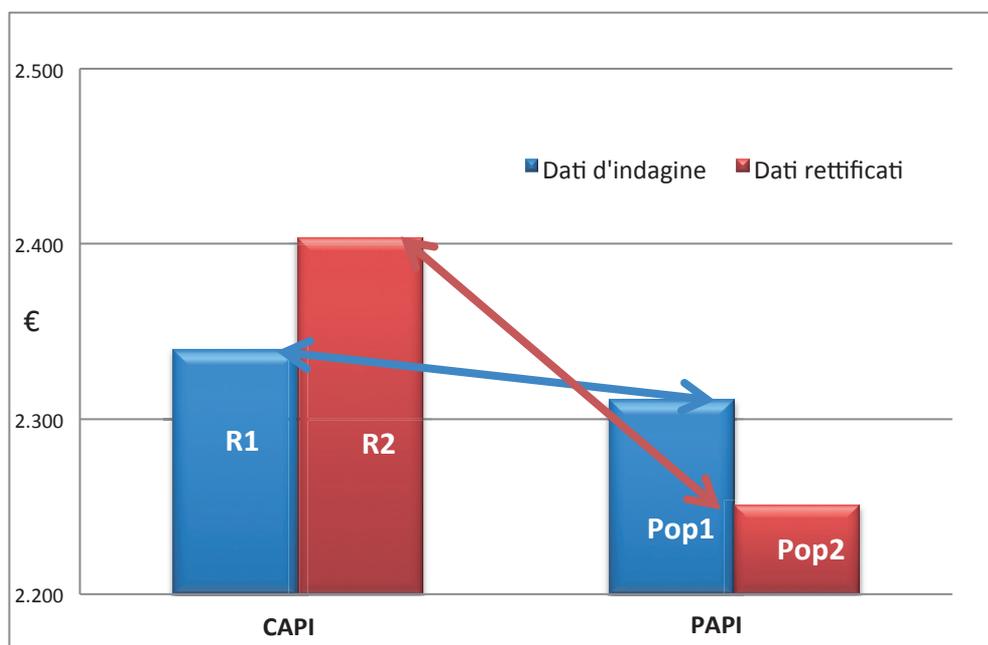
30 Si attribuisce all'ultimo trimestre una spesa effettuata in un momento precedente.

31 Più ci si allontana dal momento in cui è stato effettuato l'acquisto e più si perde l'esatta dimensione di quanto è stato speso effettivamente. In questo caso si ipotizza la percezione di una spesa più ridotta. Tuttavia, in generale, i dati disponibili confermano l'esistenza di un significativo e generalizzato *forward telescoping* mentre per accertare il *level expansion* sarebbero necessarie ulteriori verifiche che non sono necessarie per le finalità di questa ricostruzione.

Per avere un'idea visiva dell'impatto dovuto sia al metodo di riporto al mese, sia al sistema di ponderazione, si rimanda al grafico 3.1 costruito utilizzando i dati dell'indagine CAPI³², trattati sia con l'approccio R1 sia con l'approccio R2, e i dati della PAPI costruiti, inevitabilmente, solo con l'approccio R2, ma utilizzando i due diversi sistemi di ponderazione.

La ricostruzione in serie storica deve quindi riconciliare le due colonne blu del grafico, ma, inevitabilmente, è passata per il confronto tra le colonne rosse. Differenze analoghe a quelle messe in evidenza a livello nazionale sono state rilevate anche considerando il dettaglio territoriale NUTS³³ 2 (21 unità di secondo livello).

Grafico 3.1 – Spesa media mensile familiare – Anno 2013



Tali differenze sono state dunque tenute in considerazione nel raccordo macro che, come vedremo, è stato realizzato con un procedimento di riallineamento attuato in tre fasi.

32 I dati utilizzati per il grafico erano provvisori poiché l'ultimo calcolo dei valori CAPI R2 è stato fatto prima che fosse completato il trattamento dei dati di base della nuova indagine e la correzione dei valori anomali. In ogni caso un nuovo calcolo avrebbe prodotto risultati simili dato che le differenze non risentono del trattamento degli *outlier*.

33 La nomenclatura delle unità territoriali statistiche, in acronimo NUTS (dal francese *nomenclature des unités territoriales statistiques*) adottata dalla UE al secondo livello di dettaglio prevede l'identificazione, per l'Italia, delle 19 regioni più le 2 province autonome di Trento e Bolzano.

3.4 L'approccio macro di ricostruzione

Il calcolo dei coefficienti di raccordo macro fa riferimento, come già detto, a 37 macro aggregati e a 36 tipologie familiari.

L'identificazione degli aggregati di spesa e dei gruppi di famiglie è stata fatta cercando di:

- evitare l'associazione di prodotti con caratteristiche merceologiche diverse e/o incompatibili;
- evitare l'associazione di prodotti troppo diversi in termini di "funzione" (ad es. durevoli e non durevoli) e quindi troppo diversi per frequenza di spesa;
- distinguere i prodotti con una forte stagionalità da quelli con un andamento meno variabile in presenza di fattori stagionali;
- garantire una sufficiente numerosità campionaria per ciascun aggregato e tipologia di famiglia³⁴.

In generale, in base alla teoria economica sottostante l'andamento della spesa per consumi si è tenuto conto dell'ampiezza demografica della famiglia (e quindi delle economie di scala che si realizzano all'aumentare del numero di componenti), della ripartizione e dell'ampiezza del comune di residenza (e quindi dei differenti livelli e stili di consumo), dell'età della persona di riferimento della "scheda famiglia"³⁵ (come *proxy* del momento di ciclo di vita familiare) (Prospetto 3.2). La combinazione di tali caratteristiche ha portato all'individuazione di 36 gruppi.

Prospetto 3.2 – Variabili utilizzate per l'identificazione dei gruppi (*cluster*)

Ripartizione	Comuni	Ampiezza della famiglia	Età della p.r.
R1 Nord	C1 Aree metropolitane	A1 Un componente	E1 Fino a 64 a.
R2 Centro	C2 Altri comuni	A2 Due componenti	E2 65 a e più
R3 Sud e Isole		A3 Tre componenti e più	

Le consistenti differenze indotte dal disegno e dalla tecnica d'indagine hanno, dunque, richiesto un lavoro di raccordo basato su un ampio numero di voci di spesa, differenti tipologie familiari e differenti ambiti territoriali. I comportamenti di consumo e la frequenza con la quale si acquistano i prodotti (beni e servizi), infatti, è variabile e ha comportato il calcolo di ben 1332 coefficienti elementari.

La combinazione dei criteri indicati nel prospetto 3.2, ad esempio, individua il primo gruppo [denominato R1C1A1E1] nelle famiglie residenti al Nord in un'area

34 È stata fatta la verifica dell'ipotesi che i diversi gruppi di famiglie presentassero una sufficiente omogeneità *within* e una più elevata variabilità *between* per entrambe le indagini e che, simultaneamente, la distanza tra i valori medi della spesa a parità di gruppo di appartenenza non potesse essere costante. In altri termini il passaggio da PAPI a CAPI ha modificato i valori d'indagine in maniera differenziata a seconda della tipologia di spesa e della tipologia familiare

35 Quest'ultima variabile pur non essendo una *proxy* del reddito è in teoria (e in pratica), la più significativa per evidenziare dei diversi comportamenti legati alla condizione socio-economica della famiglia

metropolitana³⁶ (Tavola 3.6), formate da un solo componente con età inferiore a 65 anni (persona di riferimento della scheda).

Per ciascun gruppo, la numerosità campionaria delle due indagini (PAPI e CAPI) nei sei trimestri di sovrapposizione, pur essendo relativamente variabile, è comunque, sempre, consistente e anche la variabilità di ciascun aggregato, a livello di singolo gruppo, ha confermato che l'ipotesi adottata è coerente e statisticamente affidabile³⁷. L'unica eccezione è rappresentata dalla spesa per servizi di istruzione [Coicop 10.], per la quale il vincolo della numerosità campionaria non viene sempre soddisfatto; per tali spese la ricostruzione è stata fatta considerando solo il tipo di comune e quindi limitandosi ai sei gruppi ottenuti dalla combinazione delle tre macro-aree (nord - centro - sud) e delle due tipologie di comune (area metropolitana – altro).

Tavola 3.6 – Numerosità campionaria dei cluster (PAPI e CAPI) e universo delle famiglie (CAPI) - Terzo e quarto trimestre 2012 e anno 2013.

Cluster	Osservazioni		Famiglie
	PAPI	CAPI	
R1C1A1E1	939	683	981.246,78
R1C1A1E2	928	597	927.770,32
R1C1A2E1	842	737	932.308,98
R1C1A2E2	1026	750	852.879,62
R1C1A3E1	1720	1233	1.442.717,48
R1C1A3E2	295	187	157.262,49
R1C2A1E1	1331	928	1.136.756,44
R1C2A1E2	1151	798	1.007.482,91
R1C2A2E1	1245	1090	1.025.961,44
R1C2A2E2	1405	1049	966.615,40
R1C2A3E1	3044	2449	2.441.462,76
R1C2A3E2	574	362	282.374,38
R2C1A1E1	431	325	592.848,78
R2C1A1E2	345	278	510.751,48
R2C1A2E1	323	293	418.265,24
R2C1A2E2	409	350	444.249,14
R2C1A3E1	717	646	879.571,93
R2C1A3E2	155	144	136.309,93
R2C2A1E1	414	301	287.746,54
R2C2A1E2	489	316	305.232,12
R2C2A2E1	401	377	313.613,73
R2C2A2E2	555	418	296.487,82
R2C2A3E1	1203	861	795.817,93
R2C2A3E2	254	191	122.294,88
R3C1A1E1	551	470	416.712,39
R3C1A1E2	626	520	485.861,24
R3C1A2E1	503	418	380.479,39
R3C1A2E2	802	563	488.794,74
R3C1A3E1	1855	1366	1.429.486,44
R3C1A3E2	477	275	225.216,31
R3C2A1E1	933	685	581.013,60
R3C2A1E2	1114	842	748.864,39
R3C2A2E1	682	616	514.260,61
R3C2A2E2	1079	872	666.880,20
R3C2A3E1	2653	2131	1.916.924,17
R3C2A3E2	532	356	245.409,53
Totale	32003	24477	25.357.931,50

36 Non è necessario indicare quali siano i comuni auto-rappresentativi perché già illustrato in precedenza, quanto ricordare che si tratta di comuni ampi da un punto di vista demografico e, tra questi, naturalmente Roma, Milano, Napoli, ecc.

37 C'è una sola eccezione, della quale si parlerà diffusamente nel proseguo di questo capitolo, che riguarda le spese per l'istruzione.

Per avere coefficienti stabili che, in qualche modo, potessero includere (in modo implicito) la stagionalità e per utilizzare tutta l'informazione disponibile, il processo di ricostruzione macro è stato effettuato facendo riferimento a due diversi periodi:

- III 2012 – II 2013 [T1]
- I 2013 – IV 2013 [T2]

Dapprima sono stati calcolati i valori medi mensili CAPI e PAPI_POP2 (nuova popolazione) con riferimento alle 37 tipologie di spesa e ai 36 gruppi di famiglie nel periodo T1 e T2. Il confronto tra i due set di dati ha permesso di calcolare due serie di rapporti [Y1 e Y2] e, infine, i coefficienti finali [Y] ottenuti come media semplice dei due.

I coefficienti [Y] sono stati applicati ai singoli record individuali dell'indagine per il periodo 1997-2013 per avere un primo quadro d'insieme. Le revisioni introdotte dall'indagine CAPI non potevano essere recepite dall'indagine PAPI utilizzando i soli coefficienti Y. Le differenze nei livelli di spesa regionali non potevano essere recepiti all'interno dei *data-set* PAPI_POP2_R³⁸ e, pertanto, dopo questa prima fase si è proceduto a riallineare i dati ai totali regionali.

Questa seconda fase ha comportato il calcolo di 21 coefficienti di raccordo regionali [R] ottenuti come rapporto tra i valori medi totali della spesa mensile - CAPI e PAPI_POP2_R - L'applicazione dei coefficienti [R] al *data-set* PAPI_POP2_R ha permesso di produrre dei nuovi archivi denominati PAPI_POP2_R2. Infine (terza fase) si è proceduto a ricondurre i valori medi delle variabili ai livelli CAPI ricalcolando dei nuovi coefficienti Y³⁹

Al termine di queste elaborazioni macro si è potuto procedere alle operazioni di *fine tuning*, mettendo a confronto i dati ricostruiti con i modelli teorici del consumo e della povertà.

3.5 La ricostruzione e il confronto con i modelli teorici

Il lavoro di ricostruzione non poteva essere limitato a un riallineamento puramente "algebrico" dei dati; la duplice ricostruzione, micro e macro, doveva cioè trovare un'ulteriore validazione teorica e per questo i dati sono stati confrontati con i risultati ottenuti tramite modelli.

In particolare, si è cercato di trovare coerenza rispetto alla dinamica trimestrale⁴⁰ e alla distribuzione della spesa in termini equivalenti⁴¹.

Le componenti tradizionali delle serie storiche (ciclo-trend, stagionalità, accidentale) sono facilmente evidenziabili a partire dalle serie storiche PAPI (anche calcolate con i nuovi coefficienti di riporto all'universo) grazie alla disponibilità di stime che si riferiscono a ben 68 trimestri consecutivi. Tali stime sono un *input* importante per la Contabilità Nazionale trimestrale.

38 Quelli che recepiscono i coefficienti Y.

39 Al termine di questa terza fase è stato creato un *data-set* denominato PAPI_POP2_R3.

40 Il campione permette, in teoria, di ottenere delle stime trimestrali coerenti con l'universo di riferimento.

41 La distribuzione della spesa totale "normalizzata" con la scala di equivalenza (Carbonaro) si distribuisce come una log-normale.

La serie storica (trimestrale) y_1, y_2, \dots, y_n delle voci di spesa Y prese in considerazione dalla ricostruzione può essere rappresentata con la [4] con una componente deterministica $f(t)$ e una residuale stocastica ε_t :

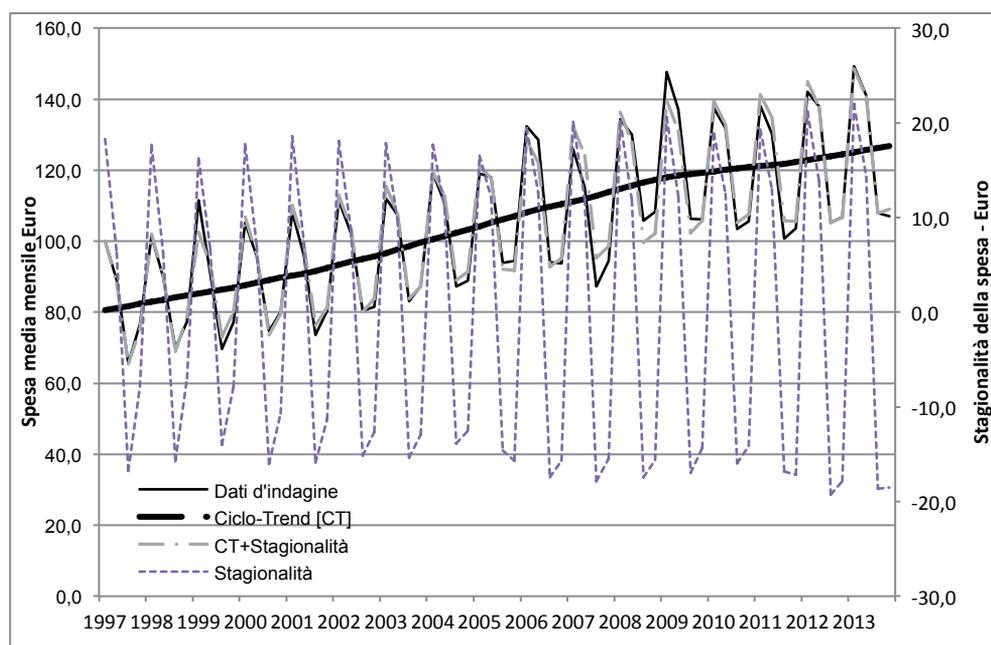
$$Y_t = f(t) + \varepsilon_t \quad [4]$$

Scomponendo la parte deterministica secondo il ciclo-trend [CT] e la stagionalità [S] possiamo giungere a una semplice notazione algebrica [5] e di conseguenza evidenziare la legge di evoluzione temporale del consumo in presenza di stagionalità.

$$f(t) = CT_t + S_t \quad [5]$$

Se prendiamo come caso di studio le spese per elettricità, gas e altri combustibili (Coicop 04.5), si può descrivere, anche con l'ausilio di una rappresentazione grafica (Grafico 3.2), il modello di riferimento.

Grafico 3.2 – Componenti trimestrali e sola componente stagionale (scala destra) della spesa media mensile familiare per elettricità, gas e altri combustibili. Anni 1997-2013 (valori in euro)



I dati teorici (parte deterministica del modello stimata con gli strumenti statistici tradizionali⁴²) presentano, in generale, un ottimo accostamento con i valori osservati. Tuttavia, la componente erratica risulta apprezzabile soprattutto in alcuni trimestri tra il 2007 e il 2009.

Le ragioni di questi scostamenti potrebbero essere molteplici (*shock* esogeni della spesa, quali stagioni invernali particolarmente calde o fredde, presenza di *outlier*

42 Si è utilizzato il modello Tramo/Seats (Gómez e Maravall, 2005) incorporato nel pacchetto Demetra © 2.2 (novembre 2008) sviluppato per Eurostat da Dossé e Hoffmann. Tramo/Seats è utilizzato per destagionalizzare le serie storiche da Istat, Eurostat, Bankitalia e Bce.

nei dati d'indagine e altri fattori), e a prescindere dalle cause, rappresentano un elemento particolarmente rilevante e influente sulla stagionalità nel periodo esaminato. Poiché la disponibilità di sei trimestri non è sufficiente per una ricostruzione *model based*⁴³ si è deciso di effettuare semplicemente uno studio sulla loro distribuzione di probabilità, da utilizzare come utile *benchmark* per la calibrazione dei coefficienti di ricostruzione.

I dati di indagine vengono utilizzati anche per il calcolo della povertà relativa. Questo calcolo, com'è noto, si basa su una metodologia ben consolidata⁴⁴, che implica l'identificazione di una soglia convenzionale e l'utilizzo della spesa equivalente (SE). In termini algebrici, l'incidenza della povertà relativa è calcolabile confrontando tale soglia con la funzione di densità di probabilità della SE che si distribuisce come una log-normale [6].

$$\text{Ln}(SE) \sim \text{Normale}(\mu, \sigma^2) \Rightarrow SE \sim \text{Log-normale}(\mu, \sigma^2) \quad [6]$$

Pertanto l'incidenza della povertà è calcolabile facendo riferimento alla [7] una volta identificato l'opportuno livello della soglia⁴⁵.

$$P(SE \leq \text{soglia}) = \int_0^{\text{soglia}} \frac{1}{\sqrt{2\pi\alpha SE}} e^{-\frac{(\ln(SE) - \alpha)^2}{2\sigma^2}} \quad [7]$$

In termini grafici è possibile rappresentare la distribuzione teorica [6] con riferimento a ciascuno dei 36 cluster di famiglie; è evidente che, per le modalità con le quali sono stati identificati i gruppi e sono stati costruiti i coefficienti di raccordo macro (descritti nelle sezioni precedenti), gli interventi finali sui dati (suggeriti dall'approccio modellistico) hanno riguardato essenzialmente i 12 cluster che si riferiscono a famiglie con tre o più componenti.

Per le famiglie composte da uno o due componenti si è già realizzato il massimo accostamento attraverso l'approccio micro (nuovi coefficienti di riporto all'universo) e macro (allineamento delle medie per ciascun cluster). Nel caso delle famiglie più ampie, invece, i dati presentano elementi di variabilità su cui è possibile agire, visto che includono comportamenti di spesa riferiti a famiglie di diversa numerosità: tre, quattro, cinque, sei, sette componenti e più. Tale numerosità influenza il livello di spesa media e, inevitabilmente induce distorsioni rispetto alla curva log-normale. L'esistenza di una significativa differenza tra le distribuzioni può essere valutata attraverso sia un approccio grafico (Grafico 3.3) sia il *dissimilarity index* (Agresti, 2002), che permette di valutare l'accostamento tra due distribuzioni teoriche (l'indice varia tra 0=max *fit* e 1=min *fit*).

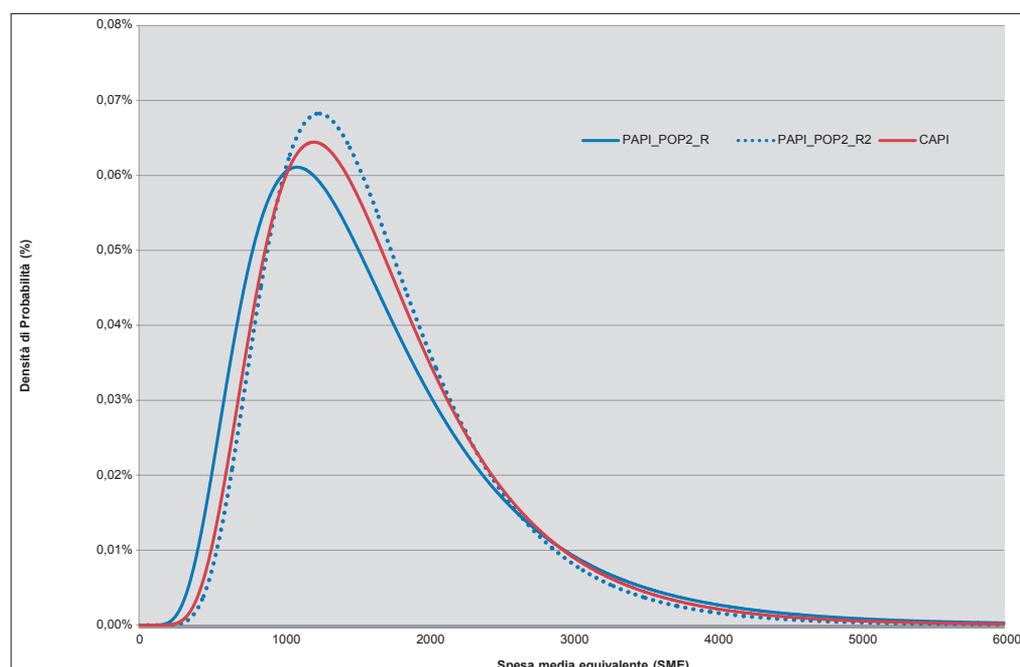
43 In realtà quando sarà disponibile una serie storica CAPI più estesa (4-5 anni) l'ipotesi di una ricostruzione "model based" potrà essere nuovamente presa in considerazione.

44 Identificata a metà degli anni '80 dalla Commissione di Indagine sull'Esclusione Sociale (CIES) che si è avvalsa del contributo di uno statista - Giuseppe Carbonaro (1985).

45 La soglia di povertà è la spesa media mensile pro-capite di un certo anno. Questo valore viene confrontato con la SME delle famiglie. Pertanto una famiglia di un solo componente è povera se il suo livello di consumo è inferiore al 60% della soglia, una famiglia di due componenti è povera se la sua spesa *complessiva* è pari alla soglia e così via utilizzando i coefficienti della "scala Carbonaro".

Nel caso in esame (Sud e Isole, aree metropolitane, età della p.r. < 65), l'indice ha un valore superiore a quello soglia (0.03-0.05) proposto da Agresti⁴⁶, l'accostamento tra le due distribuzioni non è quindi sufficiente buono. La distorsione può essere ridotta utilizzando coefficienti di raccordo costruiti sulla base della distanza tra i decili per numero di componenti.

Grafico 3.3 – Densità di probabilità della SE – Sud e Isole, aree metropolitane, età della p.r. < 65a - dati CAPI - PAPI_POP2_R1 - PAPI_POP2_R2 2013



In altri termini è stato effettuato un *fine tuning* della ricostruzione micro e macro sulla base di un modello teorico di riferimento. In tal modo è possibile calcolare i nuovi livelli della spesa per tutto il periodo 1997-2013 e, simultaneamente, anche l'incidenza della povertà assoluta e della povertà relativa.

Da un punto di vista operativo sono stati identificati i decili teorici della spesa equivalente [SE] sulla base dei dati CAPI e PAPI_POP2_R3 e il valore medio della spesa per decile. Il rapporto tra le medie ha permesso di calcolare degli ulteriori coefficienti di raccordo riferiti a ciascuna tipologia di famiglia composta da tre componenti e più. Naturalmente per applicare i coefficienti è stato necessario calcolare i valori dei decili per tutto il periodo 1997-2013.

Al termine di questa quarta fase di raccordo le stime PAPI_POP2_R4 garantivano un ottimo livello di accostamento con quelle CAPI in termini di spesa media, povertà relativa e assoluta. Naturalmente l'accostamento è stato verificato simultaneamente per le diverse stime e i diversi gruppi di famiglie (tavola 3.7), attraverso un semplice indice di variabilità delle stime, quale il coefficiente di variazione [CV].

46 Nel nostro caso, considerando la natura campionaria dei dati utilizzati, e l'errore ad essi associato, è possibile innalzare il livello di soglia di a 0.05. Le sperimentazioni sull'indice di dissimilarità sono state avviate insieme a M. D'Orazio.

Tavola 3.7 – L’impatto della ricostruzione in termini di variabilità. Valori medi dei coefficienti di variazione nei gruppi (36 cluster)

Ripartizione	PAPI			CAPI
	Vecchia	Nuovi pesi	Ricostruita	
NORD	62,6%	62,3%	56,5%	51,7%
CENTRO	61,3%	61,9%	55,0%	52,1%
SUD	60,1%	61,2%	57,3%	54,1%
ITALIA	61,3%	61,8%	56,2%	52,6%

Esaminando i dati si può verificare come la ricostruzione micro-macro abbia permesso di realizzare un accostamento significativo dei valori delle due indagini (PAPI e CAPI) non solo in termini di livelli di spesa, ma anche in termini di variabilità della spesa stessa.

Ricapitolando, la ricostruzione in serie storica è avvenuta attraverso le seguenti fasi:

1. revisione dei coefficienti di riporto all’universo in base ai nuovi vincoli della CAPI e alle nuove popolazioni censuarie disponibili;
2. definizione di un primo set di 1332 coefficienti di raccordo (36 cluster x 37 voci di spesa);
3. definizione di un secondo set di 21 coefficienti regionali per armonizzare i dati a livello territoriale NUTS 2;
4. definizione di un terzo set di 1332 coefficienti per garantire una coerenza a livello territoriale, di tipologia familiare e voce di spesa;
5. definizione di un quarto set di coefficienti (120 coefficienti diversi per ciascuno dei 17 anni della ricostruzione) per tenere conto della diversa distribuzione della spesa equivalente nelle famiglie con tre componenti e più;
6. definizione di un quinto set di coefficienti finali, che agiscono sia sui coefficienti di riporto all’universo sia sulla spesa e sono differenziati a seconda del tipo di aggregato e del livello di dettaglio, preservando la coerenza complessiva dei dati ed eliminano le piccole differenze residue tra le due indagini (imputabili a fattori diversi, gli errori campionari *in primis*).

3.6 L’esito della ricostruzione: i risultati principali

L’obiettivo della ricostruzione, com’è stato illustrato, era di avere delle nuove serie storiche per la spesa e per la povertà, relativa (1997-2013) e assoluta (2005-2013), che fossero “in linea” con i dati dell’indagine CAPI, utilizzando tutto il patrimonio informativo contenuto nell’indagine PAPI.

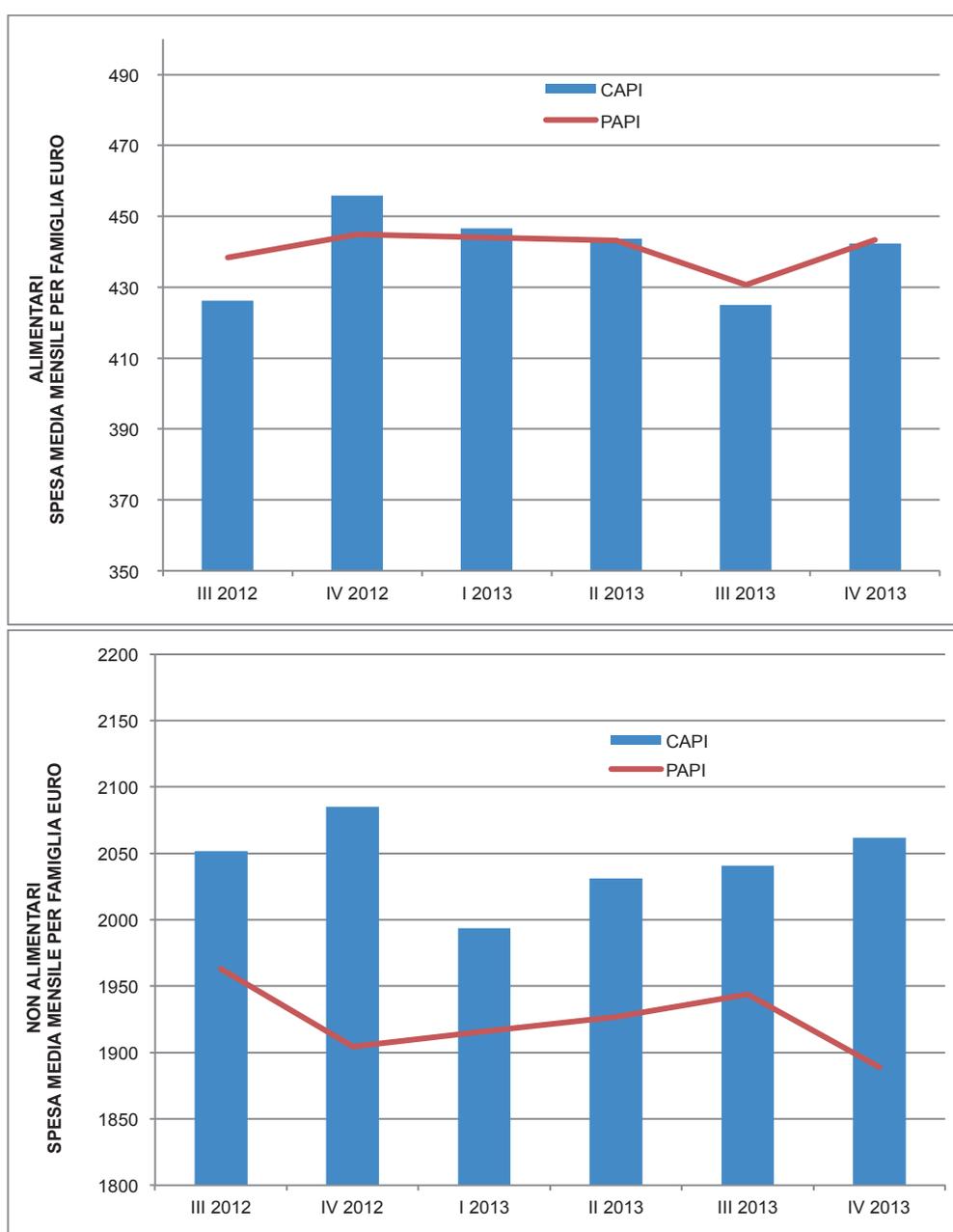
I valori pubblicati fino al 2013 e i dati della nuova indagine presentano rilevanti differenze, a seguito delle innovazioni di natura metodologica introdotte in termini di tecnica d’indagine, disegno di campionamento, questionario, periodo di riferimento delle spese e metodologia di riporto al mese della spesa rilevata.

Nell’illustrare i principali risultati, pertanto, si cercherà di evidenziare le differenze, sottolineando, per quanto possibile, il processo di transizione.

3.6.1 La ricostruzione della spesa

Nei grafici e nelle tabelle di seguito presentate saranno considerati i risultati finali con quattro prospettive diverse: i dati PAPI rilevati [PAPI], i dati PAPI rilevati e modificati in base alle nuove popolazioni e al nuovo sistema di ponderazione e calibrazione [PAPI_NP], i dati finali ricostruiti [PAPI_CAPI_R] e i dati CAPI rilevati [CAPI].

Grafico 3.4 – Spesa media mensile alimentare e non alimentare per trimestre – CAPI, PAPI



Si proporrà, innanzitutto un confronto tra i valori PAPI e quelli CAPI ovvero la spesa alimentare e non alimentare⁴⁷ così come rilevata nei sei trimestri di sovrapposizione (grafico 3.4).

La significatività e la rilevanza delle differenze evidenziate nel grafico è stata diffusamente illustrata nelle sezioni precedenti. Due indicatori offrono un'ulteriore misura di queste differenze. In particolare si fa riferimento alla Media della Distanza Assoluta Percentuale (MDAP) che individua sulla base della [8] la differenza nei "livelli" delle due serie:

$$MDAP = \frac{\sum_{III\ 2012}^{IV\ 2013} |CAPI - PAPI|}{\sum_{III\ 2012}^{IV\ 2013} CAPI} * 100 \quad [8]$$

Le differenze nella dinamica possono essere valutate in base alla [9] che identifica la Media della Distanza Prima Percentuale (MDPP⁴⁸):

$$MDPP = \frac{\sum_{III\ 2012}^{IV\ 2013} |\overline{CAPI} - \overline{PAPI}|}{\sum_{III\ 2012}^{IV\ 2013} |\overline{CAPI}| + \sum_{III\ 2012}^{IV\ 2013} |\overline{PAPI}|} * 100 \quad [9]$$

Utilizzando i dati presentati nel precedente grafico 3.4 è possibile calcolare gli indici MDAP e MDPP. Il primo assume valori di 1,2 (alimentari) e 5,9 (non alimentari) mentre il secondo varia tra 39,4 (alimentari) e 88,1 (non alimentari). Nel primo caso la lettura di questi valori è abbastanza immediata, l'indicatore individua la distanza media assoluta tra le due serie (in percentuale). I consumi alimentari CAPI nei sei trimestri in esame sono distanti in media l'1,2% rispetto ai corrispondenti PAPI; nel caso degli alimentari il 5,9%.

Il secondo indicatore permette di valutare l'accostamento tra le variazioni delle due serie. All'indicatore MDPP, utilizzato per le considerazioni sull'esito della ricostruzione, che non si presta ad una lettura immediata, è utile associare una scala di misura qualitativa:

- MDPP < 10 sincronia ottimale delle serie;
- 10 < MDPP < 25 accostamento buono;
- 26 < MDPP < 50 accostamento sufficiente;
- MDPP > 50 accostamento scarso o asincronia delle serie.

La serie CAPI dei consumi non alimentari è sostanzialmente asincrona rispetto a quella PAPI (MDPP=88,1) mentre la situazione è migliore per la spesa alimentare (MDPP=39,4) così come evidenziato dal grafico.

Le indagini CAPI e PAPI sono alla base delle stime delle povertà e le differenze ora segnalate sono riscontabili anche in quei casi. Per la povertà relativa, ad esempio, nei sei trimestri della doppia rilevazione, l'incidenza a livello nazionale delle famiglie povere si discosta di circa il 2% tra CAPI e PAPI: il valore stimato con quest'ultima era in media pari al 12,7% (2012-2013), con CAPI è più basso.

47 Naturalmente la distinzione tra spesa alimentare e non alimentare si basa sulla nuova classificazione Coicop che, come illustrato, a questo è livello di dettaglio è lievemente differente da quella adottata in precedenza. Inoltre nei consumi non alimentari sono incluse le spese per la manutenzione straordinaria dell'abitazione che non sono considerate all'interno della nuova classificazione europea.

48 Gli indicatori MDAP e MDPP sono un'estensione di quelli presentati in un precedente documento (STD/NA[99]37). Il primo ha un campo di variazione che varia tra zero e infinito mentre il secondo varia tra zero e cento.

Nel caso della povertà assoluta, inoltre, l'incidenza a livello nazionale delle famiglie assolutamente povere passa dal 5,6% stimato in precedenza (PAPI 2012-2013) ad un valore che è inferiore di circa l'1%.

I macro-dati commentati ovviamente nascondono una realtà molto articolata, con differenze tali da giustificare il lungo percorso di riconciliazione dei dati fin qui descritto.

Fissato il punto di partenza, e tenendo conto dei dati iniziali, si farà ora un rapido balzo in avanti per mettere a confronto la serie PAPI 1997-2013 con la nuova serie PAPI_CAPI_R. Il quadro finale risulta armonizzato. Utilizzando la [8] e la [9] per mettere a confronto l'indagine PAPI con i dati PAPI_CAPI_R per tutto il periodo 1997-2013 è possibile ricalcolare gli indici MDAP e MDPP per l'Italia e le 5 ripartizioni, oltre che per i consumi alimentari, non alimentari e il totale della spesa annuale (Tavola 3.8).

Tavola 3.8 – L'impatto "lordo" della ricostruzione (1997-2013) su consumi alimentari e non alimentari e totale – Italia e 5 ripartizioni - Indici MDAP e MDPP

Indicatore	Nord Ovest	Nord Est	Centro	Sud	Isole	ITALIA
Alimentari						
MDAP (livelli)	0,61	4,10	1,86	1,83	4,00	1,20
MDPP (variazioni)	10,93	13,20	11,55	7,09	8,60	7,82
Non Alimentari						
MDAP (livelli)	4,77	3,67	9,00	9,29	17,92	7,26
MDPP (variazioni)	10,18	12,33	19,95	12,28	22,28	7,86
TOTALE						
MDAP (livelli)	3,97	3,74	7,20	7,65	7,26	6,20
MDPP (variazioni)	10,90	11,92	17,08	10,84	20,26	7,81

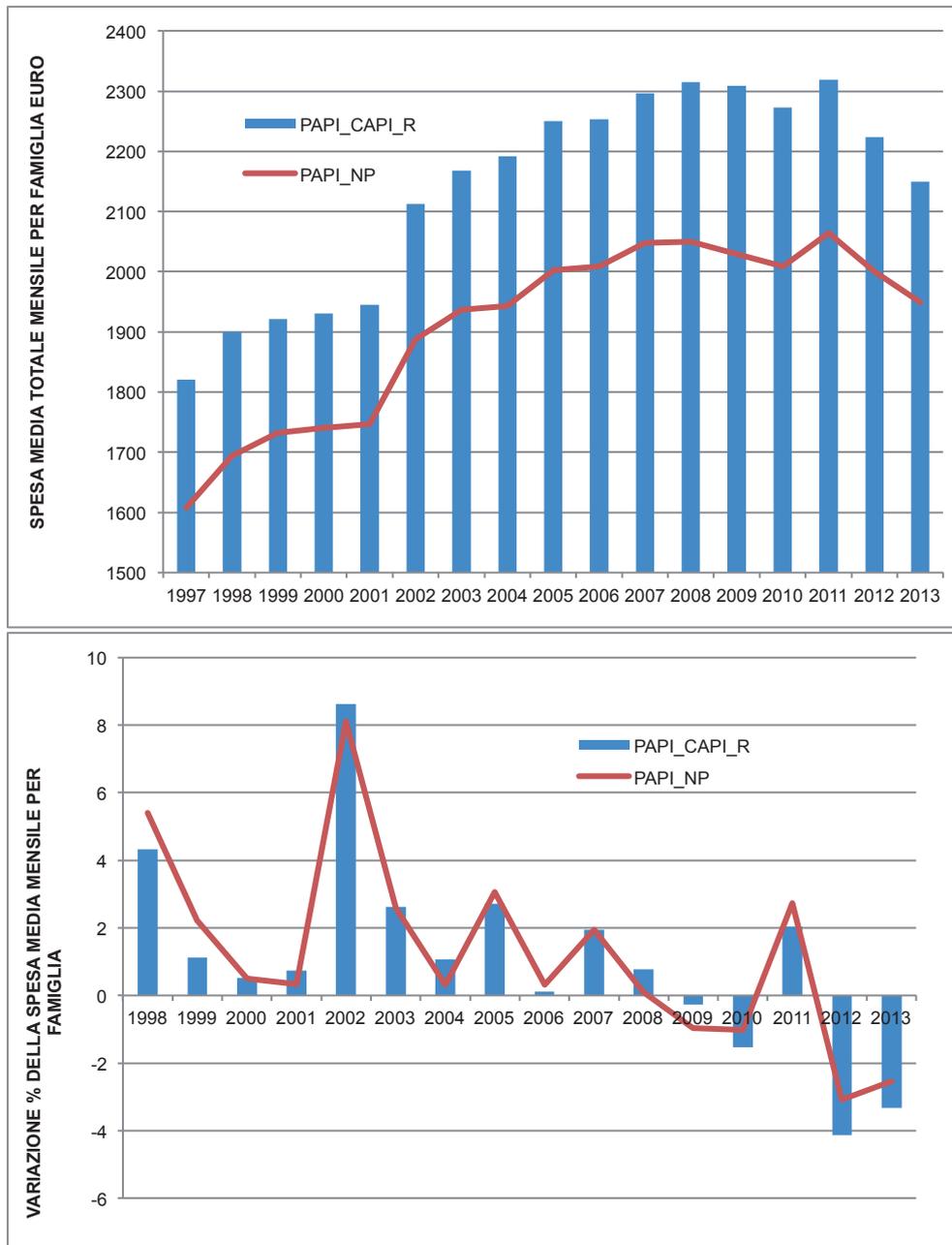
A livello globale la revisione (MDAP) ha comportato una modifica media dei consumi alimentari del 1,2%, in linea con quanto già stimato sulla base dei soli sei trimestri di sovrapposizione. Per quanto riguarda i consumi non alimentari la nuova serie si discosta, in media, del 7,26% rispetto alla precedente, un valore leggermente maggiore di quello riferito al solo periodo di doppia rilevazione (5,9%).

A livello nazionale e sul totale della spesa, la revisione media assoluta percentuale nel periodo 1997-2013 è stata del 6,20% e, com'era atteso, non è stata costante nel tempo (nell'ultimo anno ad esempio la differenza è stata inferiore al 4,7%).

Per quanto riguarda l'indicatore della dinamica (MDPP) evidenzia il più delle volte una sincronia ottimale tra le due serie. Valori più elevati si osservano al Centro e nelle Isole e ciò a causa del fatto che si è in presenza di una parziale modificazione del "rapporto di forza" tra le regioni o dell'aggregazione di due sole regioni. In particolare la Sardegna nella CAPI presenta dei valori medi significativamente diversi da quelli rilevati con l'indagine PAPI.

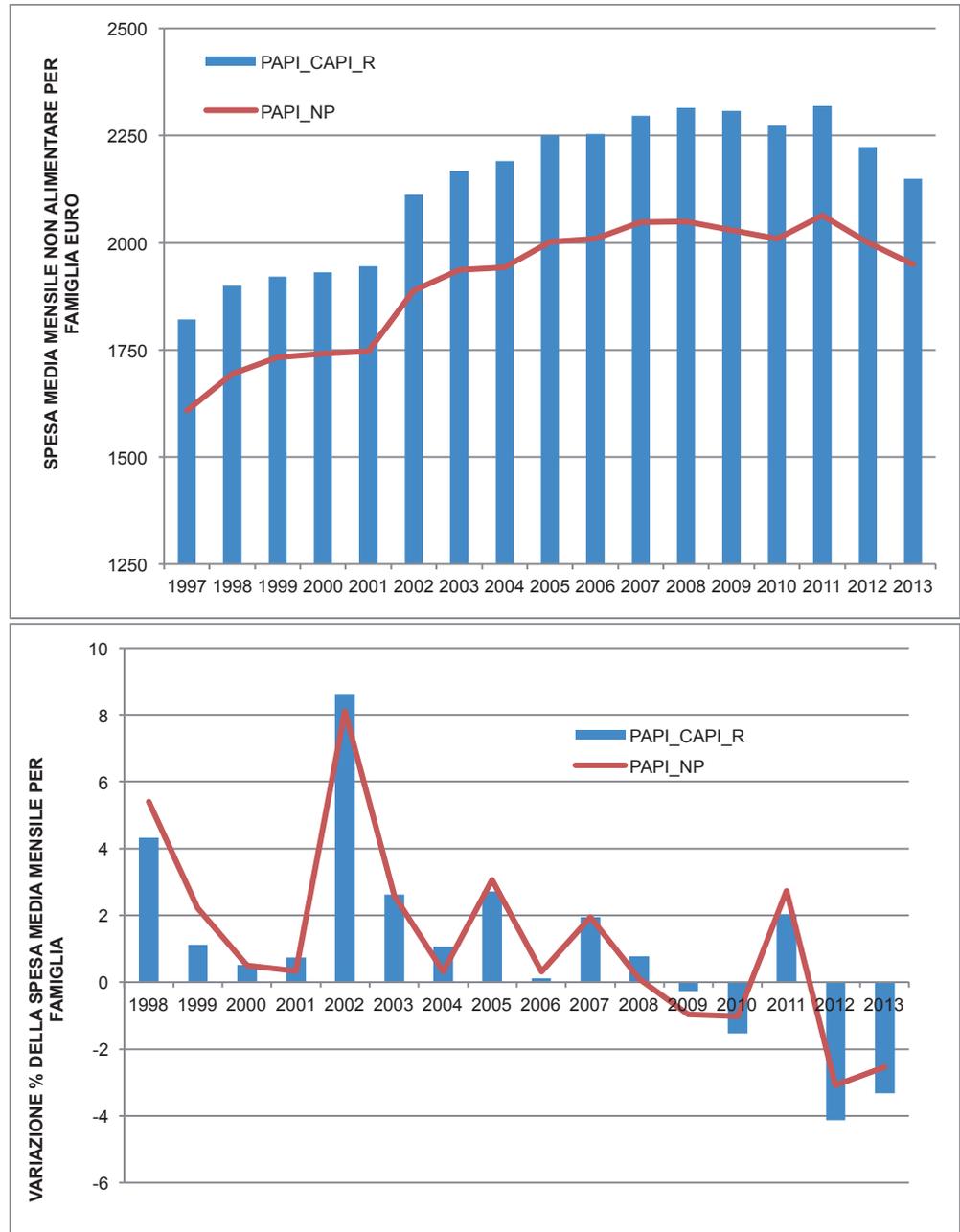
Queste valutazioni trovano conferma utilizzando l'approccio grafico. Analizzando l'andamento della spesa totale si può osservare come la nuova serie ricostruita, re-cepisca la differenza in termini di livelli assoluti conservando il profilo della dinamica in serie storica (Grafico 3.5).

Grafico 3.5 – Spesa media mensile totale e variazione percentuale rispetto all'anno precedente PAPI_CAPI_R e PAPI - Italia Anni 1997-2013



Vale inoltre la pena osservare che il valore di MDAP è piuttosto elevato nel Centro (7,2%), in particolare per la spesa non alimentare (9,0%), a causa degli effetti della revisione dei dati del Lazio. La nuova serie tuttavia presenta un andamento prossimo a quella precedente e l'accostamento dinamico (MDPP=19,95) è buono (Grafico 3.6).

Grafico 3.6 – Spesa media mensile non alimentare PAPI_CAPI_R e PAPI - Centro. Anni 1997-2013



Il confronto tra i dati PAPI_NP e quelli PAPI dà conto dell'impatto delle nuove popolazioni e dei nuovi pesi al netto di tutti gli altri interventi (micro e macro). L'indice MDPP (Tabella 3.9) mostra differenze nella dinamica simili a quelle ottenute con la ricostruzione finale a dimostrazione del fatto che l'introduzione del nuovo sistema di ponderazione e delle nuove popolazioni spiega da solo una parte consistente della nuova dinamica della spesa.

Tavola 3.9 – L’impatto del nuovo sistema di ponderazione e delle nuove popolazioni (1997-2013) – Italia e 5 ripartizioni - Indici MDAP e MDPP

Indicatore	Nord Ovest	Nord Est	Centro	Sud	Isole	ITALIA
Alimentari						
MDAP (livelli)	2,17	0,49	0,18	1,82	2,60	0,12
MDPP (variazioni)	11,33	7,67	10,82	5,35	10,91	7,22
Non Alimentari						
MDAP (livelli)	0,76	1,84	5,12	2,77	1,05	2,11
MDPP (variazioni)	9,50	11,52	15,22	11,37	18,34	8,46
TOTALE						
MDAP (livelli)	0,26	1,63	4,12	2,55	2,11	1,74
MDPP (variazioni)	9,94	11,11	12,92	10,31	18,40	8,10

Se riconsideriamo, ad esempio, i consumi non alimentari del Centro notiamo che l’indice MDPP ha un valore elevato (15,22) così come lo era quello calcolato utilizzando i dati ricostruiti (19,95) e un discorso analogo vale nel caso delle Isole.

Per quanto riguarda i livelli (MDAP), poi, come ci si poteva attendere, le differenze sono piccole e abbastanza omogenee (e in generale le revisioni hanno comportato una riduzione nei livelli di spesa).

Per depurare i dati dagli effetti del nuovo sistema di ponderazione è possibile mettere a confronto i dati PAPI_NP con quelli ricostruiti (PAPI_CAPI_R). Si può verificare, così, quale sia stato l’impatto della sola ricostruzione macro.

I dati ricostruiti [PAPI_CAPI_R] risultano più lontani dai valori PAPI_NP in termini di livelli di spesa, ma più vicini in termini di dinamica.

La rivalutazione media dei consumi alimentari (MDAP) è del 2,21% e quella della spesa non alimentare è del 9,42% (Tabella 3.10).

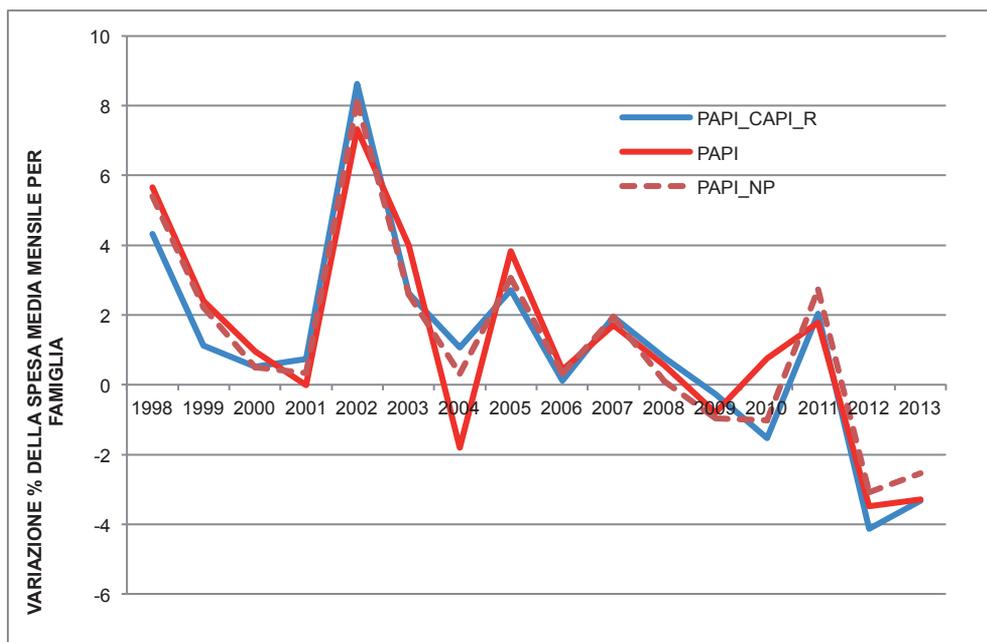
In termini di dinamica la sincronizzazione delle due serie è quasi sempre ottimale (MDPP<10) con un valore massimo ancora una volta al Centro (12,66) e dei miglioramenti significativi in tutti i casi e in tutte le ripartizioni.

Tavola 3.10 – L’impatto “netto” della ricostruzione (1997-2013) – Italia e 5 ripartizioni - Indici MDAP e MDPP

Indicatore	Nord Ovest	Nord Est	Centro	Sud	Isole	ITALIA
Alimentari						
MDAP (livelli)	0,63	4,87	0,80	2,79	5,15	2,21
MDPP (variazioni)	5,36	10,35	4,65	5,21	9,58	6,49
Non Alimentari						
MDPP (variazioni)	6,66	5,92	10,80	11,50	20,26	9,42
MDPP (variazioni)	5,11	3,68	12,66	8,59	12,63	6,49
TOTALE						
MDAP (livelli)	5,67	5,76	8,88	9,59	9,42	8,16
MDPP (variazioni)	5,23	3,62	11,32	7,57	12,28	5,96

Da notare come per le regioni centrali le stime PAPI_CAPI_R siano decisamente più vicine in dinamica ai valori PAPI_NP rispetto ai vecchi dati PAPI.

Grafico 3.7 – Confronto tra le variazioni della spesa non alimentare Centro Italia PAPI_CAPI_R, PAPI_NP e PAPI - Anni 1998-2013



Esaurita questa (parziale) rassegna dei principali risultati della ricostruzione vediamo ora quali sono state le tipologie di dati ricostruite. La scelta di procedere a una ricostruzione simultanea (micro e macro) fino al massimo punto di accostamento (PAPI_POP2_R4) ha permesso di predisporre serie storiche ricostruite ad un livello di dettaglio particolarmente ampio.

In generale, per gli anni 1997-2013, è stata ricostruita la spesa per capitolo. Inoltre sono stati ricostruiti i dati in base alle principali variabili di classificazione utilizzate per l'analisi dei dati di spesa e correntemente rilasciate (ripartizione e regione geografica, ampiezza e tipologia familiare, alcune caratteristiche della persona di riferimento della famiglia).

3.6.2 La ricostruzione delle incidenze di povertà

L'impatto della ricostruzione sulla povertà relativa, misurato considerando la distanza dei dati CAPI_R rispetto a PAPI e PAPI_NP, è molto simile a quello osservato per la spesa media.

La soglia di povertà risulta in media più elevata del 7,52% rispetto a quella calcolata con la PAPI. Le variazioni della soglia calcolata con CAPI_R [MDPP] sono molto più vicine e sincrone a PAPI_NP rispetto a PAPI (Tabella 3.11). Per quanto riguarda le incidenze il discorso cambia leggermente.

Innanzitutto l'indice sui livelli (MDPP) viene calcolato su dei valori percentuali (le incidenze) e dunque la differenza del 10,43% è una "percentuale di una percentuale"

e risente in misura maggiore dell'ordine di grandezza della variabile di riferimento. Nei 17 anni di osservazione (1997-2013) l'incidenza della povertà media ricostruita [PAPI_CAPI_R] in Italia è risultata pari al 10,45%. Sulla base dell'indice MDAP si può affermare che il suo scostamento medio in livello dalla PAPI è dell'1,09% calcolato sulla base del MDAP e dell'incidenza media e dunque un valore relativamente contenuto. D'altra parte è certo che lo scostamento della povertà nelle isole (MDAP=34,30) è il più elevato in assoluto.

Tavola 3.11 – L'impatto della ricostruzione (1997-2013) sulla povertà relativa: indici MDAP e MPP sulla soglia e sulle incidenze (famiglie) per ripartizione

Indicatore	CAPI_PAPI_R vs PAPI	CAPI_PAPI_R vs PAPI_NP	Soglia			
MDAP (livelli)	7,52	8,90				
MDPP (variazioni)	7,90	3,52				
Indicatore	Nord Ovest	Nord Est	Centro	Sud	Isole	ITALIA
	CAPI_PAPI_R vs PAPI					
MDAP (livelli)	6,77	10,22	9,30	8,82	34,30	10,43
MDPP (variazioni)	24,52	30,39	37,22	19,01	28,79	27,91
CAPI_PAPI_R vs PAPI_NP						
MDAP (livelli)	6,48	11,18	8,48	8,71	34,89	11,40
MDPP (variazioni)	20,04	24,67	39,36	19,26	31,33	23,45

Se confrontiamo l'incidenza al Nord Est e al Sud facendo riferimento ai dati CAPI_R e ai dati PAPI ci troviamo davanti nel primo caso (Nord Est) a un progressivo allontanamento delle due serie che si manifesta a partire dal 2005. Nel secondo caso (Sud) la distanza rimane sostanzialmente invariata per tutto il periodo di osservazione (Grafico 3.8).

Quando consideriamo l'indicatore MDPP sulle variazioni il ragionamento torna ad essere quello fatto a proposito della spesa e anche la scala qualitativa è la stessa. Si nota che in generale l'accostamento è buono (valori prossimi a 25 e in molti casi inferiori). Il maggior accostamento ai dati PAPI_NP è presente anche in questo caso.

Per quanto concerne il Centro l'accostamento in termini di livelli è buono (MDAP=9,30) mentre in termini di dinamica (Grafico 3.9), la differenza tra le due serie risulta a volte più accentuata anche se, probabilmente, non sempre statisticamente significativa perché inferiore all'ampiezza dell'intervallo di confidenza delle stime stesse⁴⁹.

⁴⁹ La problematica dell'intervallo di confidenza dei valori ricostruiti è stata presa in considerazione ma non può essere "risolta" data la complessità teorica di tale misura in assenza di un *data-set* completo ricostruito.

Grafico 3.8 – Confronto tra le incidenze della povertà relativa PAPI_CAPI_R e PAPI - Nord Est [sopra] Sud [sotto] Anni 1997-2013

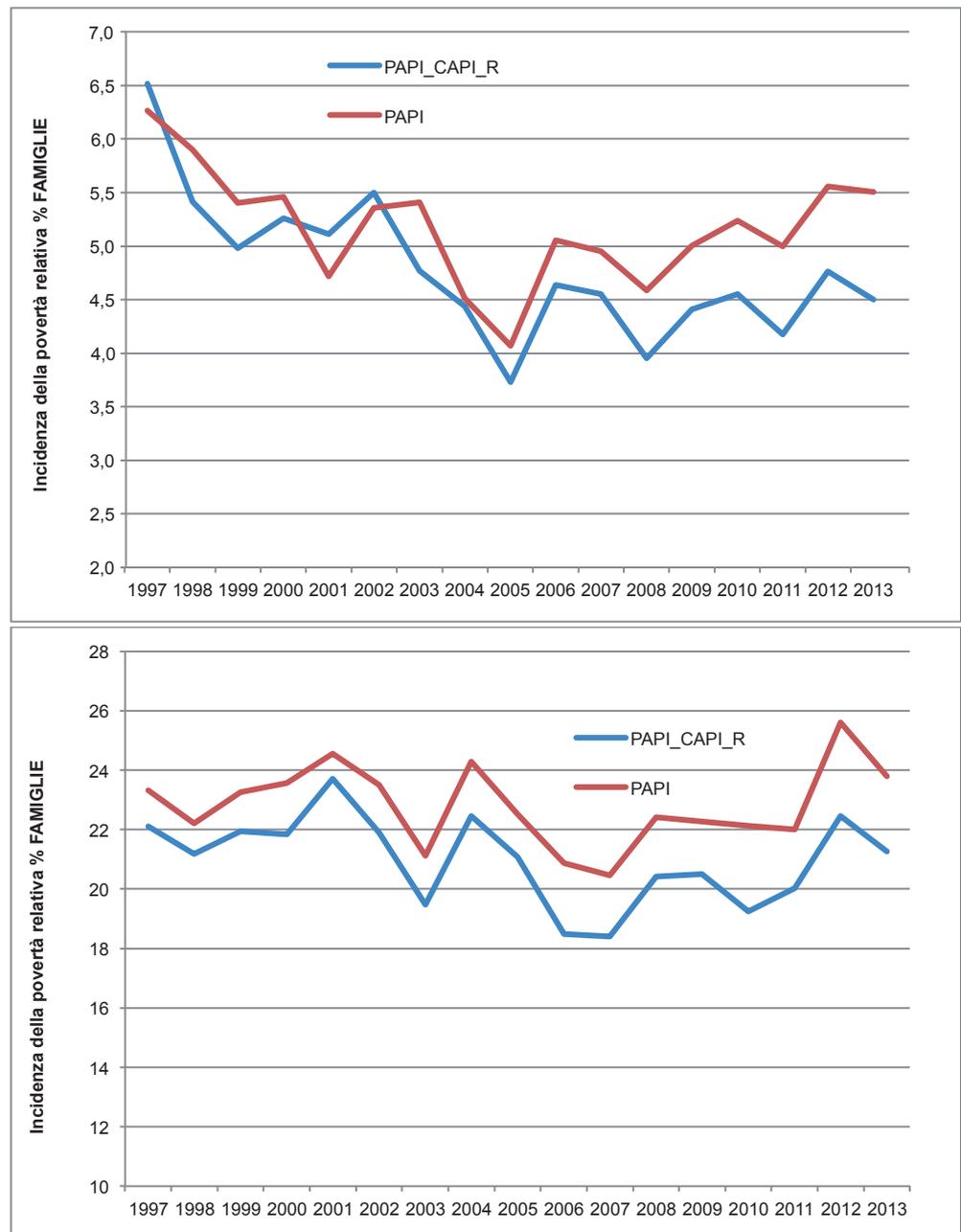
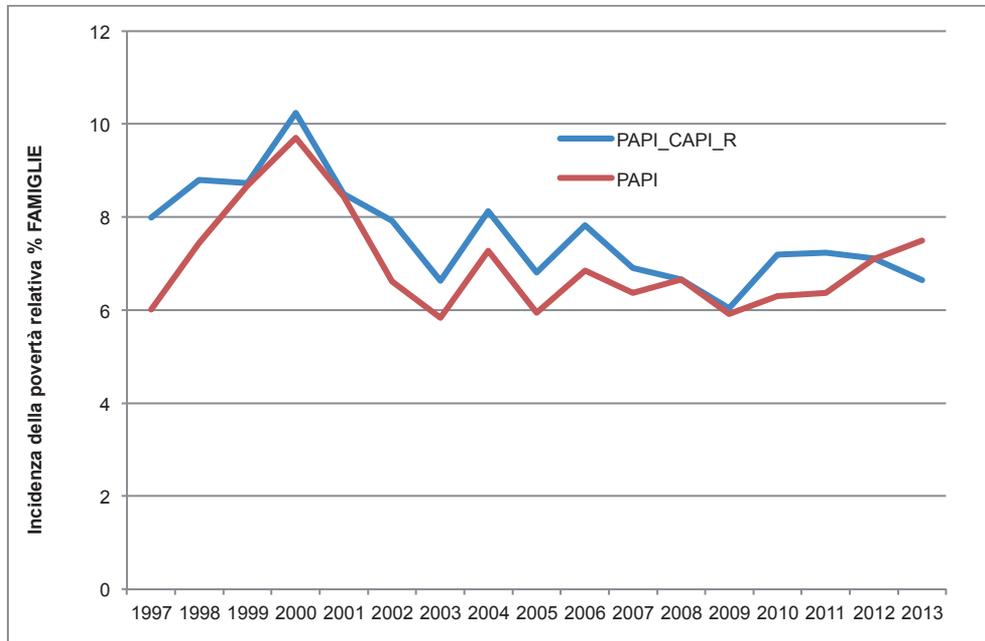


Grafico 3.9 – Confronto tra le incidenze della povertà relativa CAPI_R e PAPI Centro Anni 1997-2013



Qualche cenno, infine, sulla ricostruzione della povertà assoluta. Il nuovo andamento della povertà assoluta è simile a quello della povertà relativa con un lieve progressivo allontanamento della nuova CAPI_R dalla PAPI. Il nuovo livello della povertà assoluta sale dal 3,1% (2005) al 5,5% (2013) con i dati PAPI_CAPI_R (+2,4%) mentre passava da 3,9% a 7,9% con i dati PAPI (+4,0%).

Le incidenze di povertà relativa e assoluta sono state ricostruite a partire dal primo anno disponibile: 1997 per l'incidenza di povertà relativa a livello nazionale e ripartizionale, 2002 per le incidenze di povertà relativa a livello regionale e 2005 per le incidenze di povertà assoluta.

3.7 Conclusioni

La scelta di utilizzare un approccio micro-macro ha permesso di ricostruire la serie storica con un ampio dettaglio, riferito alle principali variabili di analisi (spesa media per capitolo, povertà relativa e assoluta) e di classificazione (ripartizione, regione, ampiezza e tipologia familiare, condizione occupazionale e professionale della persona di riferimento).

La ricostruzione include anche l'adeguamento alle popolazioni stimate a seguito delle risultanze del Censimento del 2011 (che ha comportato una revisione delle stime anagrafiche) e ai nuovi vincoli di calibrazione, introdotti nel calcolo dei pesi finali (popolazione e famiglie residenti per ripartizione e tipologia di comune, popolazione straniera residente per ripartizione e sesso, dal 2009, popolazione di 15 anni e più per ripartizione, condizione e posizione nella professione, popolazione e famiglie residenti per ripartizione e mese) per migliorare l'accuratezza delle stime, anche per sottogruppi di popolazione, e tenere conto del carattere fortemente stagionale di alcune tipologie di spesa.

I dati contenuti nelle serie storiche ricostruite, in termini sia di spesa per consumi sia di povertà, presentano rilevanti differenze con la vecchia serie storica, a seguito delle innovazioni di natura metodologica introdotte in termini di tecnica d'indagine, disegno di campionamento, questionario, periodo di riferimento delle spese e metodologia di riporto al mese della spesa rilevata. La ricostruzione in serie storica evidenzia come, soprattutto per effetto delle modifiche apportate al dettaglio di spesa e ai periodi di riferimento, la nuova serie si attesti su livelli più alti della precedente e come l'aumento sia stato più accentuato tra le famiglie appartenenti alle code della distribuzione, tra le famiglie residenti nel Mezzogiorno e tra quelle più ampie.

Sebbene su livelli mediamente più elevati, la nuova serie segue la stessa dinamica della vecchia, sia in termini di spesa sia in termini di povertà, riproponendo tutti gli andamenti osservati nel corso degli anni, soprattutto a seguito degli evidenti effetti della crisi che hanno sostanzialmente modificato le condizioni economiche della popolazione.

Grafico 3.10 – Spesa media mensile familiare PAPI_CAPI_R (nuova serie) e PAPI (vecchia serie). Anni 1997-2013

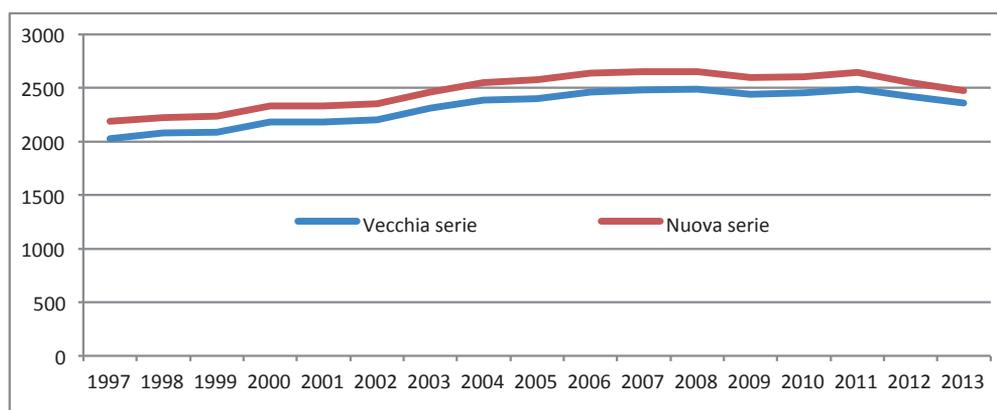


Grafico 3.11 – Incidenza di povertà relativa tra le famiglie PAPI_CAPI_R (nuova serie) e PAPI (vecchia serie). Anni 1997-2013

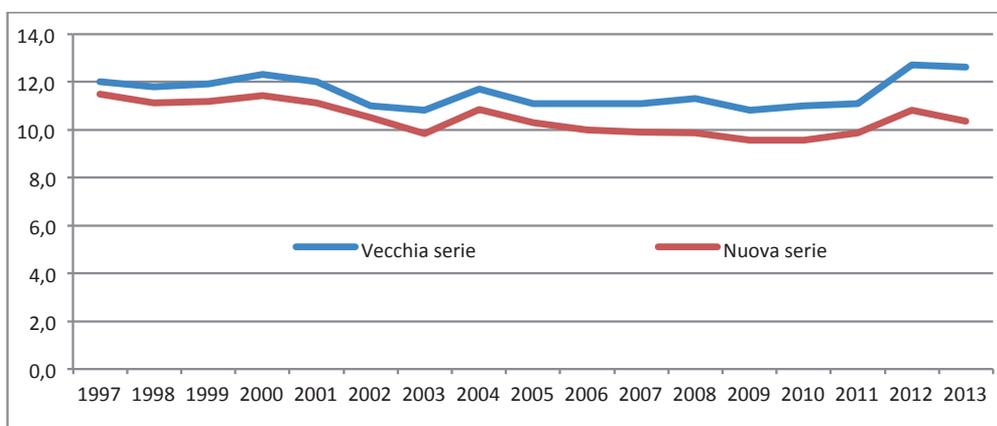
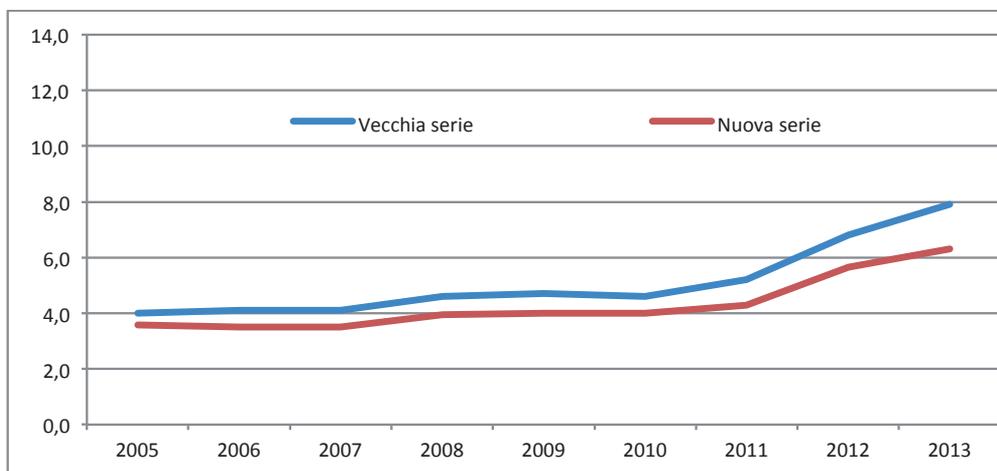


Grafico 3.12 – Incidenza di povertà assoluta tra le famiglie PAPI_CAPI_R (nuova serie) e PAPI (vecchia serie). Anni 2005-2013



GLOSSARIO

Cittadini stranieri: persone, nate in Italia o all'estero, di cittadinanza straniera o apolide.

Classificazione dei consumi individuali secondo lo scopo (Coicop): è una classificazione armonizzata a livello internazionale delle voci di spesa secondo lo scopo (Classification of Individual Consumption According by Purpose). Predisposta dalla Divisione statistica delle Nazioni unite per consentire il confronto dei comportamenti di consumo tra paesi, la Coicop è costituita da 14 capitoli di spesa, a loro volta articolati in categorie e in gruppi.

Consumi delle famiglie: l'insieme delle spese sostenute per tutti i beni e servizi destinati al soddisfacimento dei bisogni familiari e individuali. Include: i) la stima del valore del canone mensile delle abitazioni di proprietà a disposizione della famiglia comprensive delle pertinenze (fitti figurativi) in quanto rappresentano il costo del servizio "abitare" che la famiglia consuma nel corso del tempo; ii) il valore monetario dei prodotti autoconsumati dalla famiglia, provenienti cioè dalla propria impresa (individuale o familiare), dalla coltivazione diretta (inclusa la raccolta di prodotti spontanei), dall'allevamento (incluse la caccia e la pesca) o ricevuti come compenso per la propria attività lavorativa; iii) i beni e servizi, acquistati o autoprodotti, regalati a persone esterne alla famiglia. Esclude le spese sostenute per i) attività professionale; ii) trasferimenti o donazioni; iii) accrescimento del patrimonio; iv) tasse e tributi. In Italia le spese per consumi delle famiglie vengono rilevate attraverso l'Indagine sulle spese.

Consumi finali: rappresentano il valore dei beni e servizi impiegati per soddisfare direttamente i bisogni umani, siano essi individuali o collettivi. Sono utilizzati due concetti: la spesa per consumi finali e i consumi finali effettivi. La differenza fra i due concetti sta nel trattamento riservato ad alcuni beni e servizi che sono finanziati dalle amministrazioni pubbliche o dalle istituzioni senza scopo di lucro al servizio delle famiglie, ma che sono forniti alle famiglie come trasferimenti sociali in natura; questi beni sono compresi nel consumo effettivo delle famiglie, mentre sono esclusi dalla loro spesa finale (Sistema europeo dei conti, Sec 2010).

Contabilità nazionale: l'insieme di tutti i conti economici che descrivono l'attività economica di un paese o di una circoscrizione territoriale. Essa ha per oggetto l'osservazione quantitativa e lo studio statistico del sistema economico o dei sub-sistemi che lo compongono a diversi livelli territoriali.

Errore campionario: è l'errore che si commette osservando solo una parte della popolazione e non l'intera popolazione, il suo calcolo permette di costruire un intervallo di confidenza intorno alla stima puntuale. Nel 2014 la stima della spesa media mensile è risultata pari a 2.488,50 euro; poiché l'errore relativo è pari allo 0,5%, il valore che si otterrebbe osservando l'intera popolazione è compreso, con una probabilità

del 95%, tra 2.463,06 e 2.513,94 euro. Quando si scende nel dettaglio delle caratteristiche familiari, la minore numerosità campionaria fa aumentare l'errore relativo e rende quindi le stime meno precise.

Famiglia: insieme di persone legate da vincoli di matrimonio, parentela, affinità, adozione, tutela o da vincoli affettivi, coabitanti e aventi dimora abituale nello stesso comune. Una famiglia può essere costituita anche da una sola persona. L'assente temporaneo non cessa di appartenere alla propria famiglia sia che si trovi presso altro alloggio (o convivenza) dello stesso comune, sia che si trovi in un altro comune italiano o all'estero.

Fitto figurativo: è una componente non-monetaria del reddito delle famiglie o della spesa per consumi delle famiglie che vivono in case di loro proprietà; rappresenta il costo che queste dovrebbero sostenere per prendere in affitto, ai prezzi vigenti sul mercato immobiliare, un'unità abitativa con caratteristiche identiche a quella in cui vivono (al netto delle spese di condominio, riscaldamento, accessorie e con riferimento a una casa non ammobiliata). Negli studi sulla povertà e sulla distribuzione del reddito, il concetto viene esteso anche alle famiglie in usufrutto o in uso gratuito e agli inquilini con affitti agevolati, cioè inferiori ai prezzi di mercato.

Incidenza della povertà: si ottiene dal rapporto tra il numero di famiglie con spesa media mensile per consumi pari o al di sotto della soglia di povertà e il totale delle famiglie residenti.

Paniere di povertà assoluta: rappresenta l'insieme dei beni e servizi che, nel contesto italiano, vengono considerati essenziali per una determinata famiglia per conseguire uno standard di vita minimamente accettabile.

Persona di riferimento: Persona rispetto alla quale sono definite le relazioni di parentela, generalmente corrispondente all'intestatario della scheda anagrafica familiare.

Prezzi al consumo (indice dei): la variazione nel tempo dei prezzi che si formano nelle transazioni relative a beni e servizi scambiati tra gli operatori economici e i consumatori privati finali.

Per le famiglie di operai e impiegati (Foi) – La variazione nel tempo dei prezzi al dettaglio, dei beni e servizi correntemente acquistati dalle famiglie di lavoratori dipendenti.

Per l'intera collettività (Nic) – La variazione nel tempo dei prezzi relativi ai beni e servizi acquistati sul mercato per i consumi finali individuali.

Indice armonizzato dei prezzi al consumo (Ipc) – È stato sviluppato per assicurare una misura dell'inflazione comparabile a livello europeo. Si differenzia dagli altri due indici perché si riferisce al prezzo effettivamente pagato dal consumatore e perché esclude dal suo campo di definizione alcune voci che sono invece presenti nel paniere dell'indice nazionale. Inoltre, a differenza degli altri indici dei prezzi al consumo, l'indice armonizzato tiene conto anche delle riduzioni temporanee di prezzo (saldi, sconti e promozioni).

Scala di equivalenza: sono utilizzate nelle indagini statistiche per uniformare unità di analisi eterogenee (esempio consumi e redditi delle famiglie) mediante l'utilizzo

di appropriati coefficienti di correzione. L'indagine sulle spese per la stima della spesa equivalente e delle soglie di povertà relativa utilizza la scala Carbonaro (Prospetto G1).

Prospetto G1 – Scala di equivalenza Carbonaro

Ampiezza della famiglia	Scala di equivalenza
1	0,60
2	1,00
3	1,33
4	1,63
5	1,90
6	2,16
7 o più	2,40

Soglia di povertà assoluta: rappresenta la spesa minima necessaria per acquisire i beni e servizi inseriti nel paniere di povertà assoluta. La soglia di povertà assoluta varia, per costruzione, in base alla dimensione della famiglia, alla sua composizione per età, alla ripartizione geografica e alla dimensione del comune di residenza.

Soglia di povertà relativa: per una famiglia di due componenti è pari alla spesa media procapite nel Paese.

Spesa equivalente: è calcolata dividendo il valore familiare della spesa per il coefficiente della scala di equivalenza e permette di rendere direttamente confrontabili i livelli di spesa di famiglie di ampiezza diversa.

Spesa media mensile: è calcolata dividendo la spesa totale per il numero delle famiglie residenti in Italia.

Spesa mediana mensile: è il valore di spesa per consumi che divide la distribuzione di frequenza in due parti uguali (il 50% delle famiglie presenta un valore di spesa per consumi inferiore o pari alla mediana, il 50% un valore superiore). Poiché la spesa per consumi ha una distribuzione asimmetrica, la mediana risulta sempre inferiore al valore medio.

Spesa media per persona (procapite): si ottiene dividendo la spesa totale per consumi delle famiglie per il numero totale dei componenti.

Tipo di comune: *area metropolitana* - comuni con oltre 250.000 abitanti; *grandi comuni* - comuni della periferia dell'area metropolitana e comuni con 50.000 abitanti e più; *piccoli comuni* - comuni con meno di 50.000 abitanti.

Variazione congiunturale: variazione percentuale rispetto al mese o al periodo precedente.

Variazione tendenziale: variazione percentuale rispetto allo stesso mese o periodo dell'anno precedente.

RIFERIMENTI BIBLIOGRAFICI

- AAPOR., 2000. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. Ann Arbor, Michigan.
- Bali N., Federici A., 2014. *Le strategie formative degli attori coinvolti nel lavoro sul campo*. In Freguja, Romano (a cura di) *La modernizzazione delle tecniche di rilevazione nelle indagini socio-economiche sulle famiglie*, Istat, Roma. <http://www.istat.it/it/archivio/145721>
- Barcherini S., Falorni C., Grassi D., 2014. *La tecnica Capi nell'indagine sulle spese delle famiglie*. In Freguja, Romano (a cura di) *La modernizzazione delle tecniche di rilevazione nelle indagini socioeconomiche sulle famiglie*, Istat, Roma. <http://www.istat.it/it/archivio/145721>
- Bracci L, Di Leo F., Pisani S., 199. *Retrapolating Italian Annual National Accounts Data According to ESA95*. Doc OECD STD/NA(99)37, Paris.
- Bradburn N.M., 2010. *Recall Period in Consumer Expenditure Surveys Program*. NORC/ University of Chicago
http://www.bls.gov/cex/methwrkshp_pap_bradburn.pdf
- Browning M., Crossley T., Winter, J., 2014. *The Measurement of Household Consumption Expenditures*. Institute for Fiscal Studies, Working paper W14/07. Economic and social research council. London, UK. <http://www.ifs.org.uk/wps/wp201407.pdf>
- Canberra Group., 2011. *Expert Group on Household Income Statistics, Final Report and Recommendations*. Ottawa
http://www.unece.org/fileadmin/DAM/stats/groups/cgh/Canberra_Handbook_2011_WEB.pdf
- Ceccarelli C., Cutillo A., 2007. *Il Trattamento della Mancata Risposta Totale nell'Indagine EU-Silc: Una Valutazione Tramite una Misura del Cambiamento*. Congiuntura., Vol.1° Trimetre, Udine: CREF, pp. 91-112.
- Ceccarelli, C., Coppola, L., Cutillo, A., Di Laurea D., 2008. *Combining survey and administrative data in the Italian EU-Silc experience: Positive and critical aspects*. In Conference of European Statisticians - Working Session on Statistical Data Editing, 1–10. UNECE.URL <http://www.unece.org/fileadmin/DAM/stats/documents/2008/04/sde/wp.14.e.pdf>
- Ceccarelli C., Cutillo A., Di Laurea D., 2009. *Metodologie per la stima degli affitti figurativi ed impatto sulla distribuzione del reddito*. Rivista di statistica ufficiale. n. 1. Istat. Roma. http://www3.istat.it/dati/catalogo/20100913_00/1_2009.pdf
- Ceccarelli C., Pezone A., Rosati S., 2013. *L'utilizzo delle liste anagrafiche comunali nella statistica ufficiale*. Rivista di Economia Demografia Statistica, Vol. LXVII n. 3-4 – 2013 (in corso di pubblicazione).
- Crescenzi F., Ceccarelli C., 1999. *A proposal for Italian Household Budget Survey temporal weighting*, 52nd Session of the International Statistical Institute, Helsinki, August 1999, Tome LVIII – Book 1.
<https://www.stat.fi/isi99/proceedings/arkisto/varasto/cres0815.pdf>
- Crossley T., Winter J., 2013. *Asking Households About Expenditures: What Have We Learned?* Draft paper. <http://www.nber.org/chapters/c12666.pdf>
- D'Aurizio L., Grassi D., Macchia S., Manzari A., Pannuzi N., 1999. *Acquisizione e trattamento dei dati nella nuova indagine sui consumi*. Roma: Istat (Documenti Istat. n. 21).
- Deville, J.C. and Särndal, C.E., 1992. *Calibration estimators in survey sampling*. Journal of the American Statistical Association, 87 (418): 376–382.

- Di Torrice M., Dattilo B., Morricone M., Sabato M., 2014. *L'integrazione della rilevazione viaggi e vacanze nell'indagine sulle spese delle famiglie* in Freguja, Romano (a cura di) *La modernizzazione delle tecniche di rilevazione nelle indagini socioeconomiche sulle famiglie*, Istat, Roma. <http://www.istat.it/it/archivio/145721>
- Donatiello G., D'Orazio M., Frattarola D., Rizzi A., Scanu M., Spaziani M., 2014. *Statistical Matching of Income and Consumption Expenditures*. International Journal of Economic Sciences, Vol. III(3), pp. 50 – 65.7
- Eurostat, 1993. *Family Budget Surveys in the EC: Methodology and Recommendations for Harmonisation - Luxembourg 1993*.
- Eurostat, 1997. *Household budget surveys in the EU - Methodology and recommendations for harmonization 1997*. Office for Official Publications of the European Communities.
- Eurostat, 2003. *Household Budget Surveys in the EU. Methodology and recommendations for harmonisation – 2003*. Methods and nomenclatures. Luxembourg: Eurostat
http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-BF-03-003-__-N-EN.pdf
- Eurostat, 2009. *Improving data comparability for the next HBS round (2010)*. Working party living conditions. Luxembourg: Eurostat (10-12 June 2009) http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/Annexes/hbs_esms_an5.pdf
- Eurostat, 2010. *The comparability of imputed rent*. Methodologies and working papers. Luxembourg.
http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-10-022/EN/KS-RA-10-022-EN.PDF
- Eurostat, 2011. *Implementing core variables methodological guidelines*. Methodological guidelines. Luxembourg: Directorate F: Social Statistics and Information Society, Eurostat
http://epp.eurostat.ec.europa.eu/portal/page/portal/information_society/documents/Tab/CORE%20VARIABLES%20UPDATED%20GUIDELINES%20May%202011.pdf
- Eurostat, 2012. *Description of the data transmission for HBS (Reference Year) 2010 Version: Final*. Methodological guidelines. Luxembourg: Directorate F: Social Statistics, Unit F-4: Quality of Life, Eurostat
http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/Annexes/hbs_esms_an5.pdf
- Eurostat, 2013, *COICOP Five-Digit - Structure and Explanatory Notes*, documento della Unit B5, del 6 dicembre 2013.
- Ferrante M.R., 1996. *Dealing with nonignorable attrition a review*. Statistica Applicata-Italian Journal of Applied Statistics, Vol. 8, n. 4
<http://sa-ijas.stat.unipd.it/sites/sa-ijas.stat.unipd.it/files/661-669.pdf>
- Filippucci C., Ferrante M.R., 1997. *Individual diaries and expense documents in the Italian Consumer Expenditure Survey*. Journal of Official Statistics, Vol. 13, No. 2, pp 113-121. Statistic Sweden.
<http://www.jos.nu/Articles/abstract.asp?article=132113>
- Freguja C., Romano M. C., 2014, (a cura di) *La modernizzazione delle tecniche di rilevazione nelle indagini socioeconomiche sulle famiglie*, Istat, Roma. <http://www.istat.it/it/archivio/145721>
- Garofalo G., 2014. *Il Progetto ARCHIMEDE obiettivi e risultati sperimentali*. Istat Working Papers, N. 9/2014, Istat. Roma. <http://www.istat.it/it/files/2014/11/IWP-n.-9-2014.pdf>
- Goldenberg K., Ryan J., 2009. *Evolution and Change in the Consumer Expenditure Surveys: Adapting Methodologies to Meet Changing Needs*. National Bureau of Economic Research, Summer Institute. Conference on Research on Income and Wealth Cambridge, Massachusetts July 13, 2009
<http://www.bls.gov/cex/nber2009ryan1.pdf>
- INE, 2007, *Encuesta de Presupuestos Familiares*. Madrid
- Istat, 2000. *I consumi delle famiglie*. Roma: Istat (Collana Annuari, n. 4).

- Istat, 2004a. *Metodologie di stima degli aggregati di contabilità nazionale a prezzi correnti*. Collana Metodi e norme, n. 21. Roma.
http://www3.istat.it/dati/catalogo/20041202_00/
- Istat, 2004b. *Concord V. 1.0 Controllo e correzione dei dati. Manuale utente e aspetti metodologici*. Roma: Istat (Tecniche e strumenti n. 1).
- Istat, 2006a. *Il sistema di indagini sociali multiscopo*. Collana Metodi e norme, n. 31. Roma.
http://www3.istat.it/dati/catalogo/20060615_00/
- Istat, 2006b. *La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione*, collana Metodi e norme, n. 32. Roma.
http://www3.istat.it/dati/catalogo/20060830_00/met_norme_06_32_%20rilevazione_forze_lavoro.pdf
- Istat, 2007. *Indicatori della dinamica dei prezzi al consumo per alcune tipologie di famiglie. Anni 2001-2006*. Roma: Istat (Approfondimenti, 20 febbraio 2007).
http://www3.istat.it/salastampa/comunicati/non_calendario/20070220_00
- Istat, 2009. *La misura della povertà assoluta*. Roma: Istat (Collana Metodi e norme, n. 39).
http://www3.istat.it/dati/catalogo/20090422_00/
- Istat, 2011. *I consumi delle famiglie. Anno 2009*. Roma: Istat (Collana Annuari, n.16). http://www3.istat.it/dati/catalogo/20110405_00/
- Istat, 2013. *I consumi delle famiglie. Anno 2012*. Roma: Istat (Statistica report, 8 luglio 2013).
<http://www.istat.it/it/archivio/127996>
- Istat, 2014a. *La misura dell'inflazione per classi di spesa delle famiglie*. Roma: Istat (Statistiche focus, 23 gennaio 2014) <http://www.istat.it/it/archivio/146246>
- Istat, 2014b. *Reddito e condizioni di vita*. Roma: Istat (Statistica report, 30 ottobre 2013).
<http://www.istat.it/it/archivio/136932>
- Lynn P., 2009. *Methodology of Longitudinal Surveys*. Wiley series in survey methodology. Wiley and sons Ltd, Chichester. UK.
- Kott, P.S., 2006. *Using calibration weighting to adjust for nonresponse and coverage errors*. Survey Methodology, 32 (2): 133–142.
- Kreuter F., Eckman S., Jaeckle A., Kirchner A., Presser S., Tourangeau R., 2012. *Mechanisms of misreporting to filter questions*. Paper presented at the 67th AAOPR Annual Conference, Orlando, FL.
- Madow W. G., 1949, *On the Theory of Systematic Sampling, II*, Annals of Mathematical Statistics, 20, 333–354.
- Muircheartaigh C., 1996. *Measurement Errors in Panel Surveys II: some considerations*. Statistica Applicata-Italian Journal of Applied Statistics, Vol. 8, n. 4 <http://sa-ijas.stat.unipd.it/sites/sa-ijas.stat.unipd.it/files/633-645.pdf>
- Murthy M. N., 1967, *Sampling Theory and Methods*, Calcutta: Statistical Publishing Society.
- National Research Council, 2013. *Measuring What We Spend: Toward a New Consumer Expenditure Survey*. Panel on Redesigning the BLS Consumer Expenditure Surveys, Dillman D.A., House C.C. Editors. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academy Press.
- Parker J.A., Souleles N. S., Carroll C. D., 2014. *The Benefits of Panel Data in Consumer Expenditure Surveys*. Paper presented at the NBER CRIW Conference on Consumption Measurement, December 2011. <http://www.nber.org/chapters/c12674.pdf>
- Pezone A., Pellizzaro F., 2012: *The use of population registers in the 15th Italian census: challenges and preliminary evidence*. Joint UNECE/Eurostat Meeting on Population and Housing Censuses Expert Group Meeting on Censuses Using Registers, Geneva, Switzerland, 22-23 May 2012

- Pezone A., Scavalli M., 2010. *Using Population Registers as a Population List: Some Results from the Pilot Survey 2009*. Joint UNECE/Eurostat Expert Group Meeting on Register-Based Censuses, The Hague, Netherlands, 10-11 May 2010.
- Rinaldelli C., 2004. *Il riporto temporale alla spesa mensile nell'indagine sui Consumi delle famiglie: studio ed applicazione di differenti metodologie ed esperienze internazionali*. Documento interno. Istat. Roma.
- Rinaldelli C., Marrone P., 2005. *Prime riflessioni ed ipotesi sull'introduzione di una componente panel nell'indagine sui consumi delle famiglie*. Documento interno. Istat. Roma.
- Rinaldelli C., 2005. *Studio della varianza campionaria per le stime dell'indagine sui consumi in diversi schemi di rotazione*. Documento interno. Istat. Roma.
- Raughunathan T.E., Lepkowski J.M., Van Hoewyk J., Solenberger P., 2001. *A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Regression Models*. Survey Methodology, June 2001, vol. 27, No.1, pp85-95.
- Särndal, C.E., Lundström S., (2005). *Estimation in Survey with nonresponse*. New York: John Wiley & Sons Ltd.
- Silberstein A.R., Scott S., 1991. *Expenditure diary surveys and their associated errors* in P. P. Biemer, R. M. Groves, L. E. Lyberg, *Measurement errors in surveys*, Wiley & Sons, Inc., Hoboken, New Jersey, Capitolo 16 pagg. 303-326.
- Silberstein, A.R., Scott S., 2004. *Expenditure Diary Surveys and Their Associated Errors, in Measurement Errors in Surveys* (eds P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz and S. Sudman), John Wiley & Sons, Inc., Hoboken, NJ, USA.
- Silberstein, A.R., Jacobs C.A., 1989. *Symptoms of Repeated Interview Effects in the Consumer Expenditure Interview Survey*. in Kasprzyk D., Duncan G., Singh M.P. *Panel Surveys*, John Wiley and Sons. New York, pp 319-339.
- SINU, 2014. *LARN - Livelli di Assunzione di Riferimento di Nutrienti ed energia per la popolazione italiana. IV Revisione*. Coordinamento editoriale SINU-INRAN. Milano: SICS.
- United Nations, 2000. *Classifications of Expenditure According to Purpose*; New York. <http://www.stat.gov.tw/public/Attachment/9112616223571.pdf>
- U.S. Bureau of the Census, 1965. *Response Errors in Collection of Expenditures Data by Household Interviews: An Experimental Study*. Technical Paper No. 11. U.S. Government Printing Office, Washington, D.C. <https://play.google.com/books/reader?id=Ns1Ckc6aQ-IC&printsec=frontcover&output=reader&hl=it&pg=GBS.PA6>
- Zhang M., Von Sanden N., 2008. *Using Administrative Data for Seasonal Adjustment of Survey - Time Series in the Presence of a Major Measurement Change*. International Association of Official Statistics, Shanghai.