

BANCA D'ITALIA

Temi di discussione

del Servizio Studi

Mancate interviste e distorsione degli stimatori

di Luigi Cannari e Giovanni D'Alessio



Numero 172 - Giugno 1992

BANCA D'ITALIA

Temi di discussione

del Servizio Studi

Mancate interviste e distorsione degli stimatori

di Luigi Cannari e Giovanni D'Alessio

Numero 172 - Giugno 1992

La serie «Temi di discussione» intende promuovere la circolazione, in versione provvisoria, di lavori prodotti all'interno della Banca d'Italia o presentati da economisti esterni nel corso di seminari presso l'Istituto, al fine di suscitare commenti critici e suggerimenti.

I lavori pubblicati nella serie riflettono esclusivamente le opinioni degli autori e non impegnano la responsabilità dell'Istituto.

COMITATO DI REDAZIONE: *GIORGIO GOMEL, CURZIO GIANNINI, LUIGI GUISO, DANIELE TERLIZZESE;
RITA CAMPOREALE (segretaria).*

Mancate interviste e distorsione degli stimatori

di Luigi Cannari e Giovanni D'Alessio (*)

SOMMARIO

Le mancate interviste costituiscono un problema rilevante nelle indagini campionarie; per effetto di tale fenomeno i campioni possono fornire stime distorte degli aggregati oggetto di indagine, sottorappresentando quei gruppi di unità che tendono a sottrarsi alla rilevazione. Nel presente lavoro, utilizzando il panel di famiglie rilevate nell'ambito dell'indagine sui bilanci delle famiglie italiane condotta dalla Banca d'Italia, si verifica, mediante la stima di un modello di probabilità, che la propensione a partecipare all'indagine è notevolmente differenziata secondo le caratteristiche familiari. Nell'ipotesi che il modello stimato sul panel possa essere esteso, con opportune modifiche, alle famiglie contattate per la prima volta in un'indagine, il lavoro quantifica le distorsioni che il comportamento di non risposta determina sulla stima del reddito familiare medio. Applicando uno stimatore che tiene conto della propensione delle famiglie a collaborare alla rilevazione, il reddito medio familiare del 1987 si incrementa di circa il 5 per cento rispetto ai dati pubblicati dalla Banca d'Italia. La stima del reddito risulta tuttavia ancora inferiore a quella tratta dai Conti Nazionali. L'esperimento condotto, pur dimostrando che l'effetto delle mancate interviste non è trascurabile, sembra indicare che altri fattori, tra cui probabilmente il comportamento di reticenza delle famiglie a dichiarare le loro entrate, sono responsabili dei divari tra le stime campionarie e le quelle desunte dai conti nazionali.

INDICE

1 - Introduzione	p. 5
2 - Le motivazioni del comportamento di non risposta....	p. 7
3 - Un modello di probabilità di risposta.....	p. 9
4 - Un esperimento di correzione del campione complessivo.....	p. 11
5 - Conclusioni.....	p. 15
Appendice	p. 20

(*) Banca d'Italia, Servizio Studi

1 - Introduzione¹

Nelle indagini campionarie accade di frequente che le stime di alcuni indicatori rilevati presentino scostamenti significativi rispetto alle informazioni aggregate desumibili da altre fonti. Ad esempio, nell'indagine sui bilanci delle famiglie condotta dalla Banca d'Italia, la stima del reddito ottenuta riproporzionando all'universo i dati campionari risulta inferiore rispetto a quella della Contabilità Nazionale; parimenti, le attività finanziarie sono notevolmente inferiori rispetto al valore aggregato calcolato nell'ambito della contabilità finanziaria²; analogamente, la stima del numero di abitazioni possedute dalle famiglie risulta sensibilmente diversa dai dati di Censimento³.

Tale fenomeno sembra comune a molti paesi⁴ e viene spiegato con diverse motivazioni. In primo luogo si sostiene che le famiglie appartenenti a particolari segmenti (ad esempio le famiglie più ricche) tendono a sottrarsi alle rilevazioni campionarie; pertanto, anche i campioni casuali presentano distorsioni connesse con la probabilità di

1. Gli autori desiderano ringraziare un anonimo referee per gli utili suggerimenti forniti. Il lavoro, pur essendo il risultato della collaborazione dei due autori, è stato curato da Luigi Cannari per quanto riguarda i paragrafi 2, 4 e 5 e da Giovanni D'Alessio per quanto riguarda i paragrafi 1 e 3 e l'Appendice.

2. Al riguardo si veda Cannari, D'Alessio, Raimondi, Rinaldi (1990).

3. Si veda Cannari, D'Alessio (1990).

4. La dimensione degli scostamenti tra i dati campionari e le statistiche aggregate (Conti Nazionali, Contabilità finanziaria) è tuttavia largamente variabile. Si veda Cannari (1991).

collaborare delle famiglie (selectivity bias)⁵.

Secondariamente, nella rilevazione di certi fenomeni, le famiglie intervistate possono fornire informazioni sistematicamente distorte, non dichiarando il possesso o il consumo di certi beni, o dichiarando valori notevolmente inferiori rispetto a quelli effettivi (under-reporting).

Il confronto tra le stime campionarie riportate all'universo e le informazioni aggregate desumibili da altre fonti consente di ottenere una misura complessiva della distorsione, ma non permette di identificare il peso relativo delle due principali determinanti degli scostamenti.

L'elevata quota di mancate interviste⁶ potrebbe indurre a ritenere poco rilevante il peso dei fattori connessi con il comportamento di under-reporting.

D'altra parte, il confronto delle caratteristiche rilevate per i non rispondenti con quelle relative all'insieme delle famiglie che hanno partecipato all'indagine non sembra segnalare la presenza di importanti fattori di distorsione⁷.

5. Si veda Madow, Olkin (1983), Madow, Olkin e Rubin (1983), Madow, Nisselson e Olkin (1983) e Maddala (1983). Per applicazioni ai dati italiani si vedano Bracalente (1991), Cannari e Violi (1991) e Proietti (1991).

6. Nell'indagine sul 1987 il 40,3 per cento delle famiglie incluse nel campione teorico non è stato rilevato; tale percentuale è incrementata tra il 1987 e il 1989, risultando pari al 63,0 per cento nell'indagine più recente. Al fine di ottenere un campione che rispetti i vincoli imposti a priori sulle numerosità, complessive o nei singoli strati, nella fase di rilevazione le famiglie incluse nel campione, ma non partecipanti all'indagine, vengono sostituite con altre famiglie residenti nello stesso comune, estratte da elenchi di riserva appositamente predisposti.

7. Tali confronti presentano numerosi limiti, a causa delle difficoltà operative connesse con il reperimento di informazioni sui non rispondenti. Infatti, inviare nuovamente intervistatori presso un campione di non rispondenti appare impraticabile, soprattutto quando il motivo della mancata intervista deriva dal rifiuto della famiglia a collaborare. D'altra parte, anche gli sforzi volti a reperire alcune

In questo lavoro, utilizzando i dati panel rilevati nell'indagine sui bilanci delle famiglie si cerca di far luce su questi aspetti, quantificando la distorsione determinata dal comportamento di non risposta delle famiglie.

I risultati ottenuti sembrano indicare che tale distorsione, pur non essendo trascurabile, spiega solo una parte modesta della differenza tra le stime campionarie e i dati di Contabilità Nazionale.

2 - Le motivazioni del comportamento di non risposta

Un'analisi più efficace dei fattori che determinano il comportamento di non risposta può essere condotta disponendo di un panel di famiglie. In tale condizione, per le famiglie che non partecipano ad una rilevazione si dispone infatti di informazioni riguardanti i periodi precedenti.

Una rilevazione di tipo panel è stata recentemente condotta dalla Banca d'Italia. In particolare, nell'indagine sul 1989 una parte delle famiglie già intervistate sul 1987 è stata nuovamente oggetto di rilevazione.

Il tasso di partecipazione di queste famiglie all'indagine è risultato pari al 23,4 per cento delle

(Continuazione nota 7 dalla pagina precedente)
semplici informazioni sulla struttura delle famiglie non rispondenti non hanno apportato rilevanti contributi alla comprensione del fenomeno. Nell'esperienza specifica delle indagini condotte dalla Banca d'Italia, agli intervistatori è stato richiesto di raccogliere alcuni dati sulle famiglie non rispondenti del campione 1989. In particolare si è prevista la rilevazione diretta delle caratteristiche dell'immobile, l'invito alle famiglie che rifiutano l'intervista completa a fornire alcune informazioni di struttura (sesso, età e condizione professionale del capofamiglia e numero di componenti la famiglia) o il ricorso a vicini per il reperimento di tali dati in caso di assenza della famiglia da intervistare. Questa operazione non ha tuttavia prodotto risultati soddisfacenti. Le principali variabili sono risultate disponibili in un numero limitato di casi e spesso caratterizzate da problemi di qualità.

famiglie contattate⁸.

Il motivo principale della mancata reintervista risulta essere il timore nel fornire informazioni, seguito dal rifiuto a collaborare e dalla irreperibilità; tali motivi riguardano rispettivamente il 26,7, il 23,2 e il 19,8 per cento delle famiglie panel contattate⁹.

Il peso di tali motivazioni è tuttavia notevolmente diversificato al variare delle caratteristiche socio-demografiche delle famiglie contattate (tavola 1).

L'importanza dell'irreperibilità quale motivo di non risposta risulta nettamente decrescente al crescere dell'età del capofamiglia e del numero di componenti e al decrescere della dimensione del comune di residenza della famiglia. Ciò potrebbe derivare dal fatto che le famiglie con capofamiglia in età avanzata tendono a trascorrere nella loro abitazione una maggiore quota del loro tempo. A una maggiore dimensione familiare potrebbe invece essere associata una maggiore probabilità di reperire un componente della famiglia alla

8. Tale tasso risulta sensibilmente inferiore rispetto a quello delle famiglie che, nell'ambito della stessa indagine, sono state contattate per la prima volta (41,2 per cento).

9. Una famiglia è definita "non rispondente per irreperibilità" se l'intervistatore, in due visite all'indirizzo di residenza, non è in grado di contattare membri della famiglia al corrente delle scelte effettuate dal responsabile (o dai responsabili) della gestione economica della famiglia. Tra le famiglie irreperibili sono inoltre compresi i casi in cui il nominativo da intervistare non risiede all'indirizzo specificato. E' definita "non rispondente per timore" una famiglia che dichiara di non voler collaborare alla rilevazione sostenendo o lasciando intendere all'intervistatore che il motivo è connesso con il timore di comportamenti criminosi. E' definita "non rispondente per indisponibilità" una famiglia che decide di non collaborare alla rilevazione adducendo motivi diversi dal timore e non oggettivamente constatabili, almeno in linea teorica, dall'intervistatore (ad esempio mancanza di tempo o disinteresse per la rilevazione). Infine è definita "non rispondente per altri motivi" una famiglia che non collabora alla rilevazione adducendo motivi oggettivamente constatabili dall'intervistatore (ad esempio decesso o malattia della persona responsabile dell'economia familiare).

data dell'intervista.

La mancata partecipazione dovuta al timore nel fornire informazioni all'intervistatore risulta invece crescente con l'età del capofamiglia e con il reddito familiare; il peso di tale fattore risulta particolarmente elevato in corrispondenza di redditi superiori a 50 milioni (in lire 1987) e nei comuni metropolitani.

Quanto al rifiuto a collaborare per motivi differenti dal timore, i risultati non evidenziano comportamenti di semplice interpretazione. Il peso di tale fattore risulta crescente con la dimensione della famiglia (in termini di componenti e percettori di reddito) e con l'ampiezza del comune di residenza; con riferimento al reddito familiare si osserva un andamento prima crescente (fino alla classe da 40 a 50 milioni) e poi decrescente.

Quanto agli "altri motivi", questi costituiscono una categoria residuale di ancor più complessa interpretazione. Il peso elevato di tale fattore per le famiglie con capofamiglia anziano (24,9 per cento nella classe oltre 70 anni) potrebbe essere in parte associato con lo stato di salute dei componenti la famiglia o con la difficoltà a comprendere integralmente i quesiti posti con il questionario. L'andamento prima decrescente e successivamente crescente per classi di reddito appare invece connesso con il rifiuto a collaborare per indisponibilità; la somma delle due motivazioni sembra infatti essere relativamente costante per classi di reddito e segnala la mancanza di una chiara linea di demarcazione tra le due categorie.

3 - Un modello di probabilità di risposta

L'analisi appena condotta evidenzia che la mancata partecipazione all'indagine sul 1989 delle famiglie rilevate sul 1987 è connessa con le caratteristiche socio-demografiche delle famiglie stesse; per tali variabili e per quelle ad

esse correlate le statistiche basate sul panel effettivamente intervistato sarebbero distorte¹⁰.

Al fine di individuare le determinanti della propensione a rispondere, si è quindi stimato sull'insieme delle famiglie contattate ai fini della reintervista (circa 5.000) un modello di probabilità in cui la collaborazione all'indagine e le motivazioni di non risposta sono funzione delle caratteristiche socio-demografiche delle famiglie stesse.

Il modello ipotizzato è il seguente¹¹:

$$P_{ir} = \frac{e^{x_i \beta_r}}{1 + \sum_{k=1}^{R-1} e^{x_i \beta_k}} \quad r=1, \dots, R \quad (1)$$

dove P_{ir} è la probabilità dell' i -esima famiglia di non collaborare all'indagine per il motivo r ($r=1, \dots, R-1$) e P_{iR} è la probabilità di rispondere, x_i è un vettore di variabili socio-demografiche e β_r ($r=1, \dots, R-1$) sono i vettori dei parametri da stimare¹².

Come variabili esplicative sono state utilizzate tutte quelle che risultavano correlate con il tasso di risposta, sulla base dell'analisi descrittiva presentata nel precedente paragrafo.

10. Esaminando il reddito familiare rilevato con l'indagine sul 1987, si osserva che la media ottenuta sul panel rientrato è inferiore di circa il 5 per cento rispetto al corrispondente valore calcolato sulle famiglie contattate per la reintervista. Nell'effettuare il confronto il panel rientrato è stato corretto mediante una procedura di stratificazione a posteriori, allineando la distribuzione delle famiglie per strato che risulta dai dati panel con la corrispondente distribuzione della popolazione.

11. Per una descrizione delle caratteristiche e delle proprietà dei modelli di questo tipo si veda Amemiya (1983).

12. A causa del vincolo $P_{i1} + P_{i2} + \dots + P_{iR} = 1$ i vettori da stimare sono $R-1$. Nel caso in esame, senza perdita di generalità, si è posto $\beta_R = 0$.

I risultati della stima, effettuata con il metodo della massima verosimiglianza, sono riportati in appendice.

Le dummies di ampiezza comunale risultano significative e indicano che la probabilità di risposta è tanto maggiore quanto minore è la dimensione del comune di residenza della famiglia. Tale probabilità, inoltre, risulta maggiore per i capofamiglia di sesso maschile, crescente all'aumentare del numero dei componenti (figura 1) e decrescente in funzione del reddito familiare (figura 2). Per quanto riguarda l'età, l'andamento è crescente fino a 65 anni e poi decrescente (figura 3). Il livello di istruzione del capofamiglia non risulta invece significativo.

4 - Un esperimento di correzione del campione complessivo

La relazione stimata nel precedente paragrafo evidenzia come le variabili socio-demografiche influenzano il comportamento di risposta delle famiglie comprese in un panel; estendendo il risultato a più occasioni di indagine si può supporre che successive rilevazioni comportino una selezione delle famiglie che allontana sempre più la composizione del campione risultante da quello teorico. Parimenti, è presumibile che un analogo processo di selezione agisca anche sulle famiglie che per la prima volta vengono contattate ai fini di una indagine statistica.

Nell'ipotesi che le relazioni stimate sul panel possano essere estese, con opportune modifiche, al campione di famiglie contattate per la prima volta, è possibile rimuovere le distorsioni derivanti dal comportamento di non risposta.

A tal fine, si assume in primo luogo che nella popolazione di riferimento non vi siano unità statistiche caratterizzate da una probabilità nulla di rispondere (hard core nonrespondents). Secondariamente, si ipotizza che il rapporto tra la probabilità di non rispondere per l'*r*-esimo motivo e la probabilità di rispondere si modifichi al variare

della rilevazione allo stesso modo per tutte le famiglie. In particolare:

$$\begin{array}{ll}
 P'_{iR} \neq 0 & \text{per ogni } i \\
 P_{iR} \neq 0 & \text{per ogni } i \\
 (P'_{ir}/P'_{iR}) = f_r (P_{ir}/P_{iR}) & \text{per ogni } i \text{ e } r
 \end{array} \quad (2)$$

dove P'_{iR} e P_{iR} rappresentano le probabilità di risposta della famiglia i -esima rispettivamente al primo contatto e al momento della reintervista; P'_{ir} e P_{ir} sono le probabilità di non rispondere per il motivo r , rispettivamente al primo contatto e al momento della reintervista; i parametri da stimare f_r tengono conto della variazione, tra una rilevazione e l'altra, del peso relativo dei fattori di non risposta¹³.

La prima ipotesi è necessaria in quanto se le famiglie appartenenti a una certa tipologia non collaborano ad alcuna indagine statistica non è possibile ottenere stime corrette a partire dal campione osservato.

Quanto alla seconda ipotesi, va segnalato che da una rilevazione ad un'altra l'importanza relativa dei fattori che influenzano il comportamento di non risposta si modificano; ad esempio, nell'indagine sul 1989, nel passaggio dalla prima rilevazione alla successiva, il peso del fattore "indisponibilità" aumenta mentre quello del fattore "timore" diminuisce¹⁴. L'ipotesi utilizzata in questo paragrafo consente di cogliere tali variazioni di comportamento introducendo alcuni effetti fissi, connessi con la rilevazione ma non con le caratteristiche familiari. Il modello utilizzato per le famiglie contattate per la prima

13. Nel seguito del lavoro la seconda ipotesi viene resa meno restrittiva permettendo ai parametri f_r di variare per area di residenza della famiglia.

14. Si veda Banca d'Italia (1991, p.21).

volta, risulta pertanto il seguente:

$$P'_{ir} = \frac{f_r e^{x_i \beta_r}}{1 + \sum_{k=1}^{R-1} f_k e^{x_i \beta_k}} \quad r=1, \dots, R \quad (3)$$

Come indicato nel precedente paragrafo, i parametri β_r sono stati stimati applicando il modello (1) ai dati relativi alle famiglie intervistate sul 1987 e contattate per la reintervista sul 1989. I parametri f_r sono stati calcolati sulla base della relazione¹⁵:

$$f_r = Q_r / (Q_R K_r) \quad r=1, \dots, R-1 \quad (4)$$

dove Q_R rappresenta la quota di famiglie contattate per la prima volta che sono state effettivamente intervistate, Q_r la quota di famiglie che non risponde per la motivazione r e K_r rappresenta il valore medio della variabile $\exp(x_i \beta_r)$ calcolato per il complesso delle famiglie intervistate per la prima volta¹⁶.

Ottenuta la stima della probabilità che una famiglia selezionata per il campione teorico partecipi effettivamente all'indagine, uno stimatore del valore medio della variabile Y , non affetto da distorsione dovuta a fenomeni di

15. I valori f_r sono calcolati a livello di strato (3 aree geografiche per 4 classi di ampiezza demografica dei comuni).

16. Il valore medio $E[\exp(x_i \beta_r) | i \in I]$, dove I rappresenta l'insieme delle famiglie intervistate per la prima volta, equivale infatti a $E[\exp(x_i \beta_r) P'_{iR} | i \in C] / E[P'_{iR} | i \in C]$, dove C rappresenta l'insieme delle famiglie contattate per la prima volta. Sostituendo a P'_{iR} l'espressione (3) si ottiene la relazione (4).

self-selection, può essere definito nel modo seguente¹⁷:

$$Y^+ = \sum_{i=1}^N w_i Y_i \quad (5a)$$

$$w_i = \frac{1}{P_i^T P'_{iR}} \quad (5b)$$

dove i pesi w_i rappresentano l'inverso della probabilità di effettiva partecipazione all'indagine che deriva dal prodotto della probabilità teorica di inclusione (P_i^T) per la probabilità di collaborare (P'_{iR}).

Applicando lo stimatore (5a), il reddito medio delle famiglie nel 1987 risulta pari a 31,1 milioni; tale stima è superiore del 5,4 per cento rispetto a quella ottenuta con lo stimatore utilizzato dalla Banca d'Italia sull'intero campione di famiglie intervistate sul 1987. La rivalutazione appare più consistente se la stima ottenuta mediante il modello viene confrontata con quella calcolata sul sottoinsieme di famiglie che ha partecipato sia all'indagine sul 1987 che a quella sul 1989¹⁸. In questo caso il doppio

17. Per le proprietà degli stimatori basati su osservazioni campionarie con differenti probabilità di selezione si rinvia a Hansen, Hurwitz e Madow (1953, p. 59-65), Cochran (1977, p. 259-261) e, per recenti applicazioni, a Ekholm e Laaksonen (1991, p. 325-337).

18. E' da tenere presente che i valori indicati come semplice stima sono calcolati utilizzando pesi che tengono conto del tasso di partecipazione a livello di strato, essendo questa la procedura usualmente utilizzata nell'indagine della Banca d'Italia. Qualora non venisse effettuata l'operazione di stratificazione a posteriori che tiene conto dei suddetti tassi di partecipazione, la correzione basata sul modello sarebbe più consistente. Ciò deriva dal fatto che le variabili geografiche hanno un rilevante impatto sulla probabilità di risposta e sono correlate con il livello del reddito. Pertanto, correggendo per la struttura della popolazione per area geografica e ampiezza demografica dei comuni, si elimina parte della distorsione derivante dai fenomeni di non risposta.

processo di selezione che avviene tra il contatto e la prima intervista e tra la prima e la seconda intervista, conduce a un reddito medio pari a 28,3 milioni, di circa il 10 per cento inferiore a quello stimato mediante il modello.

Nonostante la correzione, tuttavia, il reddito familiare medio stimato è ancora inferiore rispetto al corrispondente valore di Contabilità Nazionale, risultando pari all'89,4 per cento di quest'ultimo.

5 - Conclusioni

Sulla base dell'analisi condotta nel presente lavoro possono trarsi le seguenti considerazioni.

a) La propensione a collaborare all'indagine dipende dalle caratteristiche socio-demografiche delle famiglie, dall'area territoriale e dall'ampiezza demografica del comune dove esse risiedono e dal reddito familiare. Le famiglie a più alto reddito tendono ad essere sottorappresentate, con la conseguente sottostima del reddito familiare medio.

b) Il metodo di correzione generalmente utilizzato, che tiene conto del tasso di partecipazione a livello di strato, rimuove solo una parte della distorsione determinata dai fenomeni di selezione.

c) Se si ipotizza che il comportamento di non risposta osservato sulle famiglie panel possa essere esteso all'insieme di famiglie contattate per la prima volta, si può definire uno stimatore che, sotto ipotesi non troppo restrittive, non è affetto da distorsioni dovute a fenomeni di self-selection. Utilizzando questo stimatore il reddito medio dell'indagine si incrementa del 5 per cento circa rispetto al valore stimato dalla Banca d'Italia.

d) La discrepanza tra le stime campionarie e i dati di Contabilità Nazionale, pur riducendosi, rimane consistente. Il divario tra le due fonti sarebbe quindi determinato principalmente da altri fattori, tra i quali il comportamento di reticenza delle famiglie a dichiarare le proprie entrate potrebbe assumere un ruolo particolarmente rilevante.

Famiglie rilevate nell'indagine sul 1987 e contattate per la reintervista sul 1989: percentuale di famiglie reintervistate e motivazioni delle mancate interviste(*)

	Famiglie contattate per la reintervista					
	Famiglie non intervistate				Famiglie reintervistate	Totale
	Irreper.	Timore	Indisp.	A.motivi		
Sesso						
Maschio	19,8	27,0	24,0	5,3	23,9	100,0
Femmina	20,2	25,1	18,4	15,8	20,5	100,0
Eta'						
Fino a 30 anni	44,2	16,3	16,8	2,2	20,6	100,0
Da 30 a 40 anni	30,2	22,5	23,3	3,0	21,0	100,0
Da 40 a 50 anni	17,3	24,8	29,2	4,2	24,5	100,0
Da 50 a 60 anni	15,3	28,6	27,8	4,1	24,2	100,0
Da 60 a 70 anni	12,7	34,7	20,5	7,0	25,1	100,0
Oltre 70 anni	10,8	29,0	12,1	24,9	23,2	100,0
Numero di componenti						
1 componente	29,4	17,2	12,0	20,0	21,4	100,0
2 componenti	20,7	30,1	18,9	7,8	22,5	100,0
3 componenti	20,3	28,4	24,4	4,1	22,8	100,0
4 componenti	16,9	27,0	27,8	4,4	23,9	100,0
5 o piu' componenti	15,0	25,8	28,6	4,1	26,5	100,0
Numero di percettori						
1 percettore	18,5	26,5	22,5	8,8	23,6	100,0
2 percettori	22,0	26,3	22,5	5,3	23,9	100,0
3 percettori	19,2	29,5	23,8	5,8	21,6	100,0
4 o piu' percettori	13,4	26,1	38,9	3,2	18,5	100,0
Classe di reddito						
Fino a 20 milioni	18,4	25,4	20,0	10,5	25,7	100,0
Da 20 a 30 milioni	18,5	26,8	22,2	6,6	26,0	100,0
Da 30 a 40 milioni	23,6	25,3	26,4	4,0	20,6	100,0
Da 40 a 50 milioni	23,3	23,6	31,6	2,0	19,6	100,0
Da 50 a 60 milioni	17,2	34,3	25,4	3,7	19,4	100,0
Da 60 a 70 milioni	20,6	28,4	25,5	4,3	21,3	100,0
Oltre 70 milioni	20,6	38,3	22,6	4,4	14,1	100,0
Ampiezza comunale						
Fino a 20.000 abitanti	17,7	24,7	21,9	7,7	28,1	100,0
Da 20 a 40.000 abitanti	19,8	26,4	22,0	5,4	26,4	100,0
Da 40.000 a 1 milioni di ab.	20,1	26,6	23,6	7,3	22,4	100,0
Oltre 1 milione di abitanti	22,9	32,9	25,8	5,6	12,9	100,0
Area geografica						
Nord	17,8	28,6	25,5	6,5	21,5	100,0
Centro	19,8	31,1	22,0	4,0	23,1	100,0
Sud e isole	21,8	22,6	21,4	8,8	25,4	100,0
Totale	19,8	26,7	23,2	6,9	23,4	100,0

(*) Dati non ponderati

Probabilità' stimate di collaborare all'indagine
e di non collaborare secondo la motivazione

Fig.1

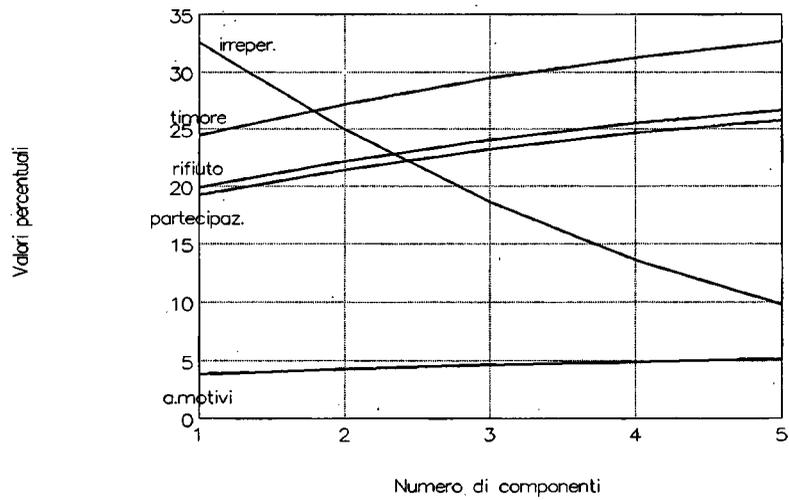


Fig.2

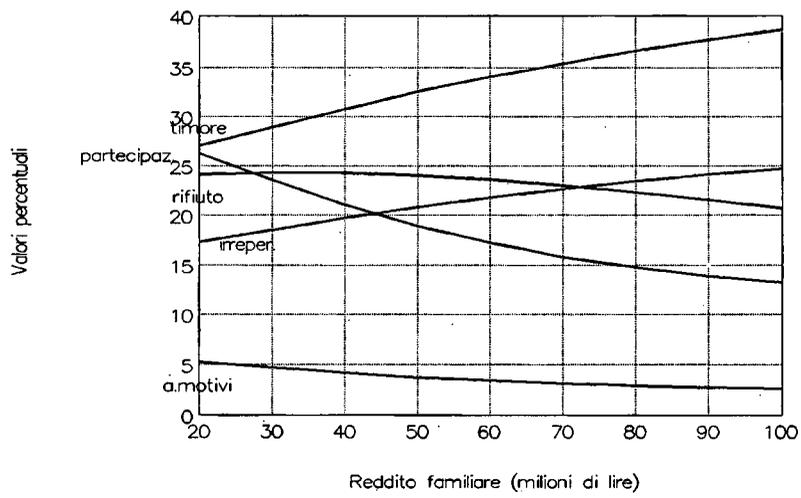
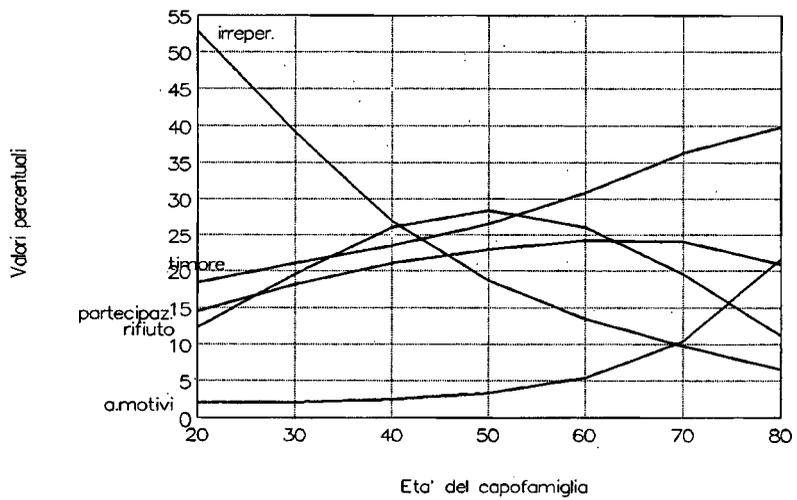


Fig.3



Reddito familiare medio, 1987

	Indagine campionaria (*)			
	Contabilità Nazionale (CN)	Contattati (1)	Intervistati sul 1987 (2)	Intervistati sul 1987 e reintervistati sul 1989 (2)
Milioni di lire	34,8	31,1	29,5	28,3
Indice (CN=100)	100,0	89,4	84,8	81,3

(*) Indagine della Banca d'Italia sui bilanci delle famiglie nel 1987. Escluso il sovracampionamento.

(1) Stima effettuata utilizzando lo stimatore (5a).

(2) Stima ottenuta mediante stratificazione a posteriori basata sui tassi di partecipazione a livello di comune.

APPENDICE

Tav.A.1

Stima della probabilità di risposta
e delle motivazioni di non risposta(*)

Parametri β_1 relativi al fattore "irreperibilità":

intercetta	3.48294660	
	(97.71)	
dac1	0.90639636	comuni oltre 1.000.000 ab.
	(90.02)	
dac2	0.24150103	comuni da 40.000 a 1.000.000 ab.
	(25.05)	
dar1	-0.46535088	area Nord
	(28.99)	
dar2	-0.28624690	area Centro
	(5.70)	
dsex	-0.27009837	maschi
	(14.51)	
eta	-0.07574424	età
	(34.36)	
eta2	0.00033830	età al quadrato
	(7.42)	
ncomp	-0.37260261	numero di componenti
	(90.45)	
y	0.02392585	reddito familiare
	(78.71)	
y2	-0.00008886	reddito familiare al quadrato
	(31.69)	

Tav.A.1 (segue)

Stima della probabilità di risposta
e delle motivazioni di non risposta

Parametri β_2 relativi al fattore "timore":

intercetta	-0.12506300	
	(0.11)	
dac1	0.90639636	comuni oltre 1.000.000 ab.
	(90.02)	
dac2	0.24150103	comuni da 40.000 a 1.000.000 ab.
	(25.05)	
dar1	0.16802230	area Nord
	(6.05)	
dar2	0.29763990	area Centro
	(8.34)	
dsex	-0.27009837	maschi
	(14.51)	
eta	-0.02727887	età
	(4.01)	
eta2	0.00033830	età al quadrato
	(7.42)	
y	0.02392585	reddito familiare
	(78.71)	
y2	-0.00008886	reddito familiare al quadrato
	(31.69)	

Tav.A.1 (segue)

Stima della probabilità di risposta
e delle motivazioni di non risposta

Parametri β_3 relativi al fattore "rifiuto":

intercetta	-1.69340640	
	(9.94)	
dac1	0.90639636	comuni oltre 1.000.000 ab.
	(90.02)	
dac2	0.24150103	comuni da 40.000 a 1.000.000 ab.
	(25.05)	
dar1	0.16802230	area Nord
	(6.05)	
dar2	0.07500846	area Centro
	(0.47)	
dsex	-0.27009837	maschi
	(14.51)	
eta	0.06258114	età
	(8.72)	
eta2	-0.00069713	età al quadrato
	(12.17)	
y	0.01745864	reddito familiare
	(36.23)	
y2	-0.00008886	reddito familiare al quadrato
	(31.69)	

Tav.A.1 (segue)

Stima della probabilità di risposta
e delle motivazioni di non risposta

Parametri β_4 relativi al fattore "altri motivi":

intercetta	-0.44738360	
	(1.28)	
dac1	0.90639636	comuni oltre 1.000.000 ab.
	(90.02)	
dac2	0.24150103	comuni da 40.000 a 1.000.000 ab.
	(25.05)	
dar1	-0.46535088	area Nord
	(28.99)	
dar2	-0.88492910	area Centro
	(20.49)	
dsex	-0.27009837	maschi
	(14.51)	
eta	-0.07574424	età
	(34.36)	
eta2	0.00108853	età al quadrato
	(89.02)	

Stima della probabilità di risposta
e delle motivazioni di non risposta

(*) Dati relativi alle famiglie intervistate sul 1987 e contattate per la reintervista sul 1989; valori delle variabili riferiti al 1987; reddito in milioni di lire; il test chi-quadrato di significatività dei singoli coefficienti è riportato tra parentesi.

Modello:
Probabilità di non risposta per il fattore r: $e^{x_i \beta_r} / (1 + \sum_k e^{x_i \beta_k})$

Probabilità di risposta: $1 / (1 + \sum_k e^{x_i \beta_k})$

Test sulle restrizioni rispetto al modello generale contenente tutte le variabili esplicative riportate nella tavola, più il livello di istruzione del capofamiglia (in anni), più la dummy ampiezza comunale da 20.000 a 40.000 abitanti, senza vincoli sui parametri β_r tra le equazioni:

rapporto di verosimiglianza (L.R.)	= 40.06
gradi di libertà	= 29
valore critico del chi-quadrato (p=0,95)	= 42.56

Significatività del modello (tutti i parametri uguali a zero ad eccezione delle intercette):

chi-quadrato	= 721.05
gradi di libertà	= 19
valore critico del chi-quadrato (p=0,95)	= 30.14

BIBLIOGRAFIA

- AMEMIYA T. (1983), Advanced Econometrics, Oxford, Oxford University Press.
- BANCA D'ITALIA (1986), Le indagini campionarie sui bilanci delle famiglie italiane, Numero speciale dei contributi all'analisi economica, Banca d'Italia, Roma.
- _____ (1991), I bilanci delle famiglie italiane nell'anno 1989, Supplementi al Bollettino Statistico, Note metodologiche e informazioni statistiche, Nuova Serie, Anno I, numero 26.
- BRACALENTE, B. (1991), I non rispondenti nelle indagini sul reddito: lo studio di un caso, documento presentato al convegno "Personal Income Distribution, Inequality and Poverty", Siena, 28-30 settembre, 1991.
- CANNARI, L. (1991), Response Errors in The Bank of Italy's Survey of Household Income and Wealth, Banca d'Italia, dattiloscritto.
- _____ - D'ALESSIO G. (1990), Housing Assets In The Bank Of Italy's Survey Of Household Income And Wealth, in Income And Wealth Distribution, Inequality And Poverty, edited by DAGUM C. e ZENGA M., Studies in Contemporary Economics, Berlino, Springer-Verlag.
- _____ - RAIMONDI, G. - RINALDI, A.I. (1990), Le attività finanziarie delle famiglie italiane, Temi di discussione, n. 136, Banca d'Italia, Roma.
- _____ - VIOLI, R. (1991), Under-reporting del reddito nell'indagine sui bilanci delle famiglie condotta dalla Banca d'Italia, documento presentato al convegno "Personal Income Distribution, Inequality and Poverty", Siena, 28-30 settembre, 1991.
- COCHRAN W.G. (1977), Sampling Techniques, New York, John Wiley & Sons.
- EKHOLM A. - LAAKSONEN S. (1991), "Weighting via Response Modeling in the Finnish Household Budget Survey", Journal of Official Statistics, Vol. 7 n. 3, 1991, pp. 325-337.
- HANSEN M.H. - HURWITZ W.N. - MADOW W.G. (1953), Sample Surveys, Methods and Theory, New York, John Wiley & Sons.

MADDALA, G.S. (1983), Limited Dependent and Qualitative Variables in Econometrics, Cambridge, Cambridge University Press.

MADOW, W.G. - OLKIN, I. (eds) (1983), Incomplete Data in Sample Surveys, Proceedings of the Symposium, New York, Academic Press.

_____ - RUBIN, D.B. (eds) (1983), Incomplete Data in Sample Surveys, Theory and Bibliographies, New York, Academic Press.

_____ - NISSELSON, H. - OLKIN, I. (eds) (1983), Incomplete Data in Sample Surveys, Report and Case Studies, New York, Academic Press.

PROIETTI, T. (1991), Tecniche di imputazione e indagini sul reddito, documento presentato al convegno "Personal Income Distribution, Inequality and Poverty", Siena, 28-30 settembre, 1991.

ELENCO DEI PIÙ RECENTI TEMI DI DISCUSSIONE (*)

- n. 146 — *La scelta del meccanismo di collocamento dei titoli di Stato: analisi teorica e valutazione dell'esperienza italiana*, di L. BUTTIGLIONE - A. PRATI (gennaio 1991).
- n. 147 — *Diversification and Performance*, di M. BIANCO (gennaio 1991).
- n. 148 — *Exchange Rate and Pricing Strategies in a Model of International Duopoly*, di P. CASELLI (gennaio 1991).
- n. 149 — *Concorrenza e redditività nell'industria bancaria: un confronto internazionale*, di V. CONTI (febbraio 1991).
- n. 150 — *Economie di scala e di diversificazione nel sistema bancario italiano*, di C. CONIGLIANI - R. DE BONIS - G. MOTTA - G. PARIGI (febbraio 1991).
- n. 151 — *Politiche di offerta e riallocazione del credito bancario negli anni ottanta*, di C. GIANNINI - L. PAPI - A. PRATI (febbraio 1991).
- n. 152 — *Stime regionali con pochi dati: analisi e simulazioni di stimatori alternativi per investimenti, occupazione e fatturato delle imprese manifatturiere*, di R. CESARI - L. F. SIGNORINI (marzo 1991).
- n. 153 — *Dinamica retributiva e differenziali salariali*, di A. GAVOSTO - P. SESTITO (luglio 1991).
- n. 154 — *Interessi reali, sistema impositivo ed effetto Sylos Labini*, di P. VAGLIASINDI (luglio 1991).
- n. 155 — *Trasformazione delle scadenze e margine d'interesse degli istituti di credito mobiliare*, di P. SABBATINI (luglio 1991).
- n. 156 — *Gli effetti della quotazione internazionale: il caso delle azioni italiane a Londra*, di F. PANETTA (agosto 1991).
- n. 157 — *Grandi e piccole imprese negli anni ottanta: la ristrutturazione dell'industria in un'analisi di dati di bilancio*, di L. F. SIGNORINI (agosto 1991).
- n. 158 — *Demand and Supply Shocks in Industrial Output*, di A. GAVOSTO - G. PELLEGRINI (novembre 1991).
- n. 159 — *I futures e le opzioni sui titoli di Stato. Un'analisi del mercato e delle prospettive in Italia*, di A. SCALIA - L. TORNETTA (novembre 1991).
- n. 160 — *Earnings Uncertainty and Precautionary Saving*, di L. GUIISO - T. JAPPELLI - D. TERLIZZESE (febbraio 1992).
- n. 161 — *Migrazioni in Europa: andamenti, prospettive, indicazioni di politica economica*, di G. GOMEL - S. REBECCHINI (febbraio 1992).
- n. 162 — *Monetary Aggregates and Monetary Policy Coordination on the Way to Economic and Monetary Union: the Role of Cross-Border Deposits*, di P. GIUCCA - A. LEVY (febbraio 1992).
- n. 163 — *Cross-Border Deposits and Monetary Aggregates in the Transition to EMU*, di I. ANGELONI - C. COTTARELLI - A. LEVY (marzo 1992).
- n. 164 — *Young Households' Saving and the Life Cycle of Opportunities. Evidence from Japan and Italy*, di A. ANDO - L. GUIISO - D. TERLIZZESE (marzo 1992).
- n. 165 — *Bequests and Saving for Retirement. What Impels the Accumulation of Wealth?*, di F. BARCA - L. CANNARI - L. GUIISO (marzo 1992).
- n. 166 — *The Microeconomics and Macroeconomics of the Permanent Income Hypothesis*, di A. DEATON (marzo 1992).
- n. 167 — *Why is Italy's Saving Rate so High?*, di L. GUIISO - T. JAPPELLI - D. TERLIZZESE (aprile 1992).
- n. 168 — *Waiting for EMU: Living with Monetary Policy Asymmetries in the EMS*, di L. BINI SMAGHI (aprile 1992).
- n. 169 — *Income and Saving in Italy: a Reconstruction*, di G. MAROTTA - P. PAGLIANO - N. ROSSI (giugno 1992).
- n. 170 — *Finance and Development: The Case of Southern Italy*, di R. FAINI - G. GALLI - C. GIANNINI (giugno 1992).
- n. 171 — *Generational Accounting: The Case of Italy*, di D. FRANCO - J. GOKHALE - L. GUIISO - L. J. KOTLIKOFF - N. SARTOR (giugno 1992).

(*) I «Temi» possono essere richiesti a:

