

Italian Survey on Consumer Expectations

Appendice metodologica

Introduzione

La *Italian Survey of Consumer Expectations* (ISCE) è un'indagine trimestrale realizzata insieme alla **Nota mensile** e alla presente **Appendice metodologica** nell'ambito delle attività dello Spoke 3 del partenariato esteso **Grins (Growing, Resilient, Inclusive and Sustainable)**.

L'indagine, svolta tramite metodologia CAWI¹, è rivolta alla popolazione di età compresa tra 18 e 75 anni residente in Italia ed è suddivisa in sei sezioni contenenti informazioni su variabili demografiche, reddito, ricchezza, consumi e aspettative su variabili individuali (consumo, reddito) e variabili macroeconomiche (inflazione, tassi di interesse nominali, crescita del PIL). L'ultima sezione è dedicata a focus tematici differenti per ogni rilevazione.

La ISCE ha tre obiettivi:

- produrre stime su aspettative e comportamenti individuali a frequenza trimestrale;
- produrre analisi di politiche per la sostenibilità economica delle famiglie;
- valutare metodi e affidabilità di aspettative soggettive.

L'indagine si basa sulle esperienze internazionali di indagini online ad alta frequenza. In particolare, la *Survey of Consumer Expectations* (SCE) della **Federal Reserve di New York** raccoglie informazioni sulle opinioni e sulle aspettative dei consumatori in materia di inflazione, occupazione, reddito e bilanci delle famiglie. La *Consumer Expectation Survey* (CES) della **Banca Centrale Europea** raccoglie mensilmente dati sulle aspettative delle famiglie su circa 10.000 famiglie delle sei maggiori economie dell'area dell'euro (Germania, Francia, Italia, Spagna, Paesi Bassi e Belgio). Anche diverse altre esperienze internazionali sono un utile riferimento, come il **Social Economic Lab di Harvard**, che analizza le determinanti delle preferenze, degli atteggiamenti e delle percezioni attraverso indagini sociali.

¹ Computer Assisted Web Interviewing

1. Il campione

Il campione è tratto da un più ampio campione rappresentativo di 120.000 individui mantenuto e aggiornato regolarmente da Doxa. L'indagine si rivolge alla popolazione di età compresa tra i 18 e i 75 anni residente in Italia, utilizzando il metodo CAWI.

Il campione è stratificato per sesso, età (18-34, 35-44, 45-54, 55-64, 65-over) e area geografica (Nord-Ovest, Nord-Est, Centro e Sud e Isole).

L'esercizio pilota è stato somministrato nel mese di settembre 2023 (100 interviste).

2. Survey Design

Il questionario è stato costruito con l'aiuto di esperti del settore e ricercatori accademici.

2.1 Il panel proprietario

L'agenzia Doxa dispone di una piattaforma web progettata e sviluppata per rispondere a specifiche esigenze di ricerca. Questo panel proprietario conta oltre 120.000 panelisti registrati, di cui 50.000 attivi (hanno completato un sondaggio nei 12 mesi precedenti o si sono iscritti nel 2021). Il tasso di risposta medio è stato del 40% e gli inviti a rispondere al sondaggio sono stati inviati agli utenti, in media, 2,5 volte al mese. I sondaggi sono ottimizzati per i diversi dispositivi (circa il 33% viene effettuato tramite cellulari).

2.2 Reclutamento dei partecipanti al panel

L'agenzia di sondaggi effettua un reclutamento periodico degli iscritti (2-3 volte l'anno) per ampliare la base di riferimento dei ricercatori online e garantire la rotazione degli iscritti (ogni anno viene sostituito il 10%-20% degli utenti attivi del panel). Particolare attenzione viene posta alla rappresentatività del panel, sia in termini sociodemografici che comportamentali. Il reclutamento prende in considerazione una serie di fonti e metodologie di reclutamento, con diverse strategie per ridurre le distorsioni nel processo di reclutamento del panel.

2.3 Profilazione del panel

I profili del panel si basano su diverse variabili socioeconomiche, tra cui:

- caratteristiche personali (sesso, età, regione, provincia, comuni, numero di abitanti, istruzione, occupazione);
- professione (tipo di professione, settore professionale, ruolo, dipartimento);
- reddito (reddito individuale e familiare e guadagni);
- ricchezza finanziaria (beni immobili, attività finanziarie, tipo di investimenti, debiti, assicurazioni);

- consumi (totale, cibo, energia, gas, salute).

2.4 Incentivi

Gli iscritti ai panel web ricevono incentivi per partecipare attivamente alla ricerca proposta. L'agenzia di sondaggi presta molta attenzione al tipo di sistema di incentivi, poiché potrebbe influenzare la decisione di aderire al panel e determinare problemi di autoselezione, o condizionare la risposta alle domande compromettendo il risultato del sondaggio. Per filtrare i partecipanti interessati solo all'incentivo, al pagamento dell'incentivo personale è associata una donazione a un ente di beneficenza senza scopo di lucro.

2.5 Fieldwork management

Durante la fase di lavoro sul campo, l'agenzia di sondaggi segue procedure rigorose per limitare i pregiudizi introdotti da chi risponde velocemente o da chi va veloce. I questionari vengono somministrati ai partecipanti in modo casuale e l'invito rimane valido per almeno una settimana (compreso il fine settimana). L'obiettivo è quello di consentire la partecipazione al sondaggio di persone che non guardano la posta elettronica ogni giorno e garantire la partecipazione di persone che tendono a non rispondere immediatamente.

2.6 Esercizio pilota

Nelle prime due settimane di settembre 2023 è stato condotto un sondaggio pilota per valutare l'efficacia del questionario proposto e identificare potenziali problemi. L'indagine pilota si è rivolta a un campione ristretto di 100 intervistati. I risultati preliminari hanno indicato un alto livello di coinvolgimento e comprensione da parte dei rispondenti. Tuttavia, sono stati individuati alcuni errori di *routing* che hanno reso necessaria una leggera revisione. Questi problemi sono stati prontamente identificati e risolti, assicurando un'elaborazione dei dati che rifletta accuratamente le informazioni fornite dagli intervistati.

3. Il questionario

Il questionario è composto da due parti. Una parte comune e stabile (circa 10-12 minuti), che sarà ripetuta in ogni ondata, e sezioni speciali di circa 5-6 minuti che possono cambiare a ogni indagine.

- A. Variabili demografiche e occupazione
- B. Reddito
- C. Ricchezza
- D. Consumi
- E. Aspettative
- F. Sezione speciale.

Il contenuto di ciascuna sezione è descritto brevemente di seguito.

A. Variabili demografiche e occupazione

La sezione A è ricca di informazioni sulle caratteristiche demografiche degli intervistati: sesso, città di residenza, livello di istruzione, stato civile, dimensione della famiglia, percettori di reddito. Per quanto riguarda l'istruzione, la ISCE raccoglie dati sul tipo e sulla specializzazione della laurea. Inoltre, viene richiesto lo stato occupazionale, distinguendo tra lavoratori dipendenti e autonomi, pensionati o in cerca di occupazione. Per coloro che sono occupati, la sezione esplora ulteriormente il settore in cui sono impiegati, fornendo un'istantanea completa della forza lavoro. Per consentire confronti, la codifica delle variabili è il più possibile simile a quella dell'indagine sul reddito e la ricchezza delle famiglie (SHIW) della Banca d'Italia.

B. Reddito

Le variabili di reddito si riferiscono al reddito mensile nel mese precedente alla rilevazione. Il reddito è raccolto in 11 fasce, oltre che con una domanda qualitativa che indica se il reddito è molto al di sotto, al di sotto, circa uguale, al di sopra, molto al di sopra della media delle famiglie italiane. Per creare statistiche descrittive si può utilizzare il punto medio degli intervalli scelti dall'intervistato. Quando gli intervalli non sono delimitati, è stato utilizzato un limite superiore ragionevole per stimare i momenti delle distribuzioni.

L'indagine si concentra sulle seguenti variabili di reddito: reddito disponibile delle famiglie, reddito da lavoro e da pensione delle famiglie, reddito totale individuale, reddito da lavoro e da pensione individuale. Ciascuna di queste variabili è raccolta al "netto delle imposte", come nella SHIW.

C. Ricchezza

La sezione C costruisce un indicatore della ricchezza totale e della partecipazione ai mercati finanziari con riferimento al mese precedente alla rilevazione. Gli intervistati riportano in 6 fasce la ricchezza finanziaria, la ricchezza reale e il debito totale, consentendo di stimare la ricchezza netta. Inoltre, indicano se sono proprietari della casa in cui vivono. Per quanto riguarda la ricchezza finanziaria, gli intervistati indicano se hanno un conto corrente e investimenti finanziari in obbligazioni, azioni, pensioni private e assicurazioni sulla vita. Le ultime domande riguardano le assicurazioni, separatamente per le assicurazioni sanitarie e per quelle contro gli infortuni.

D. Consumi

La sezione D raccoglie i consumi mensili e le categorie di consumo con riferimento al mese precedente alla rilevazione. Agli intervistati è chiesto di riportare i consumi totali (11 fasce), le bollette di gas ed elettricità (6 fasce), le spese sanitarie (6 fasce).

E. Aspettative

La sezione E è dedicata alle aspettative e alle intenzioni. Essa svolge un ruolo fondamentale nella previsione del *sentiment* economico.

La strategia comune di questa sezione consiste nel rilevare non solo la media delle variabili future (generalmente a 12 mesi di distanza), ma anche l'intera distribuzione, chiedendo agli intervistati di assegnare 100 punti a determinati intervalli delle aspettative. Ad esempio, gli intervistati indicano la probabilità (in percentuale) che il loro reddito diminuisca o aumenti entro determinati intervalli (ad esempio, diminuzione superiore all'8%, aumento tra il 2 e il 4%, ecc.). Viene poi ricavata la funzione di densità di probabilità soggettiva per ciascun intervistato e, sulla base di questa, calcolata la distribuzione delle aspettative. Si rimanda al paragrafo 5 per maggiori dettagli.

La sezione richiede le seguenti distribuzioni per la crescita prevista nei prossimi 12 mesi del reddito e delle spese individuali: reddito disponibile, reddito da lavoro e da pensione, crescita dei consumi totali, spese sanitarie, bollette del gas e dell'elettricità, prezzo della casa in cui si vive.

La sezione E chiede informazioni riguardo all'intenzione degli intervistati di acquistare specifici beni durevoli (automobili, elettrodomestici, mobili, elettronica), di chiedere un prestito nei prossimi 12 mesi e la probabilità (su una scala da 1 a 100) che il prestito venga concesso e le aspettative individuali sul tasso di interesse nominale che gli intervistati ritengono di poter investire in modo sicuro la loro ricchezza finale nei prossimi 2 mesi.

La sezione richiede, sempre in formato probabilistico, la distribuzione dell'età pensionabile attesa e del tasso di sostituzione, nonché la probabilità che specifici eventi abbiano conseguenze finanziarie per la famiglia nei prossimi 12 mesi (disoccupazione, spese sanitarie superiori a 10.000 euro, disabilità).

Utilizzando lo stesso approccio, la parte finale della sezione E chiede di fornire previsioni per i prossimi 12 mesi di quattro variabili macroeconomiche chiave: crescita del PIL, inflazione, disoccupazione e tasso di interesse nominale sui mutui. Queste aspettative possono essere facilmente confrontate con le previsioni attuali delle variabili aggregate fornite dal governo, dalle banche centrali, da altre agenzie nazionali internazionali e da altri sondaggi.

F. Sezione speciale

L'ultima sezione è dedicata a tematiche differenti per ogni rilevazione e darà la possibilità di introdurre degli "esperimenti", in cui il campione complessivo è diviso in 4/5 sottocampioni casuali per consentire disegni di trattamento e controllo.

4. Statistiche descrittive

L'Appendice Statistica contiene le statistiche descrittive dettagliate per gruppi sociodemografici (età, sesso, dimensione della famiglia, stato occupazionale, regione) per le variabili illustrate nella Nota mensile.

4.1 Missing values

Per alcune variabili, come reddito e ricchezza, è prevista la possibilità, per l'intervistato, di rispondere di non sapere o di non voler indicare la risposta. Questi casi vengono trattati utilizzando un metodo di imputazione delle risposte basato su alcune variabili sociodemografiche in grado di spiegare potenzialmente queste variabili.

Le principali variabili sociodemografiche utilizzate sono lo stato di occupazione (lavoratore dipendente, autonomo o disoccupato), la regione in cui vivono, l'età, la situazione previdenziale (in pensione o meno), il livello di istruzione. All'intervistato viene, inoltre, fornita l'informazione sul reddito mensile medio (netto) delle famiglie italiane e gli viene chiesto di indicare il suo posizionamento rispetto a tale livello.

4.2 Dati ambientali

I sondaggi online raccolti tramite CAWI comportano vantaggi e svantaggi. L'aspetto negativo è che le risposte potrebbero essere meno accurate rispetto alle interviste di persona, soprattutto quando si tratta di rispondere a domande complesse. L'aspetto positivo è che gli intervistati sono molto numerosi in tutto il Paese, provenendo da quasi 2000 comuni diversi, ed è quindi possibile unire i dati con indicatori di rischio ambientale georeferenziati. Ciò può consentire di analizzare, ad esempio, la relazione tra i rischi ambientali e alcune variabili oggetto dell'indagine.

A tal fine, ci si avvale della Piattaforma Dati GeoSafe, un software proprietario di ANIA per l'analisi dei rischi da calamità naturali sul territorio italiano secondo diversi livelli di granularità, fino alle coordinate civiche o geografiche. GeoSafe geolocalizza i rischi di catastrofi naturali sulla mappa dell'Italia, per misurare e quantificare la pericolosità e la vulnerabilità a cui un'area può essere esposta.

Lo strumento è certificato da partner accademici e istituzionali e si compone di quattro moduli: (i) rischio idraulico e idrogeologico; (ii) rischio sismico; (iii) rischio climatico.

Il modulo finalizzato alla stima del rischio idraulico (alluvioni ed esondazioni) e idrogeologico (frane) definisce una classifica di pericolosità sulla base di quattro variabili che possono essere selezionate separatamente o contestualmente, quali: (i) vicinanza a tratti fluviali; (ii) acclività dei versanti o presenza di frane; (iii) alluvioni storiche e sinistri avvenuti negli ultimi 30 anni; (iv) altezza del pescaggio dell'acqua.

Il modulo finalizzato alla stima del rischio sismico fornisce un indicatore rappresentativo della pericolosità sismica (realizzato dall'INGV) e consente di inserire i dati relativi alla tipologia di edificio, all'altezza dell'edificio e all'anno di costruzione per uno specifico edificio.

Per il rischio climatico, infine, GeoSafe fornisce un indicatore di rischio sviluppato su uno specifico dataset meteorologico (composto da precipitazioni, temperatura, neve, potenziale grandine, velocità del vento, fulmini) e collegato alle condizioni di pericolosità ed esposizione degli asset oggetto di stima.

5. Appendice

5.1. Costruzione delle aspettative e dell'indicatore di incertezza

Le distribuzioni delle aspettative riportate nella Nota e le relative tavole statistiche contenute nell'Appendice Statistica sono state costruite secondo in metodo delle *probabilistic expectations* (per distinguerle dalle stime di aspettative puntuali).

Il modo più semplice per costruire tali aspettative è quello di considerare il centro di ciascun intervallo e poi pesarlo per le probabilità assegnate da ciascun rispondente. Per l'intervallo non limitato viene imposto un *upper bound* più ampio di quelli predefiniti (in letteratura è stato dimostrato che le persone che selezionano l'ultimo intervallo hanno di solito aspettative molto estreme).

Oltre alla media di ciascuna distribuzione, è possibile costruire la deviazione standard della distribuzione di probabilità soggettiva di ciascun individuo delle variabili future. Tale valore può essere considerato come un indicatore di incertezza nella stima delle variabili future dei singoli individui. Per un dato valore medio, una deviazione standard maggiore corrisponde ad un grado incertezza maggiore. Ai fini della comparabilità tra i valori calcolati, si evidenzia che il confronto è possibile per le variabili le cui domande sono definite sullo stesso supporto.

Nell'Appendice statistica tale indicatore verrà indicato come descritto in Tabella 1 per distinguerlo dalla deviazione standard delle singole variabili che rappresenta, invece, il tipico indicatore di dispersione dei valori di una variabile in un campione di dati.

Tabella 1. Descrizione delle variabili

Variabili derivate	
Expected disposable income growth	EY
Standard deviation of disposable income growth	SDY
Expected labor income growth	EYL
Standard deviation of labor income growth	SDYL

Expected consumption growth	EC
Standard deviation of consumption growth	SDC
Expected health expenditures	ECHEALTH
Standard deviation of health expenditures	SDCHEALTH
Expected bill expenditures	ECBILL
Standard deviation of bill expenditures	SDCBILL
Expected house price growth	EHOME
Standard deviation of house price growth	SDHOME
Expected GDP growth	EGDP
Standard deviation of GDP growth	SDGDP
Expected inflation	EINFL
Standard deviation of inflation	SDINF
Expected unemployment rate	EU
Standard deviation of unemployment rate	SDU
Expected nominal interest rate	ER
Standard deviation of nominal interest rate	SDR
Expected nominal interest rate on mortgages	ERM
Standard deviation of nominal interest rate on mortgages	SDRM
Variabili descrittive	
Monthly household's income	Y
Monthly household's labor income	YL
Monthly individual income	PY
Monthly individual labor income	PYL
Household's real assets	AR
Household's financial wealth	AF
Household's debts	PF
Household's total wealth (AR + AF - PF)	W

Household's total consumption	CTOT
Household's total consumption of food	CFOOD
Household's total consumption of energy	CENER
Household's total consumption of gas	CGAS
Household's total health expenditure	CHEALTH
Tabelle	
tableB.xlsx	Reddito
tableC.xlsx	Ricchezza
tableD.xlsx	Consumi
tableE.xlsx	Aspettative