

Metodi e fonti: note metodologiche

23 febbraio 2017

Per informazioni: statistiche@bancaitalia.it
www.bancaditalia.it/statistiche/index.html

Sondaggio congiunturale sul mercato delle abitazioni in Italia

Aspetti generali

Il Sondaggio congiunturale sul mercato delle abitazioni in Italia è stato avviato nel gennaio 2009 da Banca d'Italia e Tecnoborsa e prosegue, a partire dall'edizione di ottobre 2010, anche con la cooperazione dell'Osservatorio del Mercato Immobiliare (OMI) dell'Agenzia delle Entrate¹. Dall'edizione di aprile 2016 le interviste sono condotte dal gruppo CLAS S.p.A.

L'obiettivo del sondaggio è quello di ottenere le valutazioni degli agenti immobiliari sull'andamento del mercato immobiliare sia per quanto attiene le compravendite di immobili residenziali, sia per quanto riguarda il mercato degli affitti.

La maggioranza delle informazioni raccolte sono di tipo qualitativo e volte a rilevare le opinioni degli agenti immobiliari sull'evoluzione congiunturale del mercato immobiliare nel trimestre di riferimento e in prospettiva. Le domande qualitative presenti nel questionario prevedono di solito tre o cinque modalità di risposta. La modalità centrale rappresenta un giudizio di neutralità (o stabilità) e le modalità a destra e sinistra rappresentano giudizi favorevoli e sfavorevoli rispetto al fenomeno indagato. Ad esempio, i giudizi potranno essere "migliore", "uguale", "peggiore", oppure "in aumento", "stabile", "in diminuzione", oppure "superiore", "uguale", "inferiore". I fenomeni osservati sono quantificati attraverso i saldi, calcolati come differenza fra le frequenze della modalità favorevole e sfavorevole. La modalità centrale (stabilità) non viene considerata nel calcolo.

Il sondaggio, unico nel panorama italiano per diffusione territoriale e per tipologia di informazione raccolta, viene svolto presso un panel di circa 1.300 - 1.400 agenti immobiliari che vengono intervistati trimestralmente.

L'andamento dei principali indicatori risultanti dal sondaggio è riportato nelle tavole statistiche allegate al rapporto trimestrale.

Nel seguito sono descritti nel dettaglio i vari aspetti metodologici che caratterizzano il sondaggio.

La composizione dell'universo di riferimento e del campione²

L'universo di riferimento è costituito dalle agenzie che operano in conto terzi (circa 32.000 agenzie). Le informazioni sulla distribuzione della popolazione, utilizzate per la costruzione dei pesi, sono di fonte Istat, fornite per classe di addetti e provincia (Codice Ateco 68 - Attività immobiliari).

¹ In attuazione dell'art. 23-quater del decreto legge n. 95/2012, dal 1° dicembre 2012 l'Agenzia delle Entrate ha incorporato l'Agenzia del Territorio e, conseguentemente, l'Osservatorio del Mercato Immobiliare che di questa faceva parte.

² Alcune elaborazioni sui dati delle passate rilevazioni potrebbero non coincidere con i dati pubblicati nelle edizioni precedenti del presente rapporto a causa di revisioni degli archivi nel corso del tempo.

Alcune delle liste utilizzabili per la costruzione del campione sono relative agli agenti immobiliari, altre alle agenzie, e sono le seguenti:

- a) agenti associati alla Federazione Italiana Agenti Immobiliari Professionali (FIAIP);
- b) agenti accreditati presso le Borse immobiliari;
- c) agenti corrispondenti dell'Osservatorio del Mercato Immobiliare (OMI);
- d) lista di imprese operanti nella mediazione immobiliare su beni di terzi, estratte dagli elenchi delle Camere di commercio.

A causa delle difficoltà riscontrate nell'utilizzo³ della lista camerale (in linea di principio esaustiva) l'estrazione è operata principalmente sulle liste di cui alle lettere da a) a c) e integrata con la lista camerale allo scopo di colmare eventuali carenze rispetto alla numerosità teorica e per attenuare le possibili distorsioni causate dall'utilizzo di liste relative ad associazioni di categoria.

Nelle tavole i dati vengono presentati per ripartizioni geografiche, e per aree urbane e non urbane. La frazione sondata è in genere più alta al Nord est, nel Sud e Isole e nelle aree urbane e metropolitane.

Il disegno di campionamento è stratificato. Gli strati, in totale pari a 34, sono costituiti:

- a) dai 12 comuni italiani con popolazione pari o superiore a 250.000 abitanti⁴ e dai comuni di Padova, Trieste e Messina (nel complesso sono i 15 maggiori comuni italiani per popolazione residente);
- b) dalle 15 aree intorno ai comuni del punto a), che ne formano i relativi hinterland⁵;
- c) dalle 4 macro-aree geografiche nazionali (Nord Ovest; Nord Est; Centro; Sud e Isole), con l'esclusione dei 30 strati dei punti a) e b).

Il numero di unità da rilevare in ogni strato è stato scelto con un criterio misto: una numerosità di base è stata determinata in misura proporzionale al numero di transazioni rilevate, disponibili al momento della formazione del campione, in ognuno di essi (sulla base di dati territoriali forniti dall'OMI).

All'interno dei singoli strati si è poi provveduto a fissare una consistenza minima, al fine di garantire una numerosità campionaria tale da assicurare per ognuno un errore standard accettabile per le principali stime. È stata infine aggiunta una numerosità campionaria supplementare per le aree metropolitane, per tener conto della maggiore variabilità, al loro interno, dei fenomeni di interesse dell'indagine.

³ Le informazioni di contatto delle liste camerali risultano infatti spesso incomplete o non aggiornate.

⁴ Bari, Bologna, Catania, Firenze, Genova, Milano, Napoli, Palermo, Roma (inclusa Ostia Lido), Torino, Venezia (inclusa Mestre), Verona.

⁵ Definiamo come "area urbana" l'unione del comune con almeno 250.000 abitanti e del relativo hinterland. Chiamiamo "area metropolitana" l'area urbana in cui il comune più importante conta almeno 500.000 abitanti (Genova, Milano, Napoli, Palermo, Roma, Torino). Ogni area urbana o metropolitana corrisponde a un particolare sistema locale del lavoro (Istat, I sistemi locali del lavoro 1991, Roma, 1997) definito come "l'unità territoriale identificata da un insieme di comuni contigui legati fra loro dai flussi degli spostamenti quotidiani per motivi di lavoro"; si tratta pertanto di aggregazioni di comuni che identificano mercati del lavoro omogenei. Non rilevando i vincoli amministrativi, un sistema locale può contenere, in casi residuali, comuni appartenenti a province o regioni diverse da quella del comune principale e viene individuato a partire da informazioni acquisite sul fenomeno del pendolarismo tramite il questionario del Censimento generale della popolazione.

La fase di rilevazione, il questionario, i comportamenti di risposta

Le interviste vengono condotte trimestralmente nei mesi di gennaio, aprile, luglio ed ottobre da una società di rilevazione denominata CLAS S.p.A.⁶ che svolge il sondaggio per conto della Banca d'Italia.

La rilevazione dei dati avviene tramite tecniche miste. Circa il 65 per cento delle interviste viene effettuata con la tecnica CATI (Computer Assisted Telephone Interview), mentre la restante parte viene effettuata con la tecnica CAWI (Computer Assisted Web Interview) basata su un questionario compilabile via web.

Il questionario è composto da 4 sezioni: compravendite, prezzi, prospettive, mercato delle locazioni. L'indagine raccoglie informazioni su aspetti del mercato immobiliare italiano quali le vendite delle abitazioni nel trimestre di riferimento e in termini prospettici, la struttura del mercato in cui opera l'agenzia (numero di potenziali acquirenti e venditori), i prezzi medi delle case, indicatori riguardanti il mercato degli affitti.

Occasionalmente (di solito nell'indagine relativa al mese di gennaio) viene introdotta una sezione relativa alle caratteristiche strutturali delle abitazioni vendute quali: superficie, classe energetica, condizioni e tipologia dell'immobile.

Nel corso degli anni i tassi di risposta, inizialmente pari a circa il 45 per cento, hanno subito un calo dovuto ad una molteplicità di fattori (impoverimento delle liste, crisi economica, etc.); attualmente il tasso di risposta si è assestato attorno al 30 per cento.

Il controllo di qualità dei dati e l'imputazione dei dati mancanti

I dati raccolti sono sottoposti ad un controllo di qualità. Esso mira a verificare la compatibilità dei valori inseriti nel questionario con quelli ammessi dalla domanda, la coerenza intertemporale dei dati panel, la presenza di valori anomali e la verifica dei formati numerici richiesti per l'inserimento.

Un primo controllo dei dati inseriti nel questionario viene effettuato dalla società di rilevazione tramite gli intervistatori e tramite controlli effettuati direttamente in fase di inserimento delle risposte nel dataset. Un secondo controllo viene effettuato tramite la procedura informatica di caricamento dei dati che impedisce l'acquisizione di dati esterni all'intervallo di definizione delle variabili rilevate.

Il processo di imputazione si avvale di estimatori per medie, attribuendo a ciascuna osservazione il valore medio stimato nello strato a cui appartiene l'unità da imputare. Attualmente la sola variabile imputata è il numero di case vendute. Il dato mancante viene sostituito moltiplicando il numero di agenti operanti nell'agenzia per la media di strato del numero di case vendute per agente.

La ponderazione

La procedura di ponderazione adottata si compone di due fasi. In una prima fase si usano come strati quelli originari del disegno di campionamento. Ad ogni impresa si attribuisce un peso iniziale, dato dal rapporto tra numero di imprese contenute nella cella di strato e numero di imprese del campione. Sia h la generica cella di strato e al suo interno N_h il numero di imprese

⁶ L'indagine è stata condotta fino a dicembre 2015 dalla società di rilevazione Questlab srl.

della popolazione di interesse e n_h quello del campione⁷, il peso della prima fase per ogni agenzia dello strato h è pertanto:

$$w_h^{(0)} = \frac{N_h}{n_h}$$

Per costruzione, la somma dei pesi di ogni cella restituisce pertanto la numerosità della popolazione di interesse in essa contenuta.

In una seconda fase si effettua una stratificazione a posteriori, utilizzando la tecnica del *raking*⁸, per tenere conto anche della dimensione delle agenzie immobiliari. Il peso iniziale è modificato con un fattore di aggiustamento f_k , consentendo di ottenere i pesi finali:

$$w_{hk}^{(1)} = w_h^{(0)} f_k$$

la cui somma coincide, nella generica cella di post-stratificazione k , con il numero di imprese della popolazione di interesse ivi contenute. Si aggregano le agenzie in 3 classi dimensionali: agenzie con un solo addetto, con 2-5 addetti e con 6 addetti e oltre, separatamente per aree metropolitane, aree urbane non metropolitane e altri comuni, per un totale di 9 celle di post-stratificazione. In corrispondenza di ogni indagine trimestrale tali pesi vengono ricalcolati sulla base della distribuzione della popolazione di interesse disponibile alla data più recente⁹.

Gli stimatori utilizzati e gli errori standard

Per una generica variabile quantitativa X_i , la stima della media è effettuata con uno stimatore dato da:

$$\bar{X} = \frac{1}{\sum w_i} \sum w_i X_i$$

dove w_i è il peso associato alla singola unità del campione. Il peso può essere sia quello originario di disegno sia quello ottenuto moltiplicando quest'ultimo per una variabile di scala che consente di tener conto della diversa dimensione delle agenzie del campione¹⁰.

Per variabili qualitative gli standard error vengono calcolati facendo riferimento alla distribuzione binomiale. Le stime degli errori standard delle percentuali di agenzie per il totale del campione e per alcune disaggregazioni territoriali sono riportate nella tavola A2.

⁷ Il simbolo n_h indica la numerosità campionaria effettiva. In questo modo si effettua implicitamente una correzione dei pesi per tenere conto delle mancate risposte totali (unit non-response).

⁸ La tecnica denominata iterative proportional fitting (o raking) consente di allineare in modo simultaneo i pesi campionari alla distribuzione di alcune caratteristiche note da fonti esterne. Cfr., per esempio, Verma V., Advanced sampling methods: manual for statistical trainers, Tokyo, Statistical Institute for Asia and the Pacific, 2000.

⁹ La popolazione di agenzie utilizzata è stata ottenuta da Istat, Archivio statistico delle imprese attive. Gli aggiornamenti sono normalmente rilasciati dall'Istat con riferimento all'universo di circa due anni prima. Le stime sono riviste periodicamente per tenere conto dei cambiamenti nella popolazione di riferimento dell'indagine.

¹⁰ Per tutti questi aspetti, incluse le tecniche di calcolo degli errori standard si veda ad esempio, Cicchitelli, F., A. Herzl e G.E. Montanari, Il campionamento statistico, Bologna, Il Mulino, 1994.

Il sistema BIRD per l'elaborazione a distanza dei dati dell'indagine

Il sistema per l'elaborazione a distanza dei dati raccolti nelle indagini sulle imprese condotte dalla Banca, denominato BIRD (*Bank of Italy Remote access to micro Data*)¹¹ rende disponibili anche i microdati del presente sondaggio. Il sistema offre la possibilità di svolgere elaborazioni sui dati raccolti, nel pieno rispetto della riservatezza dei dati individuali. L'utente svolge le sue analisi statistiche ed econometriche senza avere accesso diretto ai singoli microdati, sottomettendo la propria elaborazione via e-mail e ricevendo per la stessa via gli output. Sia l'input sia l'output vengono sottoposti a controlli automatici e manuali sul contenuto e sulla logica delle elaborazioni. L'utilizzo del sistema è subordinato all'accettazione, da parte della Banca d'Italia, della richiesta di rilascio di un'utenza.

¹¹ Per maggiori dettagli si visiti la pagina web <https://www.bancaditalia.it/statistiche/basi-dati/bird/index.html>