

# Caratteristiche della rilevazione e note metodologiche

## 1 - Campo e oggetto della rilevazione

La rilevazione ha cadenza mensile e copertura totale e raccoglie le informazioni sui progetti di fabbricati nuovi, residenziali e non residenziali, o di ampliamenti di fabbricati preesistenti, per i quali sia stato ritirato regolare "Permesso di costruire" presso gli uffici comunali di competenza. Le trasformazioni e le ristrutturazioni di fabbricati già esistenti, che non comportano variazioni di volumi degli stessi, non rientrano nel campo di rilevazione.

L'unità di rilevazione è costituita dalla singola opera, rappresentata da un intero fabbricato nuovo, anche se demolito e interamente ricostruito, o dall'ampliamento di un fabbricato preesistente. Due o più opere, relative allo stesso permesso di costruire, costituiscono due o più unità di rilevazione per le quali vengono compilati altrettanti modelli di rilevazione.

Lo strumento di rilevazione è costituito da un questionario cartaceo (modello ISTAT/AE), per il quale vi è l'obbligo di compilazione, a cura del richiedente il permesso di costruire, avvalendosi del modello di istruzioni denominato ISTAT/AE/Istr.

Gli uffici comunali hanno il compito di raccogliere mensilmente i questionari, controllare l'esattezza delle informazioni che vi sono riportate, completarli compilando la parte riservata al comune e inviarli mensilmente all'Istat. In caso di assenza di permessi di costruire nel mese di riferimento, il Comune deve inviare una segnalazione di attività edilizia nulla (modello ISTAT/AE/Neg). Se il Comune, nel mese di riferimento, non invia alcun questionario o segnalazione di attività edilizia nulla, è considerato non rispondente.

Il modello di rilevazione per le due tipologie di fabbricati, residenziali e non residenziali, è composto di tre parti: la prima raccoglie le notizie generali sull'opera (tempi previsti per la realizzazione, ubicazione, natura dell'opera, destinazione d'uso, richiedente il permesso di costruire, finanziamento, regime dei suoli, impianto termico, struttura portante); la seconda parte, relativa ai soli fabbricati residenziali, contiene informazioni sui piani, sul volume, sulla superficie, sul numero di abitazioni e la ripartizione delle abitazioni secondo il numero di stanze per abitazione e secondo le classi di superficie utile abitabile; la terza parte comprende le notizie relative ai soli fabbricati non residenziali e indaga sulla dimensione del fabbricato, sulla parte ad uso abitativo, sulla destinazione economica e sulla tipologia dell'opera.

## 2 - Caratteristiche della mancata risposta

Nelle indagini statistiche, sia censuarie sia campionarie, può verificarsi l'impossibilità di ottenere le informazioni da tutte le unità di rilevazione. Tale fenomeno è noto come *mancata risposta* o *incompletezza dei dati*.

Nel caso della Rilevazione dei permessi di costruire, organizzata come un'indagine censuaria di tutti i Comuni italiani e ripetuta nel tempo, la mancata risposta si riferisce a sottoinsiemi di Comuni di numerosità variabile e di composizione differente in ciascun mese.

Il fenomeno della mancata risposta si differenzia, oltre che per il numero di mesi di collaborazione, anche per la dimensione demografica e per la posizione geografica dei Comuni non rispondenti.

Un indicatore della mancata risposta dei Comuni è dato dal grado di collaborazione calcolato, sia a livello nazionale sia a livello di ciascuna ripartizione geografica, mediante la seguente formula:

$$G = 100 \frac{\sum_{i=1}^k n_i p_i}{12 \sum_{i=1}^k p_i}$$

dove  $n_i$  rappresenta il numero di mesi per cui il generico comune  $i$  ha collaborato e  $p_i$  la popolazione. Nel Prospetto 1 sono riportati i valori dell'indicatore ora definito per il totale nazionale e per ciascuna delle 5 ripartizioni geografiche italiane nel 2005.

**Prospetto 1 – Collaborazione dei comuni per ripartizione geografica - Anno 2005 (valori percentuali)**

RIPARTIZIONI GEOGRAFICHE	2005
Nord-Ovest	86,4
Nord-Est	91,9
Centro	77,7
Sud	69,4
Isole	77,1
<b>Italia</b>	<b>80,5</b>

Il grado di collaborazione a livello nazionale è pari a 80,5 per cento. A livello territoriale il valore più alto si rileva nel Nord-Est col 91,9 per cento, mentre il Sud fa registrare il livello più basso col 69,4 per cento.

Per tenere conto della dimensione demografica dei Comuni non rispondenti, l'analisi della mancata risposta è stata anche condotta suddividendo i Comuni italiani in due sottoinsiemi:

- i 160 comuni capoluogo di provincia e i non capoluogo con più di 50.000 abitanti, che comprendono una popolazione di circa 21 milioni di abitanti (36,4 per cento del totale Italia).
- i 7.940 comuni non capoluogo con meno di 50.000 abitanti che, in termini di popolazione, corrispondono a 36,7 milioni di abitanti, pari al 63,6 per cento del totale nazionale;

La mancata risposta si presenta con caratteristiche differenti per i due sottoinsiemi (Prospetto 2).

**Prospetto 2 – Collaborazione dei Comuni: numero e peso in termini di popolazione percentuale – Anno 2005**

MESI DI COLLABORAZIONE	Comuni	Popolazione
Comuni capoluogo e non capoluogo con più di 50.000 abitanti		
0	5	1,4
1	2	0,6
2	1	0,3
3	-	-
4	1	0,2
5	-	-
6	3	0,8
7	-	-
8	2	0,4
9	6	1,6
10	8	2,6
11	19	6,7
12	113	85,3
<b>Totale</b>	<b>160</b>	<b>100,0</b>
Comuni non capoluogo con meno di 50.000 abitanti		
0	2068	18,8
1	179	1,6
2	109	1,0
3	73	0,7
4	75	0,9
5	89	1,0
6	101	1,1
7	123	1,5
8	167	2,6
9	275	4,2
10	526	7,2
11	1127	14,6
12	3028	44,8
<b>Totale</b>	<b>7940</b>	<b>100,0</b>

Nel sottoinsieme dei comuni capoluogo di provincia e non capoluogo con più di 50.000 abitanti, la collaborazione è elevata; dei 160 comuni complessivi, 113 sono presenti per 12 mesi, per una popolazione pari all'85,3 per cento di quella totale del sottoinsieme.

Per quanto riguarda i 7940 comuni, 3028 sono quelli rispondenti 12 mesi, per una popolazione pari al 44,8 per cento, mentre 2068 sono quelli che non hanno mai collaborato, corrispondenti al 18,8 per cento

della popolazione totale. E' interessante notare che hanno collaborato per 11 mesi, 1127 comuni per una popolazione pari al 14,6 per cento.

### 3 – Metodo di imputazione dei dati mancanti

I risultati presentati nel precedente paragrafo sembrano indicare che le caratteristiche dei Comuni non rispondenti sono sistematicamente differenti da quelle dei Comuni rispondenti. Tale ipotesi è stata sottoposta a una verifica formale mediante la stima di un modello di regressione logistica volto a individuare l'esistenza di fattori che influenzino la probabilità di non risposta. In particolare, è risultato che la popolazione residente nel Comune (misurata al 31 dicembre 1999) e la ripartizione geografica, determinano in misura statisticamente significativa la probabilità di collaborazione all'indagine, confermando l'ipotesi che la mancata risposta non è di tipo *missing completely at random*. Tale risultato implica che l'utilizzo delle sole informazioni rilevate potrebbe condurre a risultati distorti e che è, quindi, necessario applicare un metodo di integrazione dell'informazione tramite imputazione delle mancate risposte.

La metodologia di imputazione dei dati mancanti è distinta per due sottoinsiemi di comuni: quelli capoluogo o non capoluogo con più di 50.000 abitanti e quelli restanti.

Il metodo utilizzato per i comuni del primo sottoinsieme tiene conto dell'importanza che essi assumono in termini di peso nella rilevazione e della loro elevata collaborazione complessiva. L'integrazione dei dati mensili si basa su un'analisi puntuale delle informazioni elementari che conduce a individuare l'insieme di *record* da utilizzare per l'imputazione delle mancate risposte. In sintesi, i dati mensili mancanti sono imputati sulla base di quelli forniti dal medesimo comune per i mesi contigui o, in caso di ulteriori mancate risposte, nel medesimo mese di anni contigui. Come si può evincere dai dati presentati nel Prospetto 2, anche tra i comuni capoluogo (o non capoluogo con oltre 50.000 abitanti) vi è un ristretto numero di casi di assenza totale di risposta in tutti i mesi dell'anno: per ciascuna di tali unità l'imputazione è avvenuta tramite una ricerca ad hoc di un comune donatore con caratteristiche il più possibile simili a quelle del comune per il quale vi era assenza di informazioni utilizzabili.

Nel caso dei comuni non capoluogo con popolazione inferiore a 50.000 abitanti la scelta del metodo di integrazione della mancata risposta è avvenuta mettendo a confronto tre diversi approcci: un metodo di imputazione mediante il donatore, un metodo di imputazione mediante media e un metodo di ponderazione. Per selezionare la metodologia più adatta è stato condotto, sui dati riferiti agli anni 2000 – 2001 - 2002, un esperimento di simulazione che ha utilizzato il sottoinsieme dei 1.910 comuni sempre rispondenti nei tre anni e ha riguardato una delle variabili del modello di rilevazione: il numero di abitazioni in nuovi fabbricati residenziali. I risultati della simulazione, sviluppata utilizzando una stratificazione dei Comuni in base alla ripartizione geografica e alla popolazione, hanno indicato che il migliore metodo di integrazione dei dati mancanti è quello di imputazione mediante donatore. Tale metodo consente di integrare i dati a livello di Comune e, allo stesso tempo, non porta a distorsioni significative nella distribuzione del fenomeno ( Per approfondimenti si veda Bacchini F. Iannaccone R. e Otranto E. "L'imputazione delle mancate risposte in presenza di dati longitudinali: un'applicazione ai permessi di costruzione", Contributi Istat 4/2005).

In generale per i comuni rispondenti in almeno in uno dei mesi del generico anno, all'interno di ciascun strato definito dalle variabili ausiliare scelte, il donatore viene individuato minimizzando, per ciascun comune  $j$ , la seguente funzione di distanza:

$$\min_{1 \leq k \leq r_h} \sum_{m \in M} |x_k^m - x_j^m|$$

dove  $M$  indica l'insieme dei mesi in cui l'unità  $j$  ha risposto nel corso dell'anno (i mesi possono non essere contigui) e  $r_h$  il numero dei rispondenti 12 mesi nello strato  $h$  cui appartiene il comune  $j$ .

Nel caso in cui il comune non abbia risposto in nessuno dei 12 mesi dell'anno, la selezione del donatore avviene estraendo casualmente un comune dall'insieme dei comuni rispondenti 12 mesi nello strato.

In entrambe le situazioni il donatore individuato viene utilizzato per imputare congiuntamente tutti i mesi mancanti al fine di preservare il profilo temporale del fenomeno.

In particolare il metodo del donatore è stato utilizzato per integrare i dati riferiti alle seguenti variabili *pivot*:

- numero di fabbricati, numero di abitazioni, superficie utile abitabile e volume del fabbricato nel caso di permessi di costruire relativi all'edilizia residenziale;

- numero di fabbricati, volume e superficie dei nuovi fabbricati per i permessi di costruire riguardante l'edilizia non residenziale.

Nel caso delle opere residenziali la variabile utilizzata per individuare il donatore è costituita dal numero di abitazioni mentre per le opere non residenziali è stata utilizzata la superficie del fabbricato. Una volta individuato, il comune donatore viene utilizzato per imputare tutte le variabili pivot. E' stato condotto un esperimento utilizzando come variabile la superficie utile abitabile arrivando a risultati simili.

Un'ulteriore valutazione dell'effetto dell'integrazione sui dati aggregati è stata operata paragonando la loro dinamica, misurata in termini di variazioni rispetto a dodici mesi prima, con quella che si sarebbe misurata prendendo in considerazione per ciascun mese il solo insieme dei comuni rispondenti anche a dodici mesi di distanza. In particolare la variabile relativa al numero di abitazioni in nuovi fabbricati residenziali è stata ricostruita selezionando i sottoinsiemi di Comuni rispondenti nello stesso mese negli anni 2000 e 2001 e negli anni 2001 e 2002. Dai risultati si evince che il metodo di imputazione sembra modificare in misura limitata le variazioni.

Al fine di operare una ricostruzione delle informazioni riguardanti le caratteristiche dei fabbricati riferite all'universo dei comuni e dei relativi permessi di costruire, l'imputazione dei dati mancanti è stata estesa a tutte le variabili comprese nel modello di rilevazione, acquisendo dal comune donatore individuato per le variabili pivot, anche i dati relativi a tutte le rimanenti caratteristiche dei fabbricati.