

# 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) che il richiedente il permesso di costruire ha l’obbligo di compilare; il questionario è accompagnato da un 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, sempre con cadenza mensile, 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 né questionari né la 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, ci si deve confrontare con 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 e composizione variabile 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 negli anni dal 1995 al 2006.

**Prospetto 1 – Collaborazione dei Comuni per ripartizione geografica - Anni 1995-2006 (valori percentuali)**

Ripartizione geografica	Anni											
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Nord-Ovest	87,0	86,1	85,6	89,0	88,1	89,2	90,5	88,1	88,3	86,7	86,4	85,5
Nord-Est	89,4	89,2	88,8	90,6	91,7	92,9	94,3	93,1	92,3	92,2	91,9	90,9
Centro	89,3	87,8	86,3	82,0	83,5	87,5	82,4	77,1	75,3	76,8	77,7	79,3
Sud	83,5	86,7	84,8	80,9	80,0	78,4	71,8	68,0	71,5	69,9	69,4	66,8
Isole	86,4	84,8	81,2	71,5	72,7	80,7	79,8	72,3	76,4	74,5	77,1	76,6
<b>ITALIA</b>	<b>87,0</b>	<b>87,0</b>	<b>85,6</b>	<b>83,9</b>	<b>84,1</b>	<b>85,9</b>	<b>83,8</b>	<b>80,1</b>	<b>81,0</b>	<b>80,2</b>	<b>80,5</b>	<b>79,7</b>

Il grado di collaborazione a livello nazionale diminuisce lentamente nel tempo, con valori che vanno dall'87,0 per cento del 1995 al 79,7 per cento del 2006. A livello territoriale il tasso di collaborazione più alto si rileva sempre nel Nord-Est, con livelli stabilmente superiori al 90 per cento a partire dal 1999. All'opposto le ripartizioni "Sud" e "Isole" fanno registrare in tutti gli anni i livelli più bassi dell'indicatore, con un minimo assoluto del 66,8 per cento registrato nel 2006 per il Sud. 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<sup>1</sup>.

La collaborazione del primo sottoinsieme viene misurata rappresentata in termini di numero di Comuni rispondenti (Prospetto 2) e in termini di relative quote di popolazione (Prospetto 3). Si osserva, in particolare, che l'incidenza di Comuni che rispondono per tutti i mesi dell'anno è molto elevata: in termini di popolazione la relativa quota è sempre superiore al 70 per cento, superando in molti anni l'80 per cento. Vi è poi un'incidenza significativa di comuni che rispondono 11 o 10 mesi all'anno e, all'opposto, la quota (sempre misurata in termini di popolazione) di quelli che collaborano per meno della metà dei mesi pesa di regola per pochi punti percentuali (con un massimo del 5 per cento nel 2006).

**Prospetto 2 Collaborazione dei Comuni: Comuni capoluogo e non capoluogo con più di 50.000 abitanti per mesi di collaborazione – Anni 1995-2006**

Mesi di collaborazione	Anni											
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
0	1	-	2	4	2	4	6	7	3	6	5	8
1	-	-	-	2	1	-	-	1	1	1	2	-
2	-	-	-	-	4	-	-	-	-	3	1	-
3	-	-	1	-	1	1	2	2	1	1	-	1
4	-	-	1	1	-	1	2	1	-	-	1	1
5	1	2	1	-	-	1	1	2	3	-	-	4
6	3	-	3	2	-	-	2	5	1	2	3	-
7	1	1	-	2	1	-	1	3	4	-	-	2
8	1	3	2	3	2	1	3	2	5	2	2	-
9	3	6	8	1	4	3	1	3	3	1	6	3
10	16	17	8	7	4	5	4	7	10	9	8	7
11	20	28	28	19	6	7	9	24	27	30	19	16
12	114	103	106	119	135	137	129	103	102	105	113	118
<b>Totale</b>	<b>160</b>											

<sup>1</sup> Al fine di mantenere la comparabilità dei risultati dell'analisi della collaborazione con quelli degli anni precedenti, questo sottoinsieme (7.940 comuni) comprende ancora i comuni capoluogo delle nuove province della Sardegna.

**Prospetto 3 Collaborazione dei Comuni: Popolazione dei Comuni capoluogo e non capoluogo con più di 50.000 abitanti per mesi di collaborazione – Anni 1995-2006 (in percentuale - popolazione assoluta per 1.000)**

Mesi di collaborazione	Anni											
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
0	0,3	-	0,6	1,0	0,6	1,1	1,5	2,5	0,8	1,7	1,4	2,1
1	-	-	-	1,5	1,2	-	-	0,2	0,3	0,3	0,6	-
2	-	-	-	-	1,1	-	-	-	-	1,2	0,3	-
3	-	-	0,1	-	0,2	0,2	0,6	0,6	0,3	0,3	-	1,0
4	-	-	0,2	0,2	-	0,3	1,0	0,3	-	-	0,2	0,3
5	0,3	0,5	0,3	-	-	0,3	0,3	0,8	1,5	-	-	1,6
6	1,0	-	0,7	0,6	-	-	5,1	1,7	0,5	1,3	0,8	-
7	0,2	0,2	-	0,6	0,3	-	0,4	0,9	1,3	-	-	0,6
8	0,2	1,0	0,6	0,8	0,6	0,3	1,9	2,1	3,3	0,9	0,4	-
9	1,0	1,7	2,5	0,2	2,7	1,5	0,4	6,3	1,0	0,7	1,6	0,7
10	7,8	7,3	2,6	3,5	2,2	1,9	2,3	2,1	7,9	2,8	2,6	3,0
11	6,8	13,6	12,9	5,7	6,2	3,0	4,2	10,0	10,4	16,8	6,7	11,1
12	82,6	75,8	79,5	86,0	84,8	91,4	82,3	72,7	72,7	73,9	85,3	79,7
<b>Totale</b>	<b>100,0</b>											
Popolazione	20.829	20.712	20.629	20.602	20.516	20.482	20.367	20.394	20.594	20.955	20.993	21.187

Per quanto riguarda i restanti 7.940 comuni non capoluogo con meno di 50.000 abitanti (Prospetto 4), il numero di quelli che rispondono 11 o 12 mesi all'anno dal 1998 in poi supera i 4.000, rappresentando quasi sempre oltre il 60 per cento della relativa popolazione (Prospetto 5).

**Prospetto 4 Collaborazione dei Comuni: Comuni non capoluogo con meno di 50.000 abitanti per mesi di collaborazione – Anni 1995- 2006**

Mesi di collaborazione	Anni											
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
0	558	576	719	1.278	1.492	1.384	1.626	1.875	1.856	1.987	2.068	2.235
1	84	73	112	150	158	203	196	183	193	192	179	168
2	71	66	90	96	169	136	104	104	102	118	109	102
3	74	68	85	123	120	127	92	113	96	95	73	83
4	103	88	143	154	114	108	106	95	92	106	75	85
5	138	113	161	129	111	136	90	111	121	115	89	92
6	163	163	208	148	113	125	123	113	137	142	101	119
7	284	316	307	172	150	140	141	148	116	134	123	109
8	460	557	479	209	207	177	148	170	185	162	167	149
9	868	874	820	360	340	303	213	279	278	283	275	251
10	1.239	1.254	1.182	682	536	525	307	453	484	498	526	435
11	1.678	1.618	1.546	1.367	1.170	1.146	582	871	1.148	1.109	1.127	890
12	2.220	2.174	2.088	3.072	3.260	3.430	4.212	3.425	3.132	2.999	3.028	3.222
<b>Totale</b>	<b>7.940</b>											

**Prospetto 5 Collaborazione dei Comuni: Popolazione dei Comuni non capoluogo con meno di 50.000 abitanti per mesi di collaborazione –Anni 1995- 2006 (in percentuale - popolazione assoluta per 1.000)**

Mesi di collaborazione	Anni											
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
0	4,9	4,8	6,0	12,0	11,0	10,8	12,8	16,6	15,7	17,3	18,8	19,8
1	0,7	0,6	1,0	1,2	1,3	1,5	1,9	1,5	1,8	1,6	1,6	1,4
2	0,9	0,6	0,7	1,1	4,7	1,1	1,1	1,0	1,1	1,3	1,0	1,2
3	0,6	0,8	0,8	1,0	1,2	0,9	0,7	1,4	1,1	1,0	0,7	0,8
4	1,1	0,8	1,5	1,9	1,2	0,9	0,9	1,1	1,0	1,1	0,9	0,8
5	1,4	1,0	1,7	1,4	0,8	1,9	1,2	1,3	1,3	1,1	1,0	1,2
6	1,5	1,6	2,7	1,7	0,9	1,2	1,6	1,1	1,5	1,5	1,1	1,4
7	2,8	3,2	3,1	2,1	1,4	1,7	1,7	1,5	1,5	1,4	1,5	1,2
8	5,4	6,4	4,8	1,9	1,9	1,8	2,1	2,4	2,6	2,3	2,6	1,9
9	10,0	9,9	9,3	3,8	3,0	3,7	2,7	4,1	3,7	3,6	4,2	3,6
10	14,7	14,8	13,8	7,9	4,5	6,4	4,1	6,0	6,5	6,5	7,2	6,6
11	21,4	20,9	20,3	15,8	9,5	13,8	7,6	12,5	16,0	15,6	14,6	11,2
12	34,7	34,5	34,5	48,2	58,5	54,1	61,5	49,4	46,2	45,5	44,8	49,0
<b>Totale</b>	<b>100,0</b>											
Popolazione	36.015	36.132	36.247	36.302	36.393	36.442	36.593	36.927	37.294	37.507	37.759	37.944

### 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, utilizzando le informazioni relative al mese che presenta la minima distanza rispetto alla media dei valori rilevati nei mesi con risposta. 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 con l'utilizzo della media e un metodo di ponderazione. Per selezionare la metodologia più adatta è stato condotto, sui dati riferiti agli anni 2000-2002, 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 uno dei mesi del generico anno, all'interno di ciascuno strato definito dalle variabili ausiliarie 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 riguardanti 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 totale del fabbricato. Una volta individuato, il comune donatore viene utilizzato per imputare tutte le variabili pivot. È 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. I risultati sembrano indicare che il metodo di imputazione modifica in misura limitata tale dinamica. 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.