

## Nota metodologica

La rilevazione sui permessi di costruire è condotta con copertura censuaria, raccogliendo mensilmente le informazioni sui progetti di fabbricati nuovi, residenziali e non residenziali, o di ampliamento di fabbricati preesistenti, per i quali sia stato ritirato regolare “Permesso di costruire” presso gli uffici comunali di competenza. I risultati della rilevazione (di seguito **rilevazione strutturale**) vengono diffusi a frequenza annuale su un ampio insieme di variabili raccolte mediante questionario. Per maggiori informazioni si può consultare il sito dell’Istat all’indirizzo: <http://www.istat.it/it/archivio/36534>.

Il Regolamento Ue sulle statistiche congiunturali (n. 1165/98), prevede la produzione e la trasmissione a Eurostat a cadenza trimestrale dei principali indicatori dei permessi di costruire, considerando che essi costituiscano una buona approssimazione dei futuri lavori del settore delle costruzioni, e ne anticipino quindi l’andamento.

A partire da dicembre 2003, al fine di adempiere ai requisiti fissati dal Regolamento sopra citato è stata avviata un’indagine su un campione di Comuni (di seguito **rilevazione rapida**) i cui risultati, permettono di giungere, con la dovuta tempestività, a una stima degli aggregati relativi all’universo. Dal 2003 sono, quindi, fornite a Eurostat le stime trimestrali dei permessi di costruire a 90 giorni dalla fine del trimestre di riferimento; le stime sono riviste a 180 giorni. Le variabili stimate sono un sottoinsieme molto ristretto di quelle misurate dalla rilevazione strutturale, riguardando esclusivamente due dimensioni dell’attività edilizia residenziale (il numero di abitazioni e la superficie utile abitabile) e una di quella non residenziale (la superficie).

I due processi hanno in comune la fase di raccolta dei dati forniti dai Comuni, mentre differiscono per un diverso modo di trasmissione all’Istat di Roma dei modelli mensili compilati. Nel caso della rilevazione rapida, per ridurre i tempi del processo, i Comuni trasmettono le informazioni direttamente all’Istat. La rilevazione strutturale avviene, invece, tramite una rete di raccolta intermedia costituita dagli uffici di statistica presso le CCIAA (Camera di Commercio Industria Artigianato e Agricoltura) che hanno competenza provinciale<sup>1</sup>. Di conseguenza, con riferimento agli indicatori dei permessi di costruire, si producono sia risultati preliminari a cadenza trimestrale con l’ausilio della rilevazione rapida, sia risultati annuali definitivi basati sulla rilevazione strutturale.

Ciò porta a disporre di due serie storiche:

- una trimestrale con risultati provvisori a 90 giorni, rivisti a 180 giorni dal periodo di riferimento;
- una annuale, disponibile a distanza di 18 mesi dal periodo di riferimento.

Per quel che riguarda le caratteristiche dell’indagine rapida, si deve in primo luogo osservare che il campione è costituito da 814 Comuni che mensilmente trasmettono i modelli raccolti direttamente all’Istat. Comprende 160 Comuni capoluogo e non capoluogo con più di 50.000 abitanti, che costituiscono uno strato autorappresentativo, mentre i restanti 654 Comuni, estratti casualmente secondo un disegno probabilistico, sono stratificati per ripartizione geografica e classi di popolazione in 20 gruppi (Prospetto 1).

### PROSPETTO 1. COMUNI CAMPIONE PER RIPARTIZIONE E CLASSI DEMOGRAFICHE

Numero di unità campionarie (escluso lo strato delle 160 unità autorappresentative)

Classi di popolazione	Nord-Est	Nord-Ovest	Centro	Sud-Isole	Totale
Fino a 3000	28	79	15	29	151
Da 3001 a 7000	45	57	20	42	164
Da 7001 a 13000	54	37	22	38	151
Da 13001 a 25000	28	32	14	32	106
Da 25000 a 50000	13	21	20	28	82
<b>Totale</b>	<b>168</b>	<b>226</b>	<b>91</b>	<b>169</b>	<b>654</b>

<sup>1</sup> A partire da luglio 2010, a questo tradizionale metodo di raccolta si sta sostituendo la compilazione e l’invio on line dei questionari: per maggiore informazione si consulti il sito dell’Istat all’indirizzo; <https://indata.istat.it/pdc/index.php?pas=2>.

Riguardo il sottoinsieme di Comuni autorappresentativi, in caso di mancata risposta parziale (MRP), l'imputazione dei dati aggregati mancanti avviene attraverso il valore medio che le variabili oggetto di stima (abitazioni e superfici) assumono nello stesso Comune nell'arco di 12 mesi (se il Comune ha risposto almeno una volta). In caso di mancata risposta totale (MRT), che si verifica quando il Comune non risponde su 12 mesi consecutivi, si utilizza il metodo basato sul donatore trasversale di minima distanza per classi.

Per quanto riguarda il trattamento delle mancate risposte dei restanti 654 Comuni campione, si applica una metodologia basata sulla riponderazione delle osservazioni rilevate mensilmente sui Comuni rispondenti. Inoltre, poiché al momento del calcolo delle stime la copertura del campione non è completa, è opportuno sfruttare l'informazione ausiliaria proveniente dai restanti Comuni non campione ma rispondenti alla rilevazione strutturale in tempo utile per l'inserimento nella stima stessa.

La metodologia di imputazione attualmente utilizzata, a partire dai dati aggregati per comune, opportunamente stratificati per ripartizione e classe di popolazione, stima per mese il numero totale delle abitazioni in nuovi fabbricati residenziali e la superficie totale non residenziale per tali comuni, attribuendo un peso al valore assunto dalle variabili di interesse nei Comuni rispondenti nel periodo considerato.

Lo stimatore utilizzato è denominato stimatore FABI, ed è il seguente:

$$\tilde{Y}_t = \sum_{k \in r_t} y_{t,k} a_{t,k}$$

dove  $y_{t,k}$  è la variabile d'interesse al tempo  $t$  nella  $k$ -esima unità e  $a_{t,k}$  è il peso attribuito alla stessa unità, determinato in base ai parametri stimati attraverso un modello di regressione logistica ed è pari a:

$$a_{t,k} = 1 / (\pi_{at,k} (\pi_{t,k|at} \varphi_{k|st,ms})) + ((1 - \pi_{t,k|at}) (\varphi_{k|s_t} \mathbf{m}_s))$$

in cui:

$$\pi_{at,k} = 1;$$

$$\pi_{t,k|at} = \text{pr di inclusione in } s_t \text{ (campione di 654 unità);}$$

$$\varphi_{k|st} = \text{pr di risposta;}$$

$$\mathbf{m}_s = \text{tasso di risposta;}$$

$$(1 - \pi_{t,k|at}) = \text{pr di non inclusione in } s_t \text{ (campione di 654 unità);}$$

$$\varphi_{k|s_t} = \text{pr di risposta;}$$

$$\mathbf{m}_s = \text{tasso di risposta.}$$

Dalla somma dei valori stimati per il gruppo dei 160 Comuni autorappresentativi e per quello dei restanti Comuni campione e non campione, si ottiene il numero totale delle abitazioni nei nuovi fabbricati residenziali, la relativa superficie utile abitabile e la superficie totale non residenziale per l'universo dei Comuni italiani<sup>2</sup>.

Le stime dei valori trimestrali dei permessi di costruire ottenute tramite la procedura ora accennata non possono essere del tutto coerenti, in termini di media annua, con i risultati annuali che sono derivati dall'informazione completa trattata nella rilevazione strutturale. La differenza deriva, in

<sup>2</sup> Per maggiori informazioni si rimanda a:

Bacchini F., Iannaccone R., Otranto E. "L'imputazione delle mancate risposte in presenza di dati longitudinali: un'applicazione ai permessi di costruzione italiani". Contributi Istat, n. 4/2005.

Falorsi P.D., Alleva G., Bacchini F., Iannaccone R. "Estimated based on preliminary data from a specific subsample and from respondents not included in the subsample". Statistical Methods and Applications, vol. 14, n.1, 83-99, 2005.

Garozzo S., Rallo G., La rilevazione dei permessi di costruire: il controllo e la correzione dei dati. Contributi Istat, n. 13/2008.

primo luogo, dalla necessità di utilizzare nei due contesti metodologie di imputazione dei dati mancanti sostanzialmente diverse. All'interno di quello censuario, l'imputazione poggia su un patrimonio di microdati comunali molto più ricco e articolato ma che, per considerazioni di fattibilità delle procedure, non considera la dimensione a frequenza mensile dell'informazione e non permette, quindi, la stima di dati infrannuali. Nella stima trimestrale, come già notato, la mancata risposta è invece trattata a livello aggregato, mediante uno stimatore che utilizza una tecnica di riponderazione dell'informazione pervenuta. Occorre ribadire che le statistiche annuali non sono soggette a revisioni e vengono diffuse per un insieme di variabili molto più ampio rispetto a quelle misurate per gli indicatori trimestrali.

La differenza tra somme annue delle stime trimestrali e statistiche annue ha consigliato, nella prima fase di gestione delle due rilevazioni parallele, di utilizzare i risultati dell'indagine rapida esclusivamente per la trasmissione a Eurostat che considera tali dati come stime confidenziali da inserire nel calcolo degli aggregati europei. Tuttavia, data l'importanza di questo set di indicatori congiunturali, si è scelto di mettere a punto una procedura basata su tecniche di *benchmarking* tale da rendere coerenti le serie trimestrali con i risultati annuali, e permettere di diffonderle a livello sia europeo, sia nazionale.

Nel prossimo paragrafo sono sintetizzate le principali tecniche di aggiustamento o di *benchmarking* attraverso le quali, per lo stesso fenomeno, si garantisce la coerenza tra le due diverse fonti disponibili, mentre, nell'ultimo paragrafo sono presentati i risultati del *benchmarking* per le tre variabili principali che vengono rilevate attraverso l'indagine: il numero di abitazioni in fabbricati residenziali, la superficie utile abitabile e la superficie di fabbricati non residenziali.

### Metodi per il benchmarking

Il calcolo della stima trimestrale dei dati dei permessi di costruire rappresenta un classico esempio di applicazione di tecniche di aggiustamento o di *benchmarking* attraverso le quali, per lo stesso fenomeno, si garantisce la coerenza tra le due diverse fonti disponibili:

- 1) le stime trimestrali  $x_t$ ,  $t = 1, 2, \dots, T$  (fonte a frequenza alta);
- 2) le stime annuali  $y_a$ ,  $a = 1, 2, \dots, A$  (fonte a frequenza bassa).

Attraverso le tecniche di *benchmarking* le discrepanze annuali (additive)<sup>3</sup>  $d_a$ ,

$$d_a = y_a - \sum_{t=(a-1) \times 4 + 1}^{a \times 4} x_t = y_a - x_a,$$

vengono ripartite in maniera da preservare la dinamica della fonte trimestrale (*movement preservation criterion*) e da soddisfare i livelli (vincoli) della fonte annuale, generando una nuova serie trimestrale  $x_t^*$ , tale per cui

$$d_a^* = y_a - \sum_{t=(a-1) \times 4 + 1}^{a \times 4} x_t^* = 0 \quad \forall a.$$

Si intuisce come i criteri di aggiustamento finalizzati a preservare la dinamica della fonte infrannuale debbano coinvolgere le discrepanze di almeno due anni per evitare le discontinuità della serie  $x_t$  da un anno all'altro.

Rientra nella categoria dei metodi di aggiustamento che ripartiscono  $d_a$  su più anni il metodo di Denton (1971) con le sue varianti: additiva e proporzionale, alle differenze di ordine uno e due.

Come sottolineato in Cholette e Dagum, il metodo di Denton oltre ad avere il difetto di introdurre un eventuale effetto transitorio all'inizio della serie storica (dovuto alla condizione iniziale  $x_0^* = x_0$ ), fa sì che la previsione della discrepanza annuale alla fine della serie dipenda dalle ultime due discrepanze.

<sup>3</sup> Le discrepanze proporzionali sono definite come  $d_a = y_a / \sum_{t=(a-1) \times 4 + 1}^{a \times 4} x_t = y_a / x_a$ .

Per gli indicatori dei permessi di costruire, un esercizio di simulazione ha dimostrato che la previsione dell'ultima discrepanza non riduce necessariamente l'entità della revisione dei dati trimestrali. Ciò è dovuto alle caratteristiche sia del profilo infrannuale delle stime preliminari (ad esempio negli ultimi due anni disponibili sembra essere scomparsa la componente stagionale), sia delle discrepanze annuali (prive di sistematicità).

Il metodo di Denton può essere approssimato nell'ambito di un modello di regressione più generale proposto da Cholette e Dagum (1994) all'interno del quale, non solo possono essere considerate una costante e delle variabili funzioni del tempo, ma è possibile considerare non vincolanti i *benchmark* annuali (*non binding benchmark*).

Il modello può essere così espresso (Cholette e Dagum, 1994 e 2006):

$$\mathbf{x}_T = \mathbf{R}\boldsymbol{\beta} + \mathbf{x}_T^* + \mathbf{e}, \quad E(\mathbf{e}) = 0, \quad E(\mathbf{e}\mathbf{e}') = \mathbf{V}_e$$

$$\mathbf{y}_A = \mathbf{J}\mathbf{x}_T^* + \boldsymbol{\varepsilon}, \quad E(\boldsymbol{\varepsilon}) = 0, \quad E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \mathbf{V}_\varepsilon, \quad E(\mathbf{e}\boldsymbol{\varepsilon}') = 0,$$

dove la matrice  $\mathbf{R}$ , di dimensione  $T \times H$ , contiene  $H$  regressori e  $\boldsymbol{\beta} = (\beta_1 \beta_2 \dots \beta_H)'$  è il vettore dei relativi coefficienti. Di seguito i *benchmark* annuali si considerano vincolanti e quindi  $\boldsymbol{\varepsilon} = 0$ .

Nel modello appena descritto, per la distribuzione delle discrepanze annuali, un ruolo cruciale è dato dai regressori  $\mathbf{R}$  e dalla matrice della varianza degli errori infrannuali  $\mathbf{V}_e = \boldsymbol{\Xi}^\lambda \boldsymbol{\Omega} \boldsymbol{\Xi}^\lambda$ , dove  $\boldsymbol{\Xi}$  è la matrice diagonale delle deviazioni standard degli errori,  $\sigma_1, \sigma_2, \dots, \sigma_T$ , e  $\boldsymbol{\Omega}$  contiene le autocorrelazioni degli errori per i quali generalmente si presuppone un modello autoregressivo (AR) di ordine uno. Il parametro  $\lambda$  può assumere i valori 0 (per errori omoschedastici),  $\frac{1}{2}$  o 1 (per errori eteroschedastici).

Dal punto di vista dell'interpretazione, i regressori correggono la parte sistematica delle discrepanze dovuta ad esempio alla sottostima dei livelli annuali da parte della fonte trimestrale. La parte rimanente delle discrepanze viene, invece, distribuita in maniera il più regolare possibile (*smooth*) attraverso un modello AR (per variabili flusso un modello AR(1) con parametro  $\varphi$  compreso tra 0,7 e 0,9).

Dal punto di vista analitico, la scelta di  $\mathbf{R}$  e  $\mathbf{V}_e$  permette di definire la funzione obiettivo (funzione di perdita quadratica) che in termini matriciali è espressa come:

$$\min_{\mathbf{J}\mathbf{x}_T^* = \mathbf{y}_A} \mathbf{e}'\mathbf{V}_e^{-1}\mathbf{e} = \min_{\mathbf{J}\mathbf{x}_T^* = \mathbf{y}_A} \mathbf{e}'\boldsymbol{\Xi}^{-\lambda}\boldsymbol{\Omega}^{-1}\boldsymbol{\Xi}^{-\lambda}\mathbf{e} = \min_{\mathbf{J}\mathbf{x}_T^* = \mathbf{y}_A} \mathbf{e}'_*\boldsymbol{\Omega}^{-1}\mathbf{e}_* \rightarrow \min_{\mathbf{J}\mathbf{x}_T^* = \mathbf{y}_A} \mathbf{v}'\mathbf{v}$$

$$\mathbf{e} = (\mathbf{x}_T - \mathbf{R}\boldsymbol{\beta} - \mathbf{x}_T^*), \quad \mathbf{e}_* \sim AR(1), \quad e_{*t} = \varphi e_{*t-1} - v_t$$

$$\boldsymbol{\Omega} = \begin{bmatrix} 1 & \varphi & \varphi^2 & \dots & \varphi^{T-1} \\ \varphi & 1 & \varphi & \dots & \varphi^{T-2} \\ \varphi^2 & \varphi & 1 & \dots & \varphi^{T-3} \\ \dots & \dots & \dots & \dots & \dots \\ \varphi^{T-1} & \varphi^{T-2} & \varphi^{T-3} & \dots & 1 \end{bmatrix}, \quad \boldsymbol{\Omega}^{-1} = \mathbf{P}\mathbf{P} \quad \mathbf{e} \mathbf{P} = \frac{1}{\sqrt{1-\varphi^2}} \begin{bmatrix} \sqrt{1-\varphi^2} & 0 & \dots & 0 & 0 & 0 \\ -\varphi & 1 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -\varphi & 1 & 0 \\ 0 & 0 & \dots & 0 & -\varphi & 1 \end{bmatrix}$$

La soluzione è:

$$\mathbf{x}_T^* = \mathbf{x}_T^\circ + \mathbf{V}_e \mathbf{J}' \mathbf{V}_d^{-1} (\mathbf{y}_A - \mathbf{J} \mathbf{x}_T^\circ)$$

dove  $\mathbf{x}_T^\circ = \mathbf{x}_T - \mathbf{R}\hat{\boldsymbol{\beta}}$  e  $\mathbf{V}_d = \mathbf{J}\mathbf{V}_e\mathbf{J}'$ , con  $\hat{\boldsymbol{\beta}} = -(\mathbf{R}'\mathbf{J}'\mathbf{V}_d^{-1}\mathbf{J}\mathbf{R})^{-1}\mathbf{R}'\mathbf{J}'\mathbf{V}_d^{-1}(\mathbf{y}_A - \mathbf{J}\mathbf{x}_T)$ .

La funzione di perdita quadratica in termini scalari è:

$$\min_{\mathbf{J}\mathbf{x}_T^* = \mathbf{y}_A} \left\{ \left( \frac{e_1}{\sigma_1^\lambda} \right)^2 + \sum_{t=2}^T \left( \frac{e_t}{\sigma_t^\lambda} - \varphi \frac{e_{t-1}}{\sigma_{t-1}^\lambda} \right)^2 \right\} = \left( -\varphi^2 \right)^{-1} \times \min_{\mathbf{J}\mathbf{x}_T^* = \mathbf{y}_A} \left\{ \left( \frac{e_1}{\sigma_1^\lambda} \right)^2 \left( -\varphi^2 \right) + \sum_{t=2}^T \left( \frac{e_t}{\sigma_t^\lambda} - \varphi \frac{e_{t-1}}{\sigma_{t-1}^\lambda} \right)^2 \right\}$$

Attraverso quest'ultima espressione è possibile derivare la variante alle differenze prime del metodo di Denton, modificato per escludere la condizione iniziale, ponendo  $\lambda = 0$ ,  $\mathbf{R} = 0$  e  $\varphi$  prossimo a 1 (ad esempio  $\varphi = 0,999$ ).

Rendere meno stringente quest'ultima condizione, ridurre cioè il valore di  $\varphi$ , equivale a rendere meno stringente il principio di preservazione del movimento della serie trimestrale preliminare, in quanto confrontando gli errori tra periodi adiacenti solo una quota  $\varphi$  dell'errore riferito al periodo precedente è considerata. D'altro canto, considerare un parametro  $\varphi$  prossimo a 1, implica una forte dipendenza della previsione del benchmark annuale non ancora disponibile dalle ultime due discrepanze. Quest'ultimo aspetto è di particolare rilevanza qualora si vogliano aggiustare i dati trimestrali preliminari in corso d'anno in mancanza del vincolo annuale. Tuttavia, al momento la politica di diffusione dei dati trimestrali dei permessi di costruire non considera questa eventualità, poiché in assenza di sistematicità delle discrepanze annuali gli esercizi di simulazione non hanno dimostrato la superiorità di un metodo rispetto all'altro e soprattutto dei dati trimestrali corretti su quelli preliminari.

Se nella funzione di perdita quadratica, i valori delle deviazioni standard degli errori,  $\sigma_1, \sigma_2, \dots, \sigma_T$  si sostituiscono con i valori trimestrali preliminari  $x_1, x_2, \dots, x_T$  e si pone  $\lambda = 1$ ,

$$\min_{\mathbf{X}_T^* = \mathbf{y}_A} \left\{ \left( \frac{e_1}{x_1} \right)^2 + \sum_{t=2}^T \left( \frac{e_t - \varphi \frac{e_{t-1}}{x_{t-1}}}{x_t} \right)^2 \right\} = \left( -\varphi^2 \right) \times \min_{\mathbf{X}_T^* = \mathbf{y}_A} \left\{ \left( \frac{e_1}{x_1} \right)^2 + \sum_{t=2}^T \left( \frac{e_t - \varphi \frac{e_{t-1}}{x_{t-1}}}{x_t} \right)^2 \right\}$$

si adotta uno schema in cui si considerano gli errori proporzionali (schema più adeguato in presenza di forte stagionalità, in base al quale le discrepanze vengono attribuite proporzionalmente ai valori delle stime preliminari, nel senso che una stima più elevata assorbe una maggiore discrepanza). Analogamente per  $\mathbf{R} = 0$  e per  $\varphi$  prossimo a 1 (ad esempio  $\varphi = 0,999$ ) si ricade nella variante alle differenze prime proporzionali di Denton, modificata per escludere la condizione iniziale. Nel prospetto che segue sono riportati i parametri del modello additivo di regressione di Cholette-Dagum per le diverse varianti.

**PROSPETTO 2. Varianti del modello additivo di Cholette-Dagum**

Varianti	$\mathbf{R}$	$\lambda$	$\Xi^\lambda$	$\varphi$
Correzione additiva	$-1_T$	0	$1_T$	0,729
Approssimazione di Denton modificato (differenze prime)*	0	0	$1_T$	0,999
Correzione proporzionale**	$0 \text{ o } -1_T$	1	$1_T \times x_T$	0,729
Approssimazione di Denton modificato (differenze prime proporzionali)**	0	1	$1_T \times x_T$	0,999

\* Si considera  $\mathbf{X}_T^\circ = \mathbf{X}_T + b$ , dove  $b$  rappresenta la media delle discrepanze annuali additive.

\*\* Si considera  $\mathbf{X}_T^\circ = \mathbf{X}_T \times b$  dove  $b$  rappresenta la media delle discrepanze annuali proporzionali.

**Risultati**

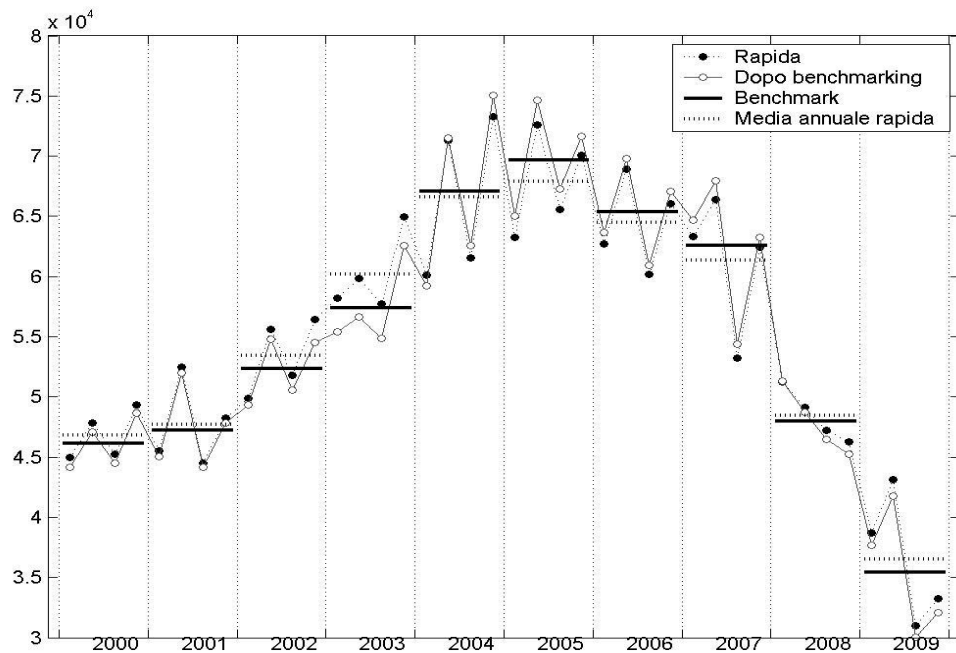
Le tecniche descritte nel paragrafo precedente sono state applicate alle tre variabili che vengono rilevate attraverso l'indagine rapida: il numero di abitazioni in fabbricati residenziali, la superficie utile residenziale e la superficie di fabbricati non residenziali.

Nell'applicazione delle tecniche di benchmarking la prima operazione riguarda il confronto tra le due fonti. Per il numero di abitazioni la più alta discrepanza in valore assoluto tra le due stime si registra nel 2003 e può essere attribuita a fattori legati all'introduzione del nuovo processo di gestione dell'indagine rapida dei permessi di costruire. Per il resto le discrepanze non sono elevate ma si osserva che per i primi quattro anni hanno segno negativo per poi assumere quasi sempre segno positivo nel periodo seguente. Sebbene le differenze in termini relativi siano contenute, il profilo delle variazioni calcolate sui dati annuali della strutturale e della rapida (Prospetto 3) segna un'inversione negli anni 2003 e 2004.

**PROSPETTO 3. Stime del numero di abitazioni in fabbricati residenziali: strutturali e campionarie**

Anno	Numero di abitazioni		Discrepanza additiva (%)	Variazioni su anno precedente	
	Strutturale	Rapida		Strutturale	Rapida
2000	184.424	187.361	-1,6	-	-
2001	189.025	190.759	-0,9	2,5	1,8
2002	209.228	213.752	-2,1	10,7	12,1
2003	229.526	240.796	-4,7	9,7	12,7
2004	268.385	266.303	0,8	16,9	10,6
2005	278.602	271.527	2,6	3,8	2,0
2006	261.455	257.891	1,4	-6,2	-5,0
2007	250.271	245.382	2,0	-4,3	-4,9
2008	191.783	193.815	-1,0	-23,4	-21,0
2009	141.857	146.084	3,2	-26,2	-24,6

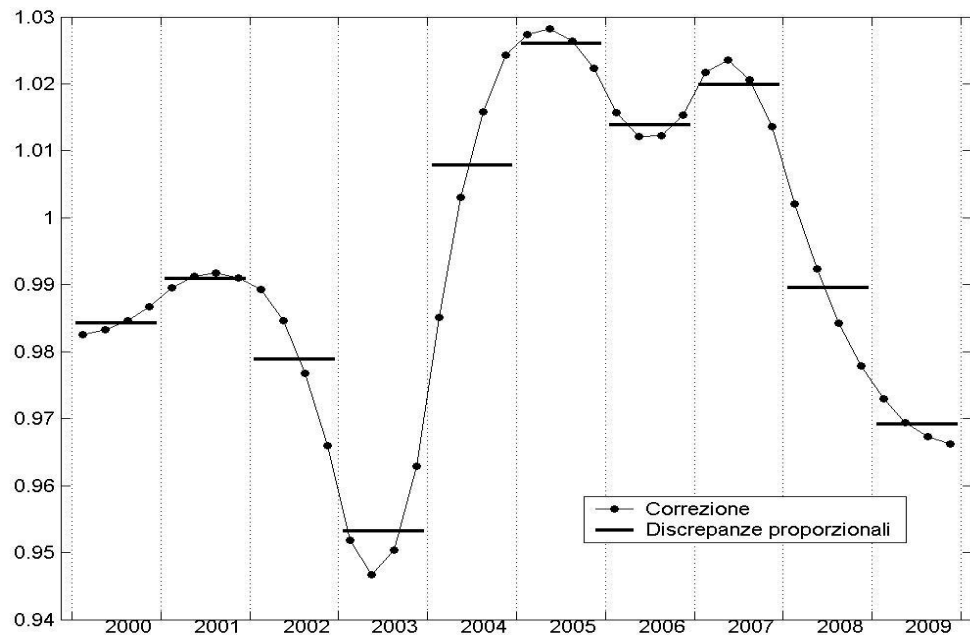
Come già accennato, per la serie del numero di abitazioni sono stati sperimentati diversi metodi di benchmarking nell'ambito del modello additivo di regressione di Cholette e Dagum. Data la ridotta ampiezza delle discrepanze, i diversi metodi hanno portato a risultati simili. Tuttavia è stato selezionato il metodo proporzionale di Denton senza condizione iniziale, in quanto il profilo della serie dopo il benchmarking è il più simile a quello della serie originale (Figura 1).

**Figura 1. Numero di abitazioni: dati trimestrali prima e dopo il benchmarking**


Ovviamente la dinamica delle due serie è molto simile ad eccezione che negli anni 2002 e 2003, che presentano una correzione più elevata.

Tale risultato è confermato anche dalla figura 2 in cui sono mostrate le discrepanze annuali tra le due serie e il coefficiente di correzione dei dati trimestrali (rapporto tra la variabile dopo e prima il benchmarking). Il coefficiente di correlazione misurato sulle variazioni tendenziali delle due serie, per il periodo 2001-2009 è molto alto (pari a 0,983).



**Figura 2. Numero di abitazioni: discrepanze annuali proporzionali e correzioni con il metodo di Denton proporzionale**


I risultati del confronto sono simili per la variabile superficie utile residenziale caratterizzata da un'alta correlazione con il numero di abitazioni: le discrepanze sono di segno negativo per gli anni fino al 2003, con un massimo di differenza in quest'ultimo anno, e di segno positivo dal 2004 al 2009 (Prospetto 4).

**PROSPETTO 4. Stime della superficie utile residenziale: strutturali e campionarie**

Anno	Superficie utile residenziale		Discrepanza additiva (%)	Variazioni % su anno precedente	
	Strutturale	Rapida		Strutturale	Rapida
2000	15.053.472	15.279.428	-1,5		
2001	15.114.416	15.313.882	-1,3	0,4	0,2
2002	16.322.853	16.681.291	-2,1	8,0	8,9
2003	17.563.337	18.410.728	-4,6	7,6	10,4
2004	19.898.618	19.558.594	1,7	13,3	6,2
2005	20.479.027	19.935.873	2,7	2,9	1,9
2006	19.143.787	18.893.571	1,3	-6,5	-5,2
2007	18.383.339	17.873.654	2,9	-4,0	-5,4
2008	14.268.787	14.096.092	1,2	-22,4	-21,1
2009	10.703.097	10.631.485	0,7	-25,0	-24,6

Anche per la superficie utile residenziale è stato applicato il metodo proporzionale di Denton senza condizione iniziale ottenendo la serie riportata in figura 3. La correzione maggiore si ottiene per gli anni 2002 e 2003 (figura 4) e il profilo della serie dopo il benchmarking rimane molto simile a quello della serie della superficie utile residenziale stimata dall'indagine rapida; il cui coefficiente di correlazione delle variazioni tendenziali è pari a 0,978.

Figura 3. Superficie utile residenziale: dati trimestrali prima e dopo il benchmarking

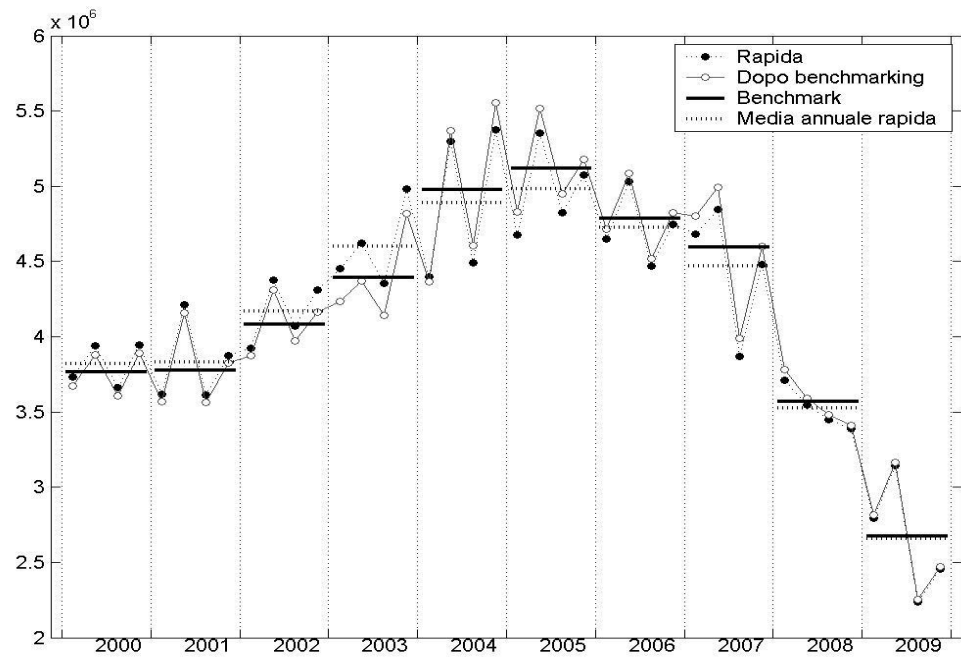
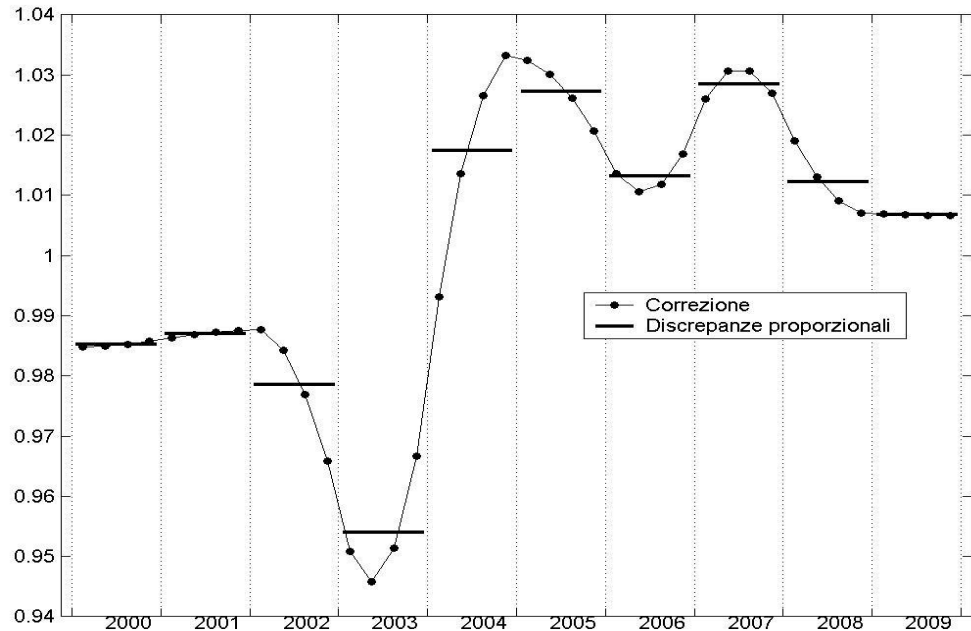


Figura 4. Superficie utile residenziale: discrepanze annuali proporzionali e correzioni con il metodo di Denton proporzionale





La variabile superficie non residenziale presenta un comportamento diverso rispetto alle variabili precedenti. Innanzitutto, la discrepanza risulta più elevata con valori del 7,6% per il 2004 e 7,9% per il 2009.

Inoltre il segno è negativo fino al 2003 (ad eccezione del 2002 in cui la differenza è nulla) e positivo dal 2004. Per tale variabile si sfruttano le proprietà del modello di Dagum - Cholette inserendo per il metodo proporzionale un regressore esterno, in modo da tenere conto dello specifico profilo delle discrepanze. In particolare, si considera un regressore di tipo *level shift* (1 per gli anni dal 2000 al 2003 e 0 per gli anni dal 2004 al 2009) che coglie il cambiamento di segno.

**PROSPETTO 5. Stime della superficie non residenziale: strutturali e campionarie**

Anno	Superficie non residenziale		Discrepanza additiva (%)	Variazioni % su anno precedente	
	Strutturale	Rapida		Strutturale	Rapida
2000	26.772.665	28.537.675	-6,2%		
2001	28.859.846	30.032.264	-3,9%	7,8%	5,2%
2002	36.161.462	36.161.462	0,0%	25,3%	20,4%
2003	28.358.843	28.864.623	-1,8%	-21,6%	-20,2%
2004	29.231.857	27.162.567	7,6%	3,1%	-5,9%
2005	25.137.913	23.971.686	4,9%	-14,0%	-11,7%
2006	24.158.269	23.958.046	0,8%	-3,9%	-0,1%
2007	24.129.655	23.009.958	4,9%	-0,1%	-4,0%
2008	22.916.306	22.123.467	3,6%	-5,0%	-3,9%
2009	17.461.360	16.185.752	7,9%	-23,8%	-26,8%

Le correzioni per la serie trimestrale della superficie non residenziale sono di entità maggiori (Figura 5) soprattutto per alcuni anni (2001, 2005 e 2007). Inoltre l'effetto dell'introduzione del regressore rende più irregolare il profilo delle correzioni (Figura 6).

Tuttavia anche per questa variabile il coefficiente di correlazione tra le variazioni tendenziali calcolate sulla serie prima e dopo il benchmarking rimane elevato (0,967).

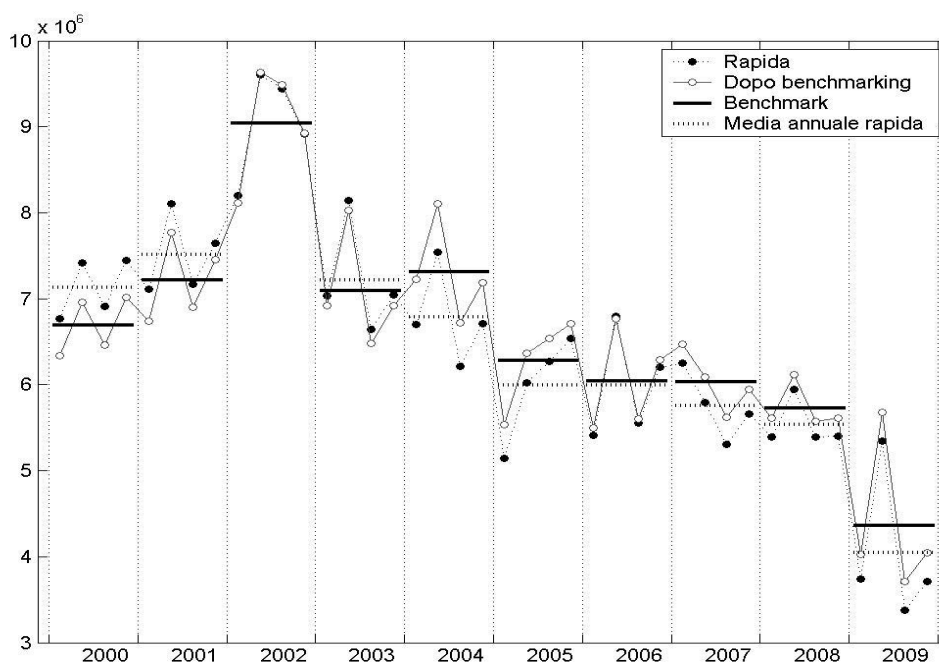
**Figura 5. Superficie non residenziale: dati trimestrali prima e dopo il benchmarking**


Figura 6. Superficie non residenziale: discrepanze annuali proporzionali e correzioni con il metodo di Denton proporzionale

