



RESEARCH DATA CENTER

REX - Remote Execution / Sistema di elaborazione a distanza

Indagine sulle imprese industriali e dei servizi

Disegno campionario e metodi utilizzati



Novembre 2022

Sommario

Sommario.....	2
Disegno campionario e metodi utilizzati.....	3
1. Rimando	3
2. L'indagine sulle imprese industriali e dei servizi.....	3
2.1 Aspetti generali.....	3
2.2 Il disegno campionario.....	3
2.3 Campione e popolazione di riferimento.....	4
2.4 La fase di rilevazione e il questionario.....	4
2.5 L'imputazione	5
2.6 I coefficienti di ponderazione	5
2.7 I risultati dell'indagine pubblicati sul sito della Banca	6

Disegno campionario e metodi utilizzati

1. Rimando

Il presente documento contiene una descrizione generale dei metodi utilizzati nell'Indagine della Banca d'Italia sulle imprese industriali e dei servizi. Per ogni maggiore dettaglio, non presente nel presente documento, sul disegno campionario e le metodologie utilizzate anno per anno, si rimanda alla **Nota metodologica** di ciascuno dei [Fascicoli](#) dedicati alle indagini e reperibili sul sito internet della Banca d'Italia.

2. L'indagine sulle imprese industriali e dei servizi

2.1 Aspetti generali

L'indagine sulle imprese dell'industria in senso stretto e dei servizi rileva in modo continuativo: informazioni anagrafiche, sull'occupazione, sugli investimenti (realizzati e previsti), sul fatturato, sulla capacità produttiva tecnica, sull'indebitamento e sui crediti commerciali. Le indagini raccolgono anche informazioni su altri fenomeni, diversi di anno in anno, che riguardano aspetti specifici di particolare interesse per la ricerca economica.

Fino all'indagine sul 1998 la rilevazione si limitava alle imprese del settore manifatturiero con 50 addetti e oltre. A partire dall'indagine sul 1999 l'universo di riferimento è stato ampliato a tutta l'industria in senso stretto, integrando il campione con imprese della Sottosezione ATECO DF (raffinerie di petrolio, trattamento di combustibili) e della Sezione C (estrazione di minerali) ed E (energia elettrica, gas e acqua). Nel 2001 l'indagine è stata estesa (con un questionario ridotto) anche alle imprese con 20-49 addetti. Dal 2002 la rilevazione è stata ulteriormente estesa alle imprese dei Servizi privati non finanziari con 20 addetti e oltre (escludendo dai servizi destinabili alla vendita, le imprese del credito e assicurazioni, i servizi pubblici e gli altri servizi sociali e personali). Dal 2006 la rilevazione si è estesa anche al settore delle costruzioni con 20 addetti e oltre. L'indagine sul 2013 ha esteso il campione delle costruzioni alle imprese con 10-19 addetti

2.2 Il disegno campionario

La popolazione di riferimento dell'indagine è divisa in strati e da ognuno di essi è estratto casualmente un certo numero di imprese, che formano il campione da rilevare (*schema di campionamento di tipo stratificato a uno stadio*). Gli strati sono costituiti dalle combinazioni di settore di attività economica, classe dimensionale (in termini di addetti) e localizzazione regionale della sede amministrativa dell'impresa.

Gli addetti per la classificazione dimensionale sono stati misurati a fine anno sino all'indagine relativa al 2003; successivamente si fa riferimento agli addetti in media d'anno (per ulteriori dettagli si veda la parte dedicata agli archivi).

Il numero di imprese da rilevare in ogni strato è stato determinato non in proporzione alla loro numerosità nella popolazione (metodo del *campione proporzionale*), ma in modo da ottenere stime affidabili della dinamica di investimenti, occupazione e fatturato: questo ha richiesto di rilevare le imprese di dimensione maggiore in misura superiore alla loro effettiva presenza nella popolazione (metodo dell'*optimum allocation to strata*, che consente di minimizzare



l'errore standard delle medie campionarie attraverso il sovracampionamento degli strati a più elevata varianza). Per particolari obiettivi di ricerca, il sovracampionamento ha riguardato anche le imprese con sede nell'Italia meridionale, anch'esse quindi maggiormente rappresentate rispetto alla loro diffusione nella popolazione italiana di imprese.

2.3 Campione e popolazione di riferimento

Il campione è un panel: le imprese rilevate nella edizione precedente sono sempre ricontattate, se fanno ancora parte della popolazione d'interesse, mentre quelle non più disposte a collaborare sono sostituite con altre omogenee per tipo di attività economica e classe dimensionale. Un'impresa che ha partecipato almeno una volta alla rilevazione si considera uscita dal campione nei casi in cui non appartiene più all'universo di riferimento per cessazione dell'attività, liquidazione o fallimento, incorporazione, o perché l'impresa cessa di appartenere alle attività economiche o alle classi dimensionali oggetto di rilevazione. Inoltre escono dal campione, talvolta solo temporaneamente, le imprese che non sono disponibili a partecipare all'indagine.

2.4 La fase di rilevazione e il questionario

La rilevazione viene svolta annualmente tra febbraio e maggio e, facendo riferimento agli eventi dell'anno precedente (anno di riferimento), attraverso interviste effettuate dalle Filiali della Banca.

I questionari sono redatti separatamente per industria e servizi, per le imprese tra 20 e 49 addetti e per le imprese con almeno 50 addetti. Questa suddivisione è generalmente finalizzata a escludere la somministrazione delle domande caratterizzate da un maggior *respondent burden* per le piccole imprese e a consentire, all'occorrenza, sezioni monografiche distinte per industria e servizi, imprese grandi e piccole. Per le costruzioni è presente un unico questionario per le imprese con almeno 10 addetti

I questionari si compongono di una parte fissa e di una variabile. Nella parte fissa vengono raccolte le informazioni anagrafiche e strutturali sull'impresa, insieme ai dati annuali su investimenti, occupazione, fatturato, risultato d'esercizio, capacità produttiva, indebitamento. La sezione variabile è, di volta in volta, dedicata a temi che richiedono specifici approfondimenti per l'analisi congiunturale o strutturale. I questionari per le imprese dei servizi sono parzialmente diversi da quelli utilizzati per le imprese industriali.

I contenuti e la struttura del questionario vengono definiti dagli economisti della Banca d'Italia nel periodo immediatamente precedente la rilevazione. Le domande nuove vengono sottoposte a test attraverso una rilevazione pilota affidata alle Filiali, al fine di verificare la comprensibilità del testo e la disponibilità delle informazioni presso le imprese.

Per occupazione, investimenti, fatturato, le informazioni vengono richieste con riferimento a tre periodi: l'anno appena concluso (dato di preconsuntivo), l'anno precedente (dato di consuntivo) e quello successivo (dato di previsione).

Dall'edizione relativa al 2010, il caricamento dei dati è stato interamente realizzato tramite una applicazione accessibile via web, che ha consentito anche il caricamento diretto da parte delle imprese a ciò disponibili. La quasi totalità dei caricamenti effettuati dalle imprese è avvenuta tramite compilazione off-line di un modulo interattivo.

2.5 L'imputazione

La percentuale di dati imputati è solitamente contenuta. Tendono ad avere un tasso più elevato di mancata risposta le domande relative a previsioni, in particolare quelle inerenti gli investimenti.

Le variabili imputate sono associate ciascuna ad una variabile flag, che assume il valore vuoto, se il valore corrispondente è stato effettivamente fornito dall'impresa, 1 se il valore è stato imputato. L'elenco delle variabili imputate è disponibile nel documento dedicato alla descrizione degli archivi.

Le metodologie adottate per l'imputazione sono basate su medie di cella o stimatori per regressione e sono finalizzate al miglioramento delle stime aggregate più che alla stima della risposta che la singola impresa avrebbe fornito. Utilizzando i flag, è sempre possibile individuare i valori imputati ed escluderli o sostituirli con le predizioni di un modello diverso.

2.6 I coefficienti di ponderazione

Ad ogni impresa del campione è associato un coefficiente di ponderazione (**PESO**) che indica in sostanza il numero di imprese della popolazione che tale impresa rappresenta¹. Ad alcune imprese rispondenti si attribuisce peso unitario: ciò accade per le imprese autorappresentative, nel caso in cui la rilevazione sia censuaria all'interno dello strato; in un numero limitato di casi si attribuisce peso unitario anche a talune imprese *outlier*, che non si ritengono rappresentative delle altre imprese dello strato per quel che riguarda l'andamento di alcune variabili fondamentali.

In teoria la somma dei pesi in ogni strato del disegno campionario dovrebbe equivalere al numero di imprese della popolazione dello strato. Si è tuttavia evitato di imporre questo requisito alle combinazioni complete di settore, attività economica e area geografica, sia perché alcune combinazioni possono risultare prive di unità campionarie, sia perché un sistema di pesi siffatto potrebbe condurre a stime eccessivamente variabili per alcuni domini di studio. Si sono quindi distinte delle celle di stratificazione, formate dalle combinazioni di classe dimensionale e settore di attività economica e delle celle di post-stratificazione, rappresentate dall'area geografica dove l'impresa è localizzata: nei due insiemi di celle, in modo indipendente, la somma dei pesi di ogni cella restituisce la numerosità della popolazione di interesse in essa contenuta. Si rinvia alla descrizione degli archivi per una esatta definizione di queste variabili.

In ogni occasione di indagine i pesi vengono ricalcolati sulla base della distribuzione della popolazione di interesse alla data più recente disponibile.

I coefficienti di ponderazione sono resi disponibili nell'archivio (variabili **PESO** e **PESORISC**). **Per evitare distorsioni nelle stime se ne raccomanda l'uso in ogni elaborazione. Dal 2007 è anche disponibile PESOADD, che riproduce la distribuzione di addetti della popolazione. Questo tipo di ponderazione è particolarmente indicato per la ponderazione delle variabili di tipo categorico, perché consente di tener conto della diversa scala dimensionale delle imprese.**

¹ E' stata anche costruita negli archivi una variabile PESORISC, contenente il PESO opportunamente moltiplicato per un fattore, di modo che la somma di tutti i PESORISC corrisponda alla numerosità campionaria. Tale variabile può risultare utile nell'uso di programmi che considerano la somma dei pesi per il calcolo dei gradi di libertà.

2.7 I risultati dell'indagine pubblicati sul sito della Banca

I principali risultati dell'indagine sono riportati nella collana Statistiche: indagine sulle imprese industriali e dei servizi; nella Relazione Annuale della Banca d'Italia e nella sua Appendice. Tutti i documenti sono consultabili sul sito internet della Banca d'Italia.

Le differenze che talvolta è possibile riscontrare tra i risultati pubblicati e quelli ricalcolati a partire dai dati micro sono attribuibili a diversi fattori, ad esempio:

1. il database reso disponibile per le elaborazioni include le revisioni di dati, se queste si sono rese disponibili successivamente alla pubblicazione;
2. al momento della pubblicazione, i coefficienti di ponderazione sono calcolati sulla base di dati della popolazione non aggiornati, non essendo disponibili quelli riferiti ai periodi più recenti. Quando i dati della popolazione si rendono disponibili i pesi vengono ricalcolati;
3. la rettifica di singoli dati e la variazione dei coefficienti di ponderazione può dare luogo a revisioni anche nel calcolo dei valori imputati, poiché i valori imputati dipendono da tutti o parte dei rimanenti valori;
4. le variabili espresse a prezzi costanti, essendo riferite all'anno più recente, sono soggette a variazione ogni anno (anche se spesso le variazioni sono di entità modesta).