

Nota metodologica

1. Introduzione

La demografia di impresa si riferisce all'analisi statistica delle caratteristiche demografiche di una popolazione di unità economiche (le imprese) ad un dato istante temporale e di come queste unità si sviluppano nel tempo, con riferimento ad un dato ambito territoriale (o settoriale o dimensionale). La creazione di nuove imprese e la chiusura di quelle non più produttive sono considerati importanti indicatori del dinamismo di un'economia. Oltre a studiare la popolazione delle imprese attive, vengono esaminate le principali caratteristiche demografiche quali la natalità, la mortalità e la sopravvivenza delle imprese.

Particolare attenzione è rivolta all'impatto che queste componenti demografiche hanno sull'occupazione. In particolare, per quanto concerne la natalità delle imprese, il loro sviluppo viene seguito per cinque anni, al fine di esaminare come e se riescono a sopravvivere e a crescere.

Il periodo di riferimento dei dati è l'anno. Ci sono due tipi di variabili nel set di dati, il numero di imprese e l'occupazione (addetti e dipendenti).

La popolazione delle imprese attive si riferisce a tutte le imprese che sono attive nel corso dell'anno, anche per un periodo di tempo limitato.

Le nascite si riferiscono ad unità nate 'da zero' (nate reali) senza il coinvolgimento, ad esempio tramite eventi di scorpori e/o fusioni, di altre unità. Anche le cessazioni si riferiscono alla mortalità delle imprese senza il coinvolgimento di altre unità (cessazioni reali nel periodo). Queste ultime vengono confermate come dato definitivo due anni dopo l'anno di riferimento, al fine di escludere la situazione di unità che possono riattivarsi.

L'occupazione è calcolata in media annua.

2. Fonti dei dati

Per la produzione di statistiche demografiche sulle imprese si fa riferimento al Registro statistico delle imprese attive. In Italia il Registro statistico delle imprese attive (ASIA) è costituito dalle unità economiche che esercitano arti e professioni nelle attività industriali, commerciali e dei servizi alle imprese e alle famiglie e fornisce informazioni identificative (denominazione e indirizzo) e di struttura (attività economica, addetti dipendenti e indipendenti, forma giuridica, data di inizio e fine attività, fatturato) di tali unità. Il Registro statistico rappresenta l'universo di riferimento unico per tutte le statistiche economiche. La sua realizzazione e l'utilizzo di tale universo costituisce il presupposto per rendere integrabili e comparabili le informazioni economiche relative alle unità che descrivono il sistema economico di un paese. Nel corso del tempo, alla tradizionale funzione di fornire la popolazione delle unità statistiche da cui poter estrarre i *frame* ed i campioni per le indagini, il Registro ha assunto ruoli sempre più rilevanti. Oggi il Registro è utilizzato direttamente per la produzione di statistiche sulla struttura della popolazione delle imprese per settore di attività economica, dimensione delle imprese e territorio, per la descrizione delle informazioni diffuse e delle metodologie implementate <http://www.istat.it/it/archivio/131578>.

3. Copertura

I dati di demografia (numero di imprese reali nate, numero di imprese reali cessate e numero di imprese sopravvivenenti e la loro relativa occupazione in termini di addetti e dipendenti) sono disponibili a partire dal 1999.

La copertura in termini di attività economica è quella indicata dal regolamento *Council Regulation (EC) No 295/2008 concerning structural business statistics, Annex IX*. A partire dal 2008, i dati di demografia coprono i settori economici dell'Industria in senso stretto (da B a E della classificazione Ateco2007), delle Costruzioni (F), del Commercio (G), del Trasporto e Magazzinaggio (H), delle Attività dei Servizi e Ristorazione (I), dei Servizi di Informazione e Comunicazione (J), delle Attività Finanziarie e Assicurative (K ad eccezione delle attività di holding, K64.2), delle Attività Immobiliari (L), delle Attività professionali, scientifiche e tecniche (M) e del Noleggio, agenzie di viaggio e servizi di supporto alle imprese (N).

4. Principali concetti e definizioni

Attualmente, la demografia delle imprese fornisce informazioni rilevanti di supporto alle decisioni politiche e indicatori a sostegno della strategia Europa 2020. Gran parte degli indicatori di demografia confluiscono nel sistema OCSE-Eurostat "Entrepreneurship Indicators Programme".

In sintesi, le principali componenti demografiche identificate sono:

- Le imprese nate reali
- Le imprese cessate reali
- Le imprese sopravvivenenti fino a cinque anni
- L'occupazione in termini di addetti e dipendenti correlata a tali componenti
- Indicatori derivati come il tasso di natalità, mortalità e i tassi di sopravvivenza e indicatori basati sull'occupazione.

La metodologia adottata per l'identificazione di tali componenti demografiche è quella prevista nel manuale di Eurostat-OCSE sulla demografia delle imprese (disponibile sul sito Eurostat). I concetti e le definizioni fanno riferimento anche al Manuale delle Raccomandazioni e indicazioni sui Registri statistici, fonte principale dei dati di demografia.

In particolare per poter definire le imprese nate e cessate non si può prescindere dal concetto stesso di impresa.

Impresa

Secondo il Regolamento Eurostat (Council Regulation (EEC) No 696/93) l'impresa è definita come "la più piccola combinazione di unità giuridiche che costituisce un'unità organizzativa per la produzione di beni o servizi, e che gode di un certo grado di autonomia nel processo decisionale, in particolare per l'assegnazione delle sue risorse correnti. Un'impresa svolge una o più attività in uno o più luoghi. Un'impresa può essere una sola unità giuridica".

Impresa attiva: Impresa operativa da un punto di vista economico (ad esempio ha utilizzato forza lavoro o realizzato fatturato) durante il periodo di riferimento, ossia l'anno.

Impresa nata reale

Definizione (Commission Regulation No 2700/98 Eurostat):

Per nascita (reale) di una impresa si intende "la creazione di una nuova attività i cui fattori produttivi scaturiscono dal nulla, ossia dove nessuna altra impresa è coinvolta nell'evento. Sono escluse tutte quelle creazioni dovute a modificazioni strutturali quali fusione, scorporo, ecc... o che risultino collegate secondo specifici criteri di continuità ad unità già esistenti". La creazione di una impresa può essere considerata una nascita (reale), se implica la creazione di nuovi fattori produttivi, e in particolare nuovi posti di lavoro.

In altre parole una nascita si verifica quando l'impresa inizia "da zero" e inizia effettivamente l'attività. La creazione di impresa può essere considerata una nascita reale se i nuovi fattori di produzione, in particolare i nuovi posti di lavoro, si creano. Se un'unità dormiente viene riattivata entro due anni, questo evento non è considerato una nascita.

Impresa cessata reale

Definizione (Commission Regulation No 2700/98 Eurostat):

Per *cessazione* (reale) di una impresa si intende “la dissoluzione di un insieme di fattori produttivi senza che questo evento comporti il coinvolgimento di altre imprese”. Sono escluse le cessazioni per trasformazioni strutturali o, applicando le stesse regole di continuità usate per identificare le imprese reali nate, per legami con altre imprese esistenti.

Imprese sopravvivenenti

Un'impresa nata nell'anno t sopravvive nell'anno $(t+1)$ se continua ad essere attiva nell'anno $(t+1)$ (**sopravvivenza senza modificazioni**).

Se l'impresa non è attiva nell'anno $(t+1)$, si ritiene che sopravviva se la sua attività è rilevata da una nuova impresa (Entrata) che ha iniziato l'attività in $(t+1)$ (**sopravvivenza per incorporazione**).

5. Principali indicatori diffusi

In linea con indicazioni del manuale Eurostat-OCSE sulla demografia delle imprese si riportano le definizioni dei principali indicatori costruiti.

5.1 Gli indicatori basati sulle Imprese

Tasso di natalità: rapporto tra il numero di imprese nate nell'anno t e la popolazione di imprese attive nell'anno t (in percentuale).

Tasso di mortalità: rapporto tra il numero di imprese cessate nell'anno t e la popolazione di imprese attive nell'anno t (in percentuale).

Tali indicatori misurano l'impatto delle imprese nate/cessate sulla popolazione di imprese attive. Possono essere calcolati per classi dimensionali, per settori di attività economica, per aree geografiche ecc...

Tasso lordo di turnover: somma tra il tasso di natalità e il tasso di mortalità.

Tasso netto di turnover: differenza tra il tasso di natalità e il tasso di mortalità.

Questi indicatori sono essenziali a cogliere la dinamica demografica complessiva.

Tasso di sopravvivenza: In generale il tasso di sopravvivenza a $(t+n)$ è il rapporto tra il numero di imprese nate nell'anno t e sopravvissute nell'anno $(t+n)$ e numero di imprese nate nell'anno t (in percentuale).

Il tasso di sopravvivenza misurato in un determinato settore economico o area geografica è una importante misura di performance imprenditoriale. Tassi di sopravvivenza troppo bassi, per esempio, potrebbero agire come disincentivo sia per i giovani imprenditori che per i potenziali creditori, ostacolando la crescita a lungo termine e l'innovazione.

5.2 Gli indicatori basati sull'Occupazione

Creazione e perdita di occupazione: Rapporto tra il numero di addetti delle imprese nate/cessate nell'anno t e l'ammontare degli addetti del totale delle imprese attive nell'anno t . Oppure rapporto tra il numero di addetti delle imprese sopravvissute nell'anno $t+n$ ($n=1,...,5$) e il numero di addetti delle imprese reali nate nell'anno t .

L'osservazione del lavoro creato dalle imprese nate o perso dalle imprese cessate (o creato/perso dalle imprese sopravvissute) fornisce un'indicazione di come la demografia delle imprese contribuisce ai cambiamenti complessivi di occupazione nell'economia.

GJT (gross job turnover): somma del numero di addetti delle imprese nate e cessate durante un anno. Questo indicatore misura la riallocazione dell'occupazione.

NJT (net job turnover): misura la variazione tra la creazione e la perdita di occupazione e si ottiene come differenza tra il numero di addetti delle imprese nate e cessate.

6. Metodo di identificazione delle principali componenti demografiche

Da un punto di vista metodologico, le componenti demografiche possono essere identificate a partire dai dati di base (i dati del Registro delle imprese) applicando una opportuna procedura sviluppata secondo una metodologia armonizzata a livello Europeo. Le principali componenti demografiche identificate sono: le imprese nate reali, le reali cessate e le imprese sopravvissute

6.1 Identificazione delle imprese reali nate

Il processo di identificazione delle imprese reali nate è basato sull'applicazione di una complessa procedura costituita da una serie di passaggi automatici e manuali che mirano ad escludere (ovvero ad identificare) le componenti **"non reali"** dall'insieme delle nuove imprese (**Entrate**) del Registro delle imprese attive con riferimento ad un periodo (anno) t .

Una formalizzazione del flusso demografico è data dalla seguente equazione:

$$N_t = N_{(t-1)} + E_t - U_{(t-1)} = A_{(t-1),t} + E_t$$

dove:

N_t = popolazione di imprese attive nell'anno t

$N_{(t-1)}$ = popolazione di imprese attive nell'anno $(t-1)$

E_t = imprese Entrate nell'anno t . Esse rappresentano quel sottoinsieme della popolazione di imprese attive nell'anno t che hanno intrapreso l'attività per la prima volta nell'anno t e che quindi sono presenti nella popolazione N_t e non in quella dell'anno precedente $(t-1)$.

$U_{(t-1)}$ = imprese Uscite nell'anno $(t-1)$. Esse sono quel sottoinsieme della popolazione di imprese attive nell'anno $(t-1)$, che cessano la loro attività nell'anno t . Queste imprese sono identificate come quelle presenti solo nell'anno $(t-1)$ e non nell'anno t .

$A_{(t-1),t}$ = imprese attive sia nell'anno $(t-1)$ che nell'anno t .

A partire dall'equazione che sintetizza il flusso demografico, l'identificazione delle imprese reali nate si articola nei seguenti passi:

Passo 1:

Il primo passo consiste nel confrontare 3 popolazioni consecutive di imprese attive (N_t , $N_{(t-1)}$ e $N_{(t-2)}$). Questa operazione consente di escludere eventuali imprese riattivate.

$t-2$	$t-1$	t	Output
$N_{(t-2)}$	Missing	E_r	Riattivazioni
Missing	Missing	E_t	Entrate

Pertanto si ha:

$$N_t = A_{(t-1),t} + [E_1 + E_r]$$

Passo 2:

Eventi di cambiamenti strutturali quali ad esempio la fusione tra due o più imprese causano la creazione di nuove unità nel registro delle imprese attive. L'identificazione delle nascite reali viene effettuata eliminando le unità nate a causa di questi eventi (scissioni, fusioni e acquisizioni). E' possibile avere informazioni su questi eventi da diverse fonti (studi pilota, fonti amministrative, fonti statistiche, ecc...).

Pertanto nella precedente equazione l'insieme E_1 viene suddiviso in 2 componenti secondo la seguente formula:

$$N_t = A_{(t-1),t} + [(E_2 + E_{ev}) + E_r]$$

dove:

E_{ev} = Entrate dovute ad eventi di trasformazione.

Passo 3:

L'ulteriore passo consiste nell'individuare ed escludere le "false" entrate, che "continuano" l'attività di altre unità.

La definizione di imprese reali nate (cessate) si basa sul concetto di identità di una unità e quindi, nell'ambito del dinamismo delle popolazioni di imprese, è di fondamentale importanza identificare quando l'osservazione di un evento modificativo determina un evento demografico (discontinuità) ovvero quando il verificarsi dell'evento non cambia l'identità essenziale di una unità (continuità).

Definizione di continuità

"Si ha continuità quando una impresa si modifica senza effetto per la sua identità essenziale determinata dai propri fattori di produzione"; i fattori di produzione sono l'insieme dei mezzi (management, lavoro, terra, capitale, impianti, processi, edifici...) che l'impresa utilizza per la sua azione produttiva. E' evidente che da un punto di vista pratico analizzare tutti (o la maggior parte) dei fattori di produzione e pesarli è estremamente difficile e costoso.

Eurostat suggerisce, come criterio pratico, di utilizzare specifici caratteri, disponibili nell'archivio di imprese, che sono correlati ai più importanti fattori di produzione. L'ipotesi di base è quella che ad un cambiamento di questi caratteri corrisponde un cambiamento nei fattori di produzione. I caratteri presi in considerazione sono:

- L'unità legale che controlla l'impresa (N)
- L'attività economica svolta dall'impresa (S)
- La localizzazione dove l'attività viene svolta (L)

La regola empirica proposta è che un'impresa è considerata una nuova unità (discontinuità) se almeno due su tre dei precedenti caratteri si modificano.

Per l'applicazione delle regole di continuità si utilizzano metodologie di *Record-Linkage* tra le unità secondo le tre variabili identificative: la denominazione d'impresa, l'indirizzo e il codice di attività economica. Ciascuna combinazione di variabili abbinate porta all'identificazione di sottopopolazioni di *matching* costituite dalle unità che presentano almeno 2 variabili abbinate. Per la definizione di *match* viene applicato un processo di *record linkage* che presuppone la standardizzazione di nomi e indirizzi e l'applicazione di opportune regole di *agreement/disagreement*. La scelta delle unità da considerare abbinate si basa su criteri deterministici. Le sottopopolazioni di *matching* sono le seguenti:

Il confronto del settore di attività e della localizzazione: unità entrate che presentano lo stesso codice di attività a 4 cifre e lo stesso indirizzo delle unità nella rimanente popolazione;

- | confronto del settore di attività e del nome;
- | confronto del nome e della localizzazione;
- | unità con legami certi provenienti da fonti statistiche e amministrative.

Questa fase prevede anche controlli manuali per le grandi imprese nuove nate.

Per sottrazione di unità, eliminando dal sottoinsieme delle Entrate E_2 quelle unità che appartengono ad almeno una delle sottopopolazioni di *matching*, si perviene alla identificazione delle imprese reali nate.

$$N_t = A_{(t-1),t} + [(RB_t + E_{cont}) + E_{ev} + E_r]$$

dove:

E_{cont} = Entrate per continuità

RB_t = Imprese Reali Nate nell'anno t .

Per cui riassumendo:

$$RB_t = E_t - E_r - E_{ev} - E_{cont}$$

6.2 Identificazione delle imprese reali cessate

Come per le imprese nate (reali), il processo di identificazione delle imprese cessate (reali) si basa sulla stessa procedura costituita da una serie di passaggi automatici e manuali che mirano ad escludere (ovvero ad identificare) le componenti "non reali" dall'insieme delle Uscite dal Registro. Sono considerate Uscite nell'anno t il sottoinsieme della popolazione di imprese attive nell'anno t , che hanno cessato la loro attività tra il 01.01 e 31.12 dello stesso anno.

Passo 1:

Date tre popolazioni successive di imprese attive, N_t , N_{t+1} e N_{t+2} , la prima componente da identificare è l'insieme delle Uscite nell'anno t (U_t). Esse rappresentano quel sottoinsieme della popolazione di imprese attive nell'anno t che hanno svolto un'attività per l'ultima volta nell'anno t . Vengono identificate confrontando la popolazione N_t con quella dell'anno successivo ($t+1$) e con quella di due anni dopo ($t+2$) ricorrendo al *matching* esatto per codici identificativi. Le Uscite sono, quindi, le imprese attive al tempo t e non attive sia in $t+1$ e in $t+2$. Il confronto con la popolazione relativa a due anni dopo è utile per escludere le imprese che possono riattivarsi.

Passo 2:

Come per le imprese reali nate, l'identificazione delle imprese cessate reali viene effettuata eliminando le unità cessate a causa di eventi di trasformazioni strutturali (scissioni, fusioni e acquisizioni).

Passo 3:

Infine, vengono applicate le regole di continuità al fine di identificare i casi in cui altre unità sono coinvolte nella cessazione di un'impresa. Come per le imprese nate, vengono considerate come variabili di abbinamento il nome dell'impresa (Ragione Sociale), l'Indirizzo e l'Attività economica.

Per l'insieme delle cessazioni "reali" è necessario disporre della popolazione di imprese attive nei due anni successivi a quello di riferimento e ciò implica la conoscenza anticipata di popolazioni che in genere non sono disponibili. Infatti, l'informazione relativa al numero di unità attive nell'anno che cessano durante lo stesso anno è nota solo con un anno di

ritardo; inoltre, tale dato risulta provvisorio poiché per eliminare i casi di riattivazione dovrà essere noto anche lo stato di attività relativo al tempo $t+2$.

Per i dati definitivi, tenuto conto della modalità di identificazione e nell'ottica di un confronto riferito all'anno t , esisterà sempre uno sfasamento temporale tra l'informazione disponibile sui tassi di natalità rispetto a quella sui tassi di mortalità. Ne consegue che il confronto tra i dati definitivi di natalità e mortalità è relativo soltanto agli anni per i quali entrambe le variabili sono disponibili.

E' comunque possibile effettuare il confronto tra natalità e cessazione rispetto allo stesso anno di riferimento t , anno per il quale i dati relativi alla mortalità sono stati stimati adottando la seguente metodologia:

Metodologia di stima delle imprese cessate

La metodologia di stima delle cessazioni "reali" al tempo t si basa sulle distribuzioni delle serie storiche dei tassi di mortalità dall'anno $t-5$ all'anno $t-1$ e sulle informazioni deducibili dalla fonte amministrativa INPS (una delle fonti utilizzate per la costruzione dell'archivio statistico delle imprese attive ASIA e per la quale si dispone di informazioni più aggiornate e relative all'anno $t+1$). La stima viene effettuata a livello aggregato secondo le principali variabili strutturali, quali l'attività economica (classi), la forma giuridica (3 tipologie), la classe di dipendenti (4 classi). In particolare, vengono individuati più di 5.600 strati, dati dall'intersezione delle modalità delle 4 variabili strutturali suddette. Per ogni strato sono calcolati i tassi di mortalità dall'anno $t-5$ all'anno $t-1$ e, solo per gli strati con classi di dipendenti maggiori di zero, viene calcolato il rapporto tra il numero di imprese dell'archivio INPS e il numero di imprese attive dell'archivio ASIA dall'anno $t-4$ all'anno t . Si impone che il tasso di mortalità nello j -esimo strato al tempo t sia funzione della serie storica dei tassi di mortalità del periodo $[t-5, t-1]$ e dei rapporti tra fonte INPS e archivio ASIA.

Formalmente:

$$\text{tasso_mortalità}(t)_j = f\left[\left(\text{tasso_mortalità}(t-i)\right)_j, w_j\right] \text{ per } j=1, \dots, n \text{ e } i=1, \dots, 5.$$

dove:

$$w_j = \begin{cases} 1 & \text{se } j \text{ è con classi di dipendenti}=0 \\ 1 - \left(\frac{n^\circ \text{imprese INPS}}{n^\circ \text{imprese ASIA}} \right)_{(t-i+1)} & \text{se } j \text{ è con classi di dipendenti}>0 \end{cases}$$

Per ogni j -esimo strato, il tasso di mortalità nell'anno t equivale ad un tasso medio ponderato di periodo. Quest'ultimo viene quindi riproporzionato utilizzando il dato INPS al tempo $(t+1)$, ovviamente per i soli strati con dipendenti >0 .

Il numero di imprese cessate "reali" al tempo t nello strato j sarà dato dal prodotto tra il tasso medio ponderato di periodo e il numero di imprese attive (archivio ASIA) nell'anno t .

Al fine di poter dare stime delle cessazioni "reali" anche per regione, il tasso di mortalità dell'anno t nello strato j è stato riproporzionato secondo il tasso medio di mortalità del periodo $(t-5, t-1)$ calcolato per ogni regione.

6.3 Identificazione delle imprese sopravvivenenti

La produzione di statistiche sulle imprese sopravvivenenti si basa essenzialmente su tre popolazioni, derivabili dal processo di individuazione delle imprese nate (cessate) reali.

- Le imprese nate in $(t-1)$ o, le imprese sopravvivenenti in $(t-1)$ nate precedentemente (RB($t-i$)) per $i=2, \dots, 5$.
- Le imprese attive nell'anno t .

c) Le entrate nell'anno t per incorporazione e il loro eventuale legame con imprese già esistenti.

Imprese nate nell' anno ($t-1$)	Imprese attive nell'anno t	Imprese incorporate da nuove imprese nell'anno t	Risultato
X			RB($t-1$) non sopravvive nell'anno t
X	X		RB($t-1$) sopravvive nell'anno t senza modificazioni
X		X	RB($t-1$) sopravvive nell'anno t con modificazioni