

# istat working papers

N.20  
2016

## **Popolazione censita e popolazione anagrafica al 2011: un'analisi esplorativa sulle differenze quantitative tra le due fonti**

*Gallo Gerardo, Tamburrano Maria Tiziana*



# istat working papers

N.20  
2016

## **Popolazione censita e popolazione anagrafica al 2011: un'analisi esplorativa sulle differenze quantitative tra le due fonti**

*Gallo Gerardo, Tamburrano Maria Tiziana*

### **Comitato scientifico**

Giorgio Alleva  
Tommaso Di Fonzo  
Fabrizio Onida

Emanuele Baldacci  
Andrea Mancini  
Linda Laura Sabbadini

Francesco Billari  
Roberto Monducci  
Antonio Schizzerotto

### **Comitato di redazione**

Alessandro Brunetti  
Romina Fraboni  
Maria Pia Sorvillo

Patrizia Cacioli  
Stefania Rossetti

Marco Fortini  
Daniela Rossi

### **Segreteria tecnica**

Daniela De Luca   Laura Peci   Marinella Pepe

## **Istat Working Papers**

**Popolazione censita e popolazione anagrafica al 2011: un'analisi esplorativa sulle differenze quantitative tra le due fonti**

N. 20/2016

ISBN 978-88-458-1915-5

© 2016

Istituto nazionale di statistica  
Via Cesare Balbo, 16 – Roma

Salvo diversa indicazione la riproduzione è libera,  
a condizione che venga citata la fonte.

Immagini, loghi (compreso il logo dell'Istat),  
marchi registrati e altri contenuti di proprietà di terzi  
appartengono ai rispettivi proprietari e  
non possono essere riprodotti senza il loro consenso.

# Popolazione censita e popolazione anagrafica al 2011: un'analisi esplorativa sulle differenze quantitative tra le due fonti<sup>1</sup>

Gallo Gerardo, Tamburrano Maria Tiziana

## Sommario

*L'obiettivo di questo lavoro è effettuare un'analisi esplorativa sui dati di popolazione, di flusso e di stock, che possono aver determinato le differenze tra il dato censuario e il dato anagrafico alla fine del 2011. L'ipotesi di ricerca consiste nel valutare se il divario tra popolazione censita e popolazione anagrafica, registrato nel 2011, sia da ascrivere del tutto alla sotto copertura censuaria, oppure anche agli effetti della dinamica migratoria con l'estero che ha caratterizzato in maniera significativa l'ultimo periodo intercensuario. In particolare, si ipotizza che le differenze di popolazione tra anagrafe e censimento, siano da imputare non solo agli errori censuari, ma anche alle mancate notifiche da parte della popolazione residente (italiana e straniera) che si è trasferita all'estero nel periodo 2002-2011.*

*L'analisi condotta mostra che le mancate notifiche connesse alla mobilità internazionale della popolazione residente rappresentano uno dei fattori latenti delle differenze quantitative tra anagrafe e censimento*

**Parole chiave:** Censimento e anagrafe, mobilità con l'estero, mancate notifiche e impiego di altre fonti amministrative.

## Abstract

*The goal of this study is to explore population data on flows and stocks to derive quantitative information on observed differences between Population Census outcomes and Population Registers' counts at the end of 2011.*

*The research question is to evaluate if the gaps recorded in 2011 by different sources of population counts should be related more to the Census error coverages than the international migratory balance which characterized the first decades of the new millennium. In particular, the basic hypothesis is that differences between population census data and population data coming from local registers could be attributed to nationals living abroad and foreigners who have left permanently the country: both population groups did not notify the local authorities of their departure in the period 2002-2011.*

*The study underlines that the missing notifications connected to the emigration abroad of usually resident population in Italy represent the most important hidden factor of the gaps between population register data and census results.*

**Keywords:** Census and population register data, international mobility, update process of population registers, missing notifications and use of administrative sources.

---

<sup>1</sup> Gli articoli pubblicati impegnano esclusivamente gli Autori, le opinioni espresse non implicano alcuna responsabilità da parte dell'Istat.

## Indice

	<b>Pag.</b>
<b>1 Premessa</b> .....	6
<b>2 Gli esiti del confronto tra censimento e anagrafe: il problema degli individui non trovati al censimento</b> .....	6
<b>3 Fonti e metodi di confronto</b> .....	9
3.1 Indicatori .....	10
<b>4 Il modello statistico per l'analisi esplorativa</b> .....	12
<b>5 Conclusioni</b> .....	16
<b>Riferimenti bibliografici</b> .....	18

## 1. Premessa

A distanza di oltre due anni dalla diffusione dei dati definitivi del 15° Censimento Generale della popolazione e delle abitazioni e quasi un anno dopo la conclusione del processo di revisione anagrafica post-censuaria, l'Istat ha a disposizione numerose informazioni per approfondire lo studio sulle componenti che possono aver caratterizzato gli scarti tra popolazione censita e popolazione anagrafica al 2011.

Inoltre, la messa a punto di un Sistema Integrato di Microdati (SIM), che si compone di dati individuali provenienti da diverse fonti non anagrafiche, fa sì che gli studi sulle statistiche di popolazione possano disporre di un patrimonio informativo fino ad oggi impensabile.

L'obiettivo di questo lavoro è quindi quello di effettuare un'analisi esplorativa delle informazioni quantitative che, attraverso l'analisi dei dati di stock e di flusso del periodo 2002-2011, possono aver determinato, almeno in parte, le differenze tra il dato censuario e il dato anagrafico all'8 ottobre 2011.

L'ipotesi di ricerca consiste nel valutare se il divario tra popolazione censita e popolazione anagrafica sia da attribuire alla dinamica migratoria con l'estero, che ha caratterizzato in maniera significativa l'ultimo periodo intercensuario, e agli effetti che questa comporta negli stock di popolazione. In particolare, si ipotizza che le differenze di ammontare tra censimento e anagrafe al 2011, oltre che dovuti ad una significativa sotto-copertura censuaria (pari a circa 640.000 unità secondo l'indagine di copertura del censimento), siano da attribuire anche alle mancate notifiche da parte della popolazione residente (italiana e straniera) che si trasferisce all'estero. Questo ammontare, difficile da quantificare in valore assoluto, è da mettere in relazione allo stock degli individui non trovati al censimento.

È importante, però, considerare che la coerenza tra dato censuario e dato anagrafico dei comuni italiani dipende non solo dalle loro capacità di performance nella produzione di dati di flusso e di stock, ma anche dai comportamenti "opportunistici" da parte dei cittadini che spesso "non notificano" i cambiamenti della loro dimora abituale, ad esempio, nel caso delle residenze di comodo all'estero.

Il paper si articola nei seguenti paragrafi: il secondo paragrafo tratta degli esiti del confronto censimento-anagrafe 2011 e delle principali poste di bilancio che determinano le differenze quantitative tra dati di censimento e dati di anagrafe; nel terzo paragrafo sono illustrate le fonti di dati, gli indicatori e il metodo di analisi utilizzato per individuare eventuali fattori latenti delle informazioni di base prese in esame; nel quarto paragrafo sono riportati i risultati dell'analisi multivariata realizzata attraverso il metodo delle componenti principali; infine, nell'ultimo paragrafo, si riportano delle considerazioni di sintesi e alcuni spunti di riflessione anche in prospettiva futura.

## 2. Gli esiti del confronto tra censimento e anagrafe: il problema degli individui non trovati al censimento

Secondo i dati definitivi del 15° Censimento generale della popolazione e delle abitazioni, la popolazione residente in Italia ammonta a 59.433.744 unità. Rispetto al 2001, quando si contarono 56.955.744 residenti, l'incremento è del 4,3% da attribuire esclusivamente alla crescita della componente straniera. Alla data di riferimento del censimento (09.10.2011), gli individui presenti nelle Liste Anagrafiche Comunali (LAC) sono pari a 61.249.486 unità. Lo scarto tra popolazione delle LAC e popolazione censita supera 1 milione e 800 mila unità; questa differenza risulta di gran lunga più ampia rispetto ai censimenti passati, sia in termini assoluti che relativi (Tabella 1). Gli scarti tra le due fonti di dati sono imputabili ad entrambe le poste in gioco: da un lato, la sovra-copertura delle anagrafi che si determina sia per una tenuta dei registri anagrafici comunali non sempre adeguata e puntuale (fino al 2001, non tutti i comuni avevano effettuato la revisione delle anagrafi), sia a seguito delle mancate comunicazioni da parte dei cittadini, non solo stranieri, quando cambiano la loro dimora abituale da e per l'estero; dall'altro, alcune delle persone non trovate al censimento potrebbero non essere realmente assenti dal territorio ma trovarsi solo "temporaneamente" fuori dal luogo di dimora abituale, determinando di fatto una sotto-copertura del censimento.

**Tabella 1 - Confronto tra popolazione anagrafica e risultati definitivi dei Censimenti del 1981, 1991, 2001 e 2011**

Fonte di dati	Popolazione residente	Differenza tra popolazione censita e popolazione anagrafica <sup>(1)</sup>	
		Assoluta	Percentuale
Anagrafe al 1.01.1981	57.140.355		
Censimento 1981, dato definitivo	56.556.911	-583.444	-1,03
Anagrafe al 1.01.1991	57.746.163		
Censimento 1991, dato definitivo	56.778.031	-968.132	-1,71
Anagrafe al 1.01.2001	57.844.017		
Censimento 2001, dato definitivo	56.995.744	-848.273	-1,49
Anagrafe all'8.10.2011 (LAC)	61.249.486		
Censimento 2011, dato definitivo	59.433.744	-1.815.742	-3,05

Fonte: Nostra elaborazione su dati Istat

Note: (1) La popolazione anagrafica (quella al 1 gennaio) di ciascun anno censuario fa riferimento alla "popolazione calcolata". Questo ammontare consiste nell'applicare al dato della popolazione residente censita (o popolazione legale) l'equazione generale della popolazione:  $P_t + \text{Nati} - \text{Morti} + \text{Immigrati} - \text{Emigrati} = P_{t+1}$ .

Tuttavia, è il caso di osservare che, nel corso delle operazioni di confronto censimento-anagrafe, la determinazione degli "individui non trovati al censimento" da parte degli Uffici comunali di censimento (Ucc) è avvenuta dopo diversi tentativi di contatto, dichiarati dagli operatori comunali sul Sistema di Gestione della Rilevazione (SGR), e che la consegna dei questionari di censimento del 2011 si è basata sulla spedizione postale attraverso gli indirizzi delle LAC piuttosto che con la consegna a mano tramite rilevatore.<sup>2</sup>

Ad ogni modo, alla determinazione della complessiva differenza tra popolazione censita e popolazione iscritta nelle LAC concorrono due poste di bilancio: la prima è costituita dagli individui censiti e non iscritti in anagrafe, pari a circa 680 mila unità, che rappresentano la misura della sotto-copertura delle LAC determinata a seguito del confronto censimento-anagrafe; la seconda è rappresentata dagli individui non trovati al censimento e iscritti nelle LAC, pari a 2 milioni e 379 mila unità, che costituiscono la sovra-copertura delle LAC<sup>3</sup> secondo il censimento 2011 (Tabella 2).

<sup>2</sup> I rilevatori sono stati impiegati sul campo solo per il recupero dei questionari quando questi non erano pervenuti presso gli Ucc (recupero delle mancate risposte).

<sup>3</sup> I dati di sovra-copertura e sotto-copertura delle LAC, riportate nella tabella 2, fanno riferimento al momento della chiusura di SGR. Conseguentemente ad alcuni aggiustamenti effettuati in occasione della validazione dei dati di popolazione legale, i dati della tabella 2 non corrispondono, per poche unità, ai dati caricati nel Sistema di Revisione dell'Anagrafe (SIREA), vale a dire la piattaforma *on line* utilizzata dai comuni per la revisione anagrafica post-censuaria.

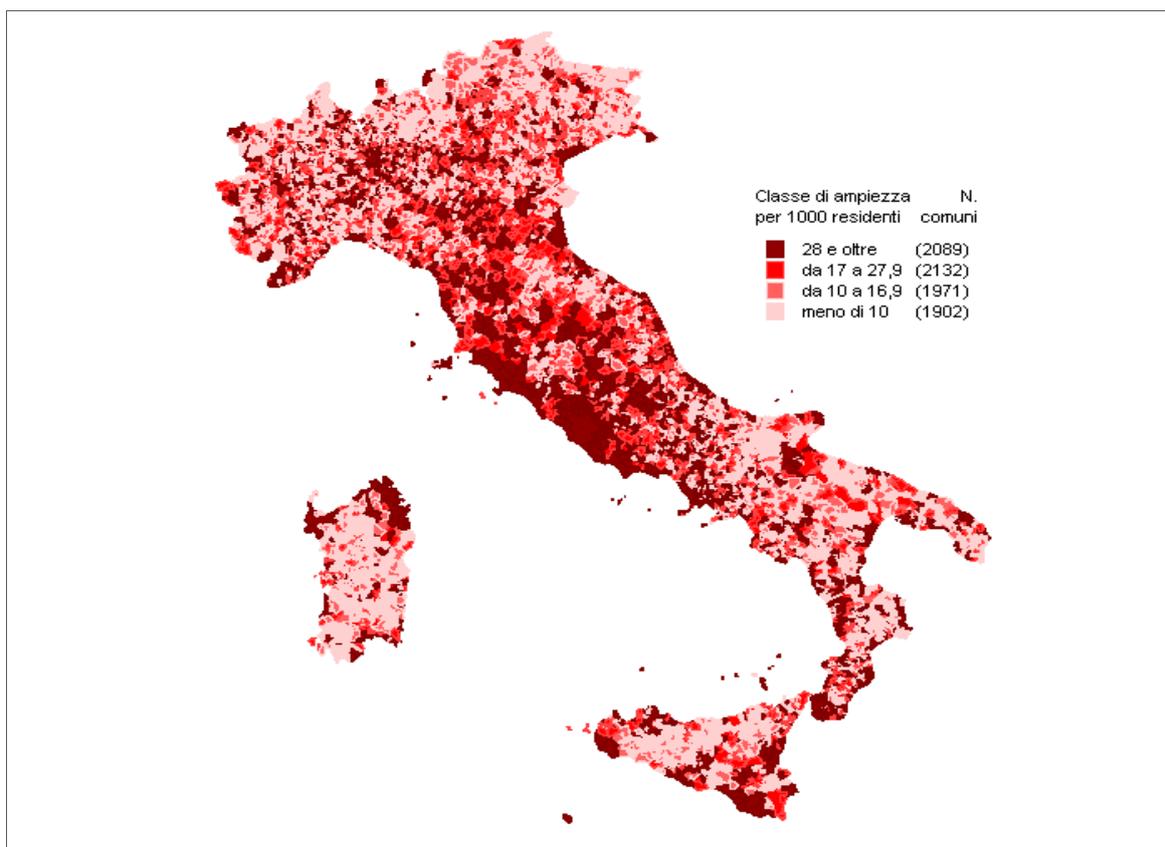
**Tabella 2 – Individui censiti e non iscritti in anagrafe e individui iscritti in anagrafe e non trovati al censimento. Censimento 2011, Valori assoluti e percentuali**

Classi di ampiezza demografica dei Comuni e ripartizioni geografiche	Individui censiti e non iscritti in anagrafe all'8 ottobre 2011				Individui iscritti in anagrafe all'8 ottobre 2011 e non trovati al censimento			
	Totale	Incidenza per 1.000 censiti	di cui:		Totale	Incidenza per 1.000 iscritti in anagrafe	di cui:	
			Cittadini stranieri				Cittadini stranieri	
			Valori assoluti	Incidenza per 1.000 censiti			Valori assoluti	Incidenza per 1.000 iscritti in anagrafe
Fino a 5.000 abitanti	76.606	7,4	14.649	24,3	207.284	19,8	88.626	130,9
5.001 – 20.000	163.931	9,1	36.185	30,1	497.116	27,1	222.822	160,2
20.001 – 50.000	118.610	10,6	28.561	42,9	382.341	33,4	150.963	191,3
50.001 – 100.000	77.759	12,3	18.470	49,1	277.743	42,5	112.515	238,8
oltre i 100.000	244.993	17,9	92.406	78,3	1.015.120	70,2	440.732	288,0
Italia Nord-Occidentale	174.649	11,1	57.025	40,0	593.179	36,6	308.119	183,5
Italia Nord-Orientale	107.928	9,4	36.511	33,5	342.438	29,3	194.041	155,3
Italia Centrale	191.307	16,5	65.868	68,0	726.775	59,7	343.159	275,2
Italia Meridionale	137.900	9,9	21.593	55,8	454.322	31,7	116.114	240,8
Italia Insulare	70.115	10,6	9.274	59,6	262.890	38,4	54.225	270,0
<b>Italia</b>	<b>681.899</b>	<b>11,5</b>	<b>190.271</b>	<b>47,2</b>	<b>2.379.604</b>	<b>38,9</b>	<b>1.015.658</b>	<b>209,0</b>

Fonte: Elaborazione su dati Istat 2012

I censiti non presenti in LAC sono stati rilevati soprattutto nei comuni con più di 100.000 abitanti (quasi 245 mila, pari a circa il 36 per cento del totale), ma raggiungono quote significative anche nei comuni tra 5.000 e 20.000 abitanti (il 24 per cento del totale).

In riferimento agli individui non trovati al censimento, i valori più consistenti si registrano soprattutto nei comuni di maggiore dimensione demografica (quelli con più di 100.000 abitanti) dove si registrano più di 1 milione di unità (oltre il 40 per cento del totale nazionale), ma anche nei comuni tra 5.000 e 20.000 abitanti (circa il 20 per cento). In termini relativi, la media nazionale delle persone non censite è pari a 39 unità per 1.000 iscritti in anagrafe, con maggiore intensità nelle ripartizioni del Centro (quasi il 60 per mille) e delle Isole (intorno alla media nazionale). Nel caso dei cittadini stranieri l'incidenza dei non trovati al censimento è di oltre 200 unità ogni 1.000 stranieri iscritti in anagrafe, anche in questo caso con proporzioni molto elevate nei grandi comuni (più di 280 irreperibili per 1.000 stranieri iscritti in anagrafe) e nelle regioni del Centro Italia (275 irreperibili per 1.000 stranieri iscritti in anagrafe). È da notare che il fenomeno delle persone non trovate al censimento è determinato anche dalle iscrizioni anagrafiche di comodo nelle zone di villeggiatura della costa tirrenica, delle colline interne del Lazio e della Toscana, ma anche nei comuni litoranei di Calabria, Sicilia, Sardegna, Liguria e del Friuli-Venezia Giulia (Figura 1).

**Figura 1 - Individui iscritti in anagrafe e non trovati al censimento**

Fonte: Elaborazione su dati Istat 2012

Appare quindi opportuno analizzare gli individui non trovati al censimento dato che questo segmento di popolazione caratterizza in modo significativo il divario tra popolazione censita e popolazione anagrafica. A riguardo, le informazioni prese in esame considerano non solo i flussi nel periodo intercensuario ma anche l'integrazione tra dati di anagrafe, quali le rettifiche post-censuarie, e dati provenienti da fonti non anagrafiche del SIM.

### 3. Fonti e metodi di confronto

L'elevato patrimonio informativo disponibile a seguito dell'ultima rilevazione censuaria, rappresenta il punto di forza del presente lavoro. La base dati di partenza fa riferimento, come è stato già messo in evidenza, alla lista degli "individui non trovati al censimento ma iscritti nelle liste anagrafiche comunali"; questa lista è stata sottoposta al processo di revisione post-censuaria tramite il nuovo sistema SIREA (Sistema di REvidione delle Anagrafi), messo a punto dall'Istat, che ha consentito di registrare on line le operazioni di revisione anagrafica del censimento 2011. Complessivamente, in SIREA sono stati trattati dagli Uffici di anagrafe dei comuni circa 3 milioni di posizioni anagrafiche: 2,38 milioni individui non trovati al censimento e poco più di 681 mila censiti ma non presenti in LAC.

In secondo luogo, sono stati utilizzati i dati aggregati di flusso, a livello di singolo comune, relativi al "Bilancio demografico annuale 2002-2011 della popolazione residente", risultanti dalla Rilevazione annuale "Movimento e calcolo della popolazione residente" (modello Istat P.2). L'uso dei dati anagrafici garantisce l'omogeneità tra i dati di stock e quelli di flusso necessari per il calcolo dei tassi e inoltre, queste informazioni hanno il vantaggio di distinguere, nell'ambito della dinamica demografica italiana, le migrazioni interne da quelle con l'estero.

Allo scopo di effettuare valutazioni sui motivi per cui l'aggregato di popolazione dei "non tro-

vati al censimento” (2,38 milioni individui) sia di così ampie dimensioni, i dati individuali delle liste di SIREA sono stati integrati con le informazioni nominative di altre fonti amministrative non anagrafiche, definite nel Sistema Integrato di Microdati (SIM). Questo sistema è una struttura informativa di base realizzata mediante l’integrazione concettuale e fisica di microdati acquisiti da fonti amministrative e statistiche ed è stato organizzato in modo da supportare i processi di produzione dell’Istat. In particolare, gli obiettivi del processo di integrazione del SIM sono:

- identificare ogni individuo in fonti diverse con un codice univoco e stabile nel tempo;
- definire per ogni individuo le relazioni logiche e fisiche, nel tempo e nello spazio, tra le informazioni disponibili da fonti diverse.

L’attribuzione di codici identificativi univoci validi nel tempo e per tutte le fonti a disposizione è certamente una condizione essenziale per favorire l’utilizzo dei dati amministrativi per fini statistici.

In particolare, gli archivi utilizzati per analizzare gli individui “non trovati al censimento” sono stati i seguenti:

- a) la popolazione censita al 2011;
- b) le persone decedute risultanti nella base dati dell’Anagrafe Tributaria al 31.12.2011;
- c) le persone risultanti iscritte all’Anagrafe Consolare all’8 ottobre 2011;
- d) le persone che hanno partecipato al mercato del lavoro in Italia, per i quali sono risultati nel corso del 2011 versamenti contributivi all’INPS come lavoratore dipendente, autonomo, subordinato (sia pubblico che privato), para-subordinato e interinale (è compreso anche l’aggregato dei lavoratori in agricoltura, i lavoratori domestici come colf e badanti);
- e) le persone risultanti nel Casellario dei Pensionati al 31.12.2011;
- f) le persone risultanti nell’Anagrafe degli Studenti al 31.12.2011;
- g) le persone per le quali risulta un trasferimento di residenza (modello Istat APR/4 di iscrizione e di cancellazione) in uno dei Comuni italiani nel corso del 2011 e nei due anni precedenti l’anno censuario (2010 e 2009);
- h) i permessi di soggiorno validi al 31.12.2011;
- i) i lavoratori domestici, le colf e le badanti;
- j) i lavoratori autonomi dell’agricoltura.

Pertanto, attraverso tecniche di record linkage e utilizzando il codice fiscale come chiave di aggancio, i dati individuali del SIM sono stati abbinati con un sottoinsieme dell’archivio di SIREA, composto da oltre 1,1 milioni di individui non trovati al censimento che, dopo le azioni di verifica effettuate dagli Uffici di anagrafe, hanno confermato la dimora abituale nel comune di iscrizione anagrafica. Di questo aggregato, però, circa 300 mila individui non hanno segnali di presenza in altri archivi amministrativi del SIM. Questo nuovo sottoinsieme ha consentito di calcolare, a livello comunale, un indicatore molto utile per l’analisi esplorativa del presente lavoro.

### 3.1 Indicatori

Sulla base delle fonti sopra descritte, sono stati calcolati per tutto il periodo intercensuario 2002-2011 i seguenti indicatori: tasso di incremento medio annuo del saldo naturale (SN), tasso di incremento medio annuo del saldo migratorio interno (SMI), tasso di incremento medio annuo del saldo migratorio con l’estero (SME), mentre i movimenti definiti dalle iscrizioni e cancellazioni “per altri motivi” del periodo 2002-2011 non sono stati considerati in quanto fanno riferimento alle operazioni di rettifica anagrafica connesse al censimento del 2001.

Il calcolo dei tassi di incremento medio annuo sopra descritti è quello composto “istantaneamente”, e si basa sul rapporto tra la variazione della popolazione nell’intervallo 0-t ( $tP-0P$ ) e il numero di anni persona vissuti dalla popolazione in tale intervallo:  $t \cdot (tP-0P) / \ln(tP/0P)$ . Pertanto, i tassi d’incremento naturale, quelli di migratorietà netta (interna e con l’estero) sono stati calcolati ponendo al denominatore gli anni persona ottenuti nel modo appena indicato. Come si evince dall’equazione qui di seguito riportata, viene così garantita l’uguaglianza tra il tasso d’incremento

complessivo e la somma algebrica dei tassi d'incremento naturale e migratorio del periodo intercensuario 2002-2011:<sup>4</sup>

$$r = \frac{\ln\left(\frac{{}_tP}{{}_0P}\right)}{t} = \frac{({}_tP - {}_0P)}{t \cdot \left(\frac{{}_tP + {}_0P}{2}\right)} = \frac{\Delta t \cdot SN + \Delta t \cdot SMI + \Delta t \cdot SME}{t \cdot \left(\frac{{}_tP + {}_0P}{2}\right)} = sn + smi + sme$$

Questo tipo di approccio permette di calcolare di quante unità si accresce una popolazione in un determinato intervallo di tempo per ogni unità costitutiva. Per il calcolo occorrono tre elementi: la numerosità della popolazione a date successive, l'entità dell'incremento totale e il tempo durante il quale avviene tale incremento.

Oltre agli indicatori di flusso del periodo intercensuario, sono stati definiti i seguenti indicatori di stock all'8 ottobre 2011:

- dagli esiti del confronto censimento anagrafe, a) il tasso di variazione percentuale tra la popolazione delle LAC e la popolazione censita; 2) il tasso di stranieri non trovati al censimento;
- dagli esiti di SIREA è stato considerato il tasso di irreperibilità al censimento (x 1.000 iscritti in anagrafe);
- dall'accoppiamento delle informazioni di SIREA e SIM, la percentuale di individui che confermano la dimora abituale in SIREA ma non hanno segnali di presenza in SIM.

#### 4. Il modello statistico per l'analisi esplorativa

Al fine di rispondere al quesito di ricerca e ridurre l'insieme delle variabili di partenza in un numero inferiore di fattori latenti, è stata utilizzata l'analisi multivariata. In particolare, l'analisi delle componenti principali (ACP). I nuovi fattori rappresentano le combinazioni lineari delle variabili osservate e sono costruiti in modo da concentrare la maggior parte della varianza delle informazioni di partenza in poche "componenti" estratte per prime. Queste componenti, per definizione tra loro incorrelate, riassumono la complessità delle 8.092 unità statistiche (i comuni italiani) che nel modello di analisi sono state pesate per l'ammontare della popolazione residente censita al 2011.

L'ipotesi di base è che l'osservazione simultanea dei tassi di incremento medio annuo dei dati di flusso della popolazione residente nel periodo 2002-2011 (il saldo naturale, il saldo migratorio interno e il saldo migratorio con l'estero) e le misure di stock al 2011, desunte dagli esiti del confronto censimento-anagrafe (tasso di variazione tra censimento e anagrafe, tasso di stranieri non trovati al censimento) e dalla revisione anagrafica post-censuaria (tasso di irreperibilità accertata), consenta di determinare le "nuove componenti di sintesi" che meglio rappresentano, per gli 8.092 comuni italiani, le variabili di partenza. Inoltre, l'indicatore relativo all'integrazione di alcune informazioni del SIM con gli esiti della revisione anagrafica di SIREA (% di individui che confermano la dimora abituale ma senza segnali di presenza sul territorio) ha permesso di esplorare meglio i fattori latenti degli scarti osservati tra popolazione censita e popolazione anagrafica all'8 ottobre 2011.

Nel complesso, le informazioni utilizzate nel modello possono essere adeguatamente riassunte con le prime tre componenti che riproducono complessivamente una quota di varianza spiegata superiore al 75% (Prospetto 1).

La prima componente, che mostra circa il 38% di varianza spiegata, definisce l'accuratezza dei dati di stock e di flusso osservati ed è da mettere in relazione con i problemi di copertura sia della

<sup>4</sup> Nel caso in cui non vi sia stata alcuna variazione di popolazione durante il periodo considerato, gli anni persona della formula [1] si ottengono moltiplicando la popolazione iniziale per il numero t di anni componenti l'intervallo considerato.

rilevazione censuaria (stranieri non trovati al censimento) che dei registri anagrafici (cancellazioni per irreperibilità censuaria). Questa componente è fortemente saturata in positivo dagli irreperibili al censimento e dalla componente straniera non censita ma mostra anche una correlazione negativa degli scarti tra popolazione censita e popolazione anagrafica.

### Prospetto 1 – Coefficienti di correlazione tra gli indicatori e i primi tre fattori dell'ACP

Indicatori dell'ACP <sup>(a)</sup>	Contributo dei fattori		
	1°	2°	3°
Tasso di variazione (%) tra Liste anagrafiche comunali 2011 e Popolazione censita 2011	<b>-0,928</b>	-0,105	0,046
Tasso di irreperibilità al censimento 2011 desunti da SIREA (x1.000censimento iscritti in anagra-	<b>0,866</b>	-0,198	-0,037
Tasso di Stranieri Non trovati al censimento 2011 (x 1.000 stranieri residenti)	<b>0,786</b>	0,407	-0,073
(%) Individui che confermano la dimora abituale di SIREA <sup>(b)</sup> ma senza segnali di presenza in	0,268	<b>0,737</b>	0,049
Tasso di incremento medio-annuo 2002-2011 del Saldo Migratorio Estero	0,501	<b>-0,711</b>	0,032
Tasso di incremento medio-annuo 2002-2011 del Saldo Naturale	0,080	0,330	<b>0,807</b>
Tasso di incremento medio-annuo 2002-2011 del Saldo Migratorio Interno	-0,198	-0,346	<b>0,744</b>
% di Varianza spiegata	37,9	20,9	17,0
<b>% Cumulata di varianza spiegata</b>	<b>37,9</b>	<b>58,8</b>	<b>75,8</b>

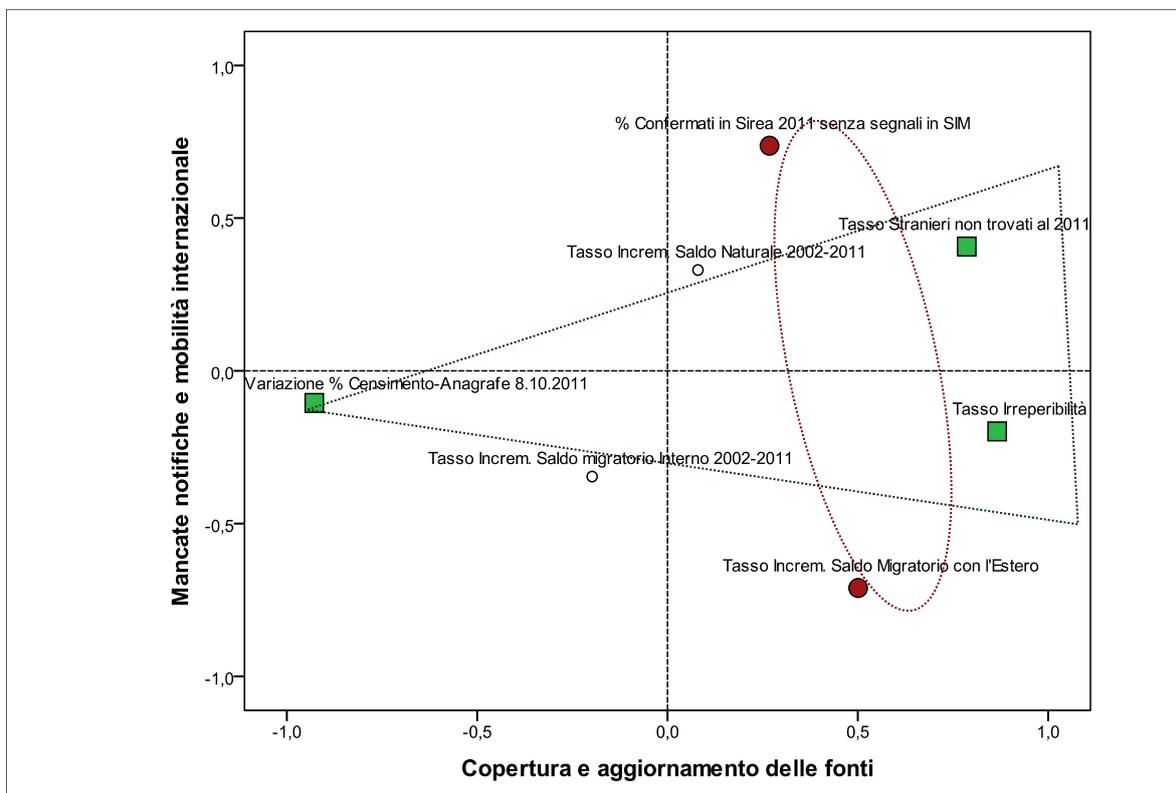
Fonte: Elaborazione su dati Istat di vari anni (2002-2011)

La seconda componente è definita da due variabili osservate che presentano quasi lo stesso peso fattoriale ma sono di segno opposto: l'incremento medio annuo del Saldo Migratorio con l'estero (con una correlazione negativa) e gli individui non censiti che confermano la dimora abituale in SIREA ma non hanno segnali di presenza nel SIM (con una correlazione positiva). La combinazione lineare di questa componente con le variabili osservate lascia supporre che uno dei fattori latenti delle differenze tra anagrafe e censimento sia da mettere in relazione con quella componente di popolazione che potrebbe aver lasciato l'Italia nel periodo intercensuario 2002-2011 senza darne comunicazione in anagrafe.

Il terzo fattore è definito dal saldo naturale e dal saldo delle migrazioni interne ed esprime la dinamica "endogena" della popolazione residente, vale a dire quella caratterizzata dal saldo naturale (nascite e decessi) e dal saldo migratorio interno dei comuni italiani.

Nella figura 2 le variabili di partenza sono state proiettate secondo le coordinate delle prime due componenti che possiamo definire, nel caso della prima, "copertura e aggiornamento delle fonti", e per la seconda, "mancate notifiche e mobilità internazionale".

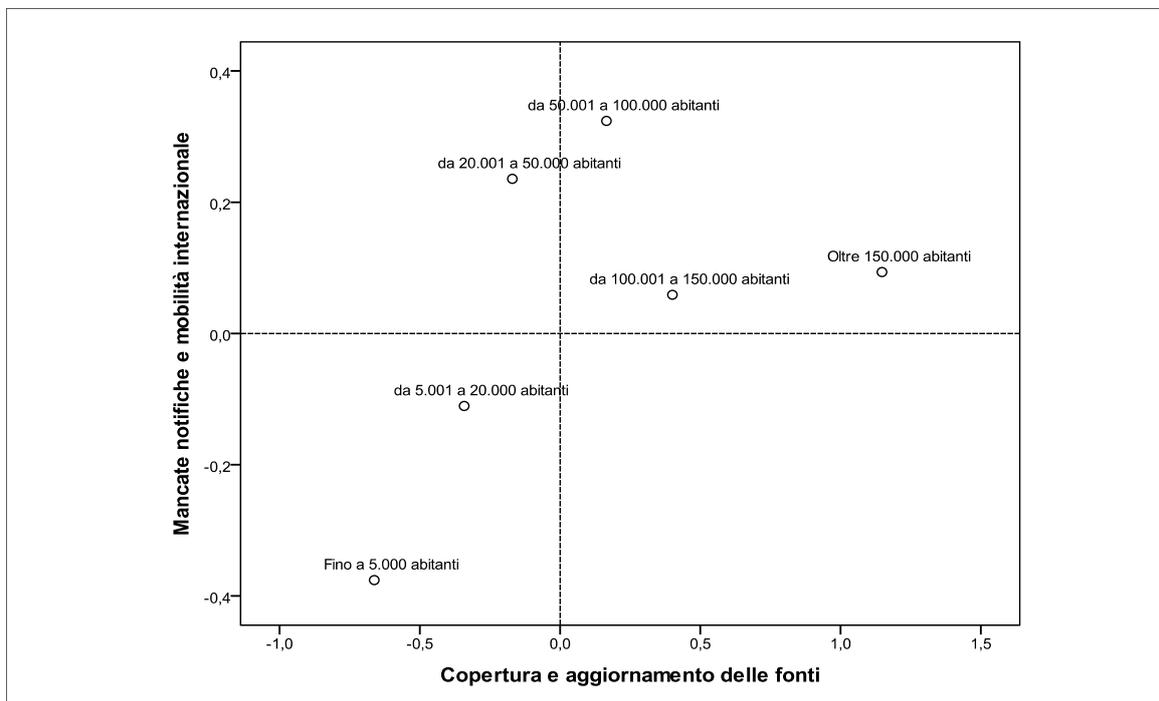
Figura 2 – Proiezione delle variabili di partenza sui primi due fattori dell'ACP



Proiettando le classi dimensionali dei comuni sul piano fattoriale delle prime due componenti, si osserva che i comuni più grandi (soprattutto quelli con oltre 150 mila abitanti) mostrano tassi di irreperibilità e di stranieri non trovati al censimento più elevati, determinando scarti più ampi tra dati di fonte anagrafica e dati di censimento (Figura 3).

Invece, se si considera l'asse delle mancate notifiche e della mobilità internazionale, si può osservare che sono soprattutto i comuni medio grandi (quelli tra 50 mila e 100 mila abitanti) che forniscono il contributo più significativo. Di contro, i comuni con meno di 20 mila abitanti presentano bassi livelli di mobilità con l'estero e una irreperibilità al censimento poco significativa.

**Figura 3 – Proiezione delle classi di ampiezza demografica dei comuni secondo le coordinate della prima e della seconda componente dell'ACP**



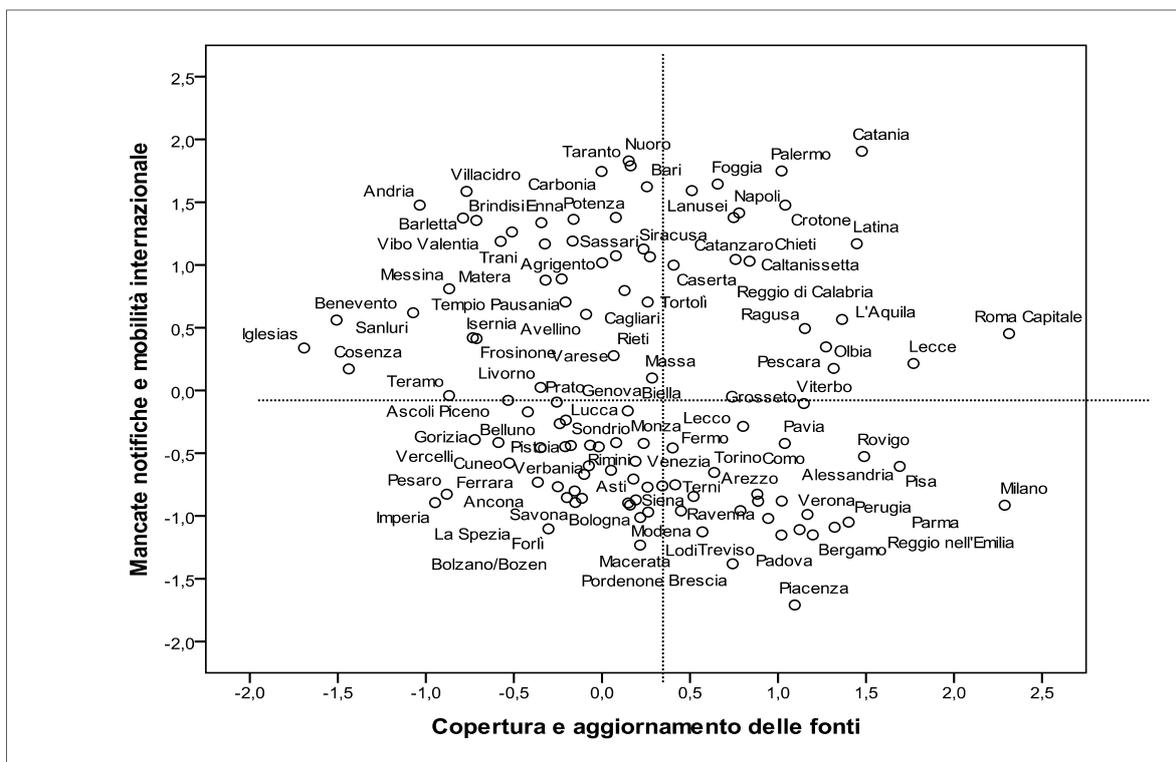
Tuttavia, al fine di arricchire la nostra analisi con un livello territoriale più fine, è risultato utile proiettare i comuni capoluogo di provincia sul piano fattoriale secondo le coordinate delle prime due componenti.

Nella figura 4 è possibile osservare che i comuni del primo quadrante si caratterizzano per una più elevata irreperibilità al censimento con valori sempre più crescenti man mano che ci si allontana dal baricentro. Le unità che si ritrovano nella parte superiore, al di là della bisettrice, si contraddistinguono anche per una più alta incidenza di persone che confermano la dimora abituale in SI-REA ma non hanno segnali di presenza nel SIM. Si tratta prevalentemente dei comuni capoluogo di provincia del Mezzogiorno (Catania, Palermo, Napoli, Bari, Taranto e Crotone) ma anche di diversi comuni dell'Italia centrale. Tra questi, Roma Capitale si colloca nella parte più a destra del primo quadrante a seguito della elevata percentuale di stranieri non trovati al censimento.

Inoltre, se ci si sposta verso il semiasse negativo della prima componente, nel secondo quadrante del piano fattoriale si osservano i comuni caratterizzati anche da differenze significative tra popolazione censita e popolazione anagrafica (ad esempio, Benevento e Cosenza). Così anche i comuni posti nel terzo quadrante, man mano che ci si sposta verso sinistra lungo il primo asse, si registrano variazioni crescenti tra censimento e anagrafe (si osservano, Ascoli Piceno, Gorizia e Vercelli).

Tuttavia, alcuni comuni del terzo quadrante, posti nella parte bassa del secondo asse, sono caratterizzati anche da una scarsa irreperibilità al censimento, da un saldo migratorio con l'estero significativo e, nel complesso, mostrano una variazione tra censimento e anagrafe assai più contenuta rispetto agli altri (si evidenziano, Bolzano, Pordenone, Bologna e Forlì).

**Figura 4- Proiezione dei comuni capoluogo di provincia secondo le coordinate della prima e della seconda componente**



Nel quarto quadrante, caratterizzato da una significativa mobilità internazionale e man mano che ci si sposta verso destra da alti livelli di irreperibilità, si collocano i comuni con una più alta incidenza di cittadini stranieri. Si tratta prevalentemente dei grandi centri urbani del Nord-ovest (si osservano, Milano, Torino, Brescia, Bergamo, Lodi e Pavia), del Nord-est (Venezia, Verona, Treviso, Padova, Modena, Ravenna, Piacenza e Parma) e dell'Italia centrale (soprattutto, Pisa, Siena, Arezzo, Macerata, Perugia e Terni).

## 5. Conclusioni

I risultati di questo lavoro fanno emergere l'esistenza di alcune importanti criticità nei conteggi di popolazione che, pur essendo già note ed in parte indipendenti dal processo di produzione della statistica ufficiale, inducono gli addetti ai lavori a prendere decisioni importanti sulle attività future.

Alla vigilia di una nuova stagione censuaria, caratterizzata da profonde innovazioni metodologiche e orientata verso un maggiore sfruttamento delle fonti amministrative e di indagini ad hoc per valutarne il livello di copertura, le informazioni desunte dall'ultima esperienza censuaria, se combinate con i dati di flusso degli anni intercensuari, possono fornire indicazioni preziose per le statistiche di popolazione.

È ormai noto che, negli ultimi decenni, il dato che più caratterizza la dinamica demografica del nostro paese è la consistente immigrazione straniera dall'estero che rappresenta l'unico fattore di crescita della popolazione. Tuttavia, la mobilità internazionale comporta un impatto notevole sulla qualità dei dati di popolazione che spesso risente degli effetti negativi, dovuti non solo alla difficoltà di rilevare gli stranieri dimoranti abitualmente in Italia, ma anche alla necessità di considerare il problema delle mancate cancellazioni da parte di coloro che lasciano il paese per trasferirsi all'estero.

L'analisi esplorativa condotta sulle variabili che hanno un forte impatto nelle differenze quantitative tra anagrafe e censimento mostrano che uno dei fattori latenti è rappresentato dalle mancate notifiche connesse alla mobilità internazionale della popolazione residente. Questo risultato è confermato, peraltro, anche dal confronto dei dati di flusso a livello internazionale. Infatti, se si osservano i dati sui trasferimenti di residenza degli anni 2008-2012, diffusi dall'OCSE, poco meno di 100 mila persone provenienti dall'Italia si sono iscritte nelle anagrafi tedesche e altre 40 mila in quelle svizzere, a fronte di poco più di 70 mila cancellazioni verso questi due paesi registrate dall'Istat nello stesso periodo.

Pertanto, le mancate comunicazioni relative ai trasferimenti all'estero rappresentano un aspetto di notevole rilevanza per le innovazioni di processo e di prodotto che caratterizzeranno i conteggi di popolazione nell'immediato futuro.

## Bibliografia

- Bolasco S. 1999. *Analisi Multidimensionale dei dati*. Carocci editore, Roma.
- Cortese A., Greco M. 1993. *Il grado di copertura del censimento demografico 1991: considerazioni sulla base del confronto con le risultanze anagrafiche*. Sistema Statistico Nazionale-Istat, Roma.
- Di Franco G. 2011. *Tecniche e modelli di analisi multivariata*. FrancoAngeli editore, Milano.
- Ferrara R., Fornari R. 2013. “Analisi in Componenti Principali”. In *Tecniche di Analisi Multivariata* (a cura di) B. Aragona, Liguori Editore, Napoli.
- Istat. 2012. *Il Censimento in pillole*. Comunicato stampa. Roma. <http://www.istat.it/it/files/2012/12/Comunicato-stampa.pdf> (19 dicembre 2012).
- Istat. 2012. *L'Istat diffonde la popolazione legale dei Comuni italiani*. Comunicato stampa. Roma. [http://www.istat.it/it/files/2012/12/popolazione\\_legale.pdf](http://www.istat.it/it/files/2012/12/popolazione_legale.pdf) (19 dicembre 2012).
- Istat. 2012. *Il Censimento della popolazione straniera*. Comunicato stampa. Roma. [http://www.istat.it/it/files/2012/12/scheda\\_stranieri.pdf](http://www.istat.it/it/files/2012/12/scheda_stranieri.pdf) (19 dicembre 2012).
- Istat. 2012. *Esiti del confronto tra censimento e anagrafe*. Comunicato stampa. Roma. <http://www.istat.it/it/files/2012/12/scheda-confronto-censimento-anagrafe.pdf> (19 dicembre 2012).
- Istat. 2012. *Il processo di produzione dei dati di popolazione legale*. Comunicato stampa. Roma. [http://www.istat.it/it/files/2012/12/nota-metodologica\\_censimento\\_popolazione.pdf](http://www.istat.it/it/files/2012/12/nota-metodologica_censimento_popolazione.pdf) (19 dicembre 2012).
- Livi Bacci M. 1994. *Introduzione alla demografia*. Loescher editore.
- Maggino F. 2005. *L'analisi dei dati nell'indagine statistica*. Vol I e II, Università di Firenze.
- OCDE (2015). *International Migration Database*. <https://stats.oecd.org/Index.aspx?DataSetCode=MIG>
- Preston, S.H., Heuvaline P., Guillot M. 2001. *Demography: Measuring and Modeling Population Processes*. Blackwell publishers.
- Strozza S., Benassi F., Ferrara R., Gallo G. 2014. *La recente evoluzione demografica nei maggiori ambiti urbani italiani e il fondamentale ruolo degli stranieri*. Franco Angeli anno XLV – n. 109