

PREVISIONI DELLA POPOLAZIONE RESIDENTE E DELLE FAMIGLIE | BASE 1/1/2020

Futuro della popolazione: meno residenti, più anziani, famiglie più piccole

Le previsioni sul futuro demografico in Italia restituiscono un potenziale quadro di crisi. La **popolazione residente è in decrescita**: da 59,6 milioni al 1° gennaio 2020 a 58 mln nel 2030, a 54,1 mln nel 2050 e a 47,6 mln nel 2070.

Il **rapporto tra giovani e anziani sarà di 1 a 3 nel 2050** mentre la popolazione in età lavorativa scenderà in 30 anni dal 63,8% al 53,3% del totale.

La crisi demografica sul territorio: entro 10 anni l'81% dei Comuni avrà subito un calo di popolazione, l'87% nel caso di Comuni di zone rurali.

Previsto in crescita il numero di famiglie ma con un numero medio di componenti sempre più piccolo. **Meno coppie con figli, più coppie senza**: entro il 2040 una famiglia su quattro sarà composta da una coppia con figli, più di una su cinque non avrà figli.

50,7 anni

L'età media della popolazione nel 2050

Da 45,7 anni nel 2020.

2048

L'anno in cui i decessi potrebbero doppiare le nascite (784mila contro 391mila)

10,3 milioni

Le persone destinate a vivere sole nel 2040

Da 8,6 milioni nel 2020.

www.istat.it

UFFICIO STAMPA
tel. +39 06 4673.2243/44
ufficiostampa@istat.it

CONTACT CENTRE
tel. +39 06 4673.3102
contact@istat.it



Calo graduale ma continuo della popolazione

Sulla base dello scenario di previsione “mediano” si prevede una decrescita della popolazione residente nel prossimo decennio: da 59,6 milioni al 1° gennaio 2020 (punto base delle previsioni) a 58 milioni nel 2030, con un tasso di variazione medio annuo pari al -2,9‰. Nel medio termine la diminuzione della popolazione risulterebbe più accentuata: da 58 milioni a 54,1 milioni tra il 2030 e il 2050 (tasso di variazione medio annuo pari al -3,4‰) (Prospetto 1).

Nel lungo termine le conseguenze della dinamica demografica prevista sulla popolazione totale si fanno più importanti. Tra il 2050 e il 2070 la popolazione diminuirebbe di ulteriori 6,5 milioni (-6,4‰ in media annua). Sotto tale ipotesi la popolazione totale ammonterebbe a 47,6 milioni nel 2070, conseguendo una perdita complessiva di 12,1 milioni di residenti rispetto a oggi.

Le previsioni demografiche sono, per costruzione, tanto più incerte quanto più ci si allontana dall'anno base. L'evoluzione della popolazione totale rispecchia tale principio già dopo pochi anni di previsione. Nel 2050 il suo intervallo di confidenza al 90% (ovvero che il suo presunto valore cada tra due estremi con probabilità pari al 90%) oscilla tra 51 e 57,5 milioni. Venti anni dopo si è tra 41,1 e 54,9 milioni.

Così, se dal lato più favorevole la popolazione potrebbe subire una perdita di “soli” 4,7 milioni tra il 2020 e il 2070, dall'altro si potrebbe pervenire a un calo di ben 18,6 milioni. Confermando quanto emerso in precedenti studi previsivi, sia dell'Istat sia di istituzioni internazionali (*Eurostat* e *United Nations Population Division*), risulta pressoché certo che la popolazione andrà incontro a una diminuzione. Infatti, sebbene non sia esclusa l'eventualità che la dinamica demografica possa condurre a una popolazione nel 2070 più ampia di quella odierna, la probabilità empirica che ciò accada è minima, risultando pari all'1% (percentuale di casi favorevoli all'evento sul totale delle simulazioni condotte).

La questione investe tutto il territorio, pur con differenze tra Centro-nord e Mezzogiorno. Sempre secondo lo scenario mediano, nel breve termine si prospetta nel Nord (-1,3‰ annuo fino al 2030) e nel Centro (-2,2) una riduzione della popolazione meno importante rispetto al Mezzogiorno (-5,4). Nel periodo intermedio (2030-2050), e ancor più nel lungo termine (2050-2070), tale tendenza si rafforza, con un calo di popolazione in tutte le ripartizioni geografiche ma con più forza in quella meridionale. Nel Nord, in genere meno sfavorito, la riduzione media annua sarà dell'1,4‰ nel 2030-2050 e del 4,3‰ nel 2050-2070, contro -6,9 e -10,3‰ nel Mezzogiorno.

Anche l'evoluzione della popolazione nelle ripartizioni geografiche è contrassegnata da incertezza. Per il Nord è difficile individuare la direzione del cambiamento demografico, se orientato alla crescita, come indicato dal limite superiore dell'intervallo di confidenza al 2070 (28,2 milioni), oppure alla decrescita, guardando all'inferiore (20,8). La sua popolazione finale, quindi, ricade tra due valori rispettivamente al di sotto e al di sopra di quello nell'anno base, per quanto lo scenario mediano (24,3 milioni) indichi come più probabile una sua diminuzione. Centro e Mezzogiorno presentano, invece, ipotesi al 2070 i cui valori massimi (11,2 e 15,4 milioni) sono inferiori a quelli delle rispettive popolazioni iniziali.

PROSPETTO 1. POPOLAZIONE RESIDENTE PER RIPARTIZIONE GEOGRAFICA. SCENARIO MEDIANO E INTERVALLO DI CONFIDENZA AL 90%. Anni 2020-2070, 1° gennaio, dati in milioni (*)

Ripartizione geografica	2020	2030	2040	2050	2070
Nord	27,6	27,3	27,0	26,5	24,3
		[27,0 / 27,5]	[26,2 / 27,9]	[24,9 / 28,3]	[20,8 / 28,3]
Centro	11,8	11,6	11,3	11,0	9,7
		[11,5 / 11,7]	[11,0 / 11,7]	[10,3 / 11,7]	[8,4 / 11,2]
Mezzogiorno	20,2	19,1	18,0	16,7	13,6
		[19,0 / 19,3]	[17,6 / 18,5]	[15,9 / 17,6]	[11,9 / 15,4]
ITALIA	59,6	58,0	56,4	54,1	47,6
		[57,5 / 58,4]	[54,8 / 58,1]	[51,0 / 57,5]	[41,1 / 54,9]

(*) Tra parentesi quadre valori sottostanti gli intervalli di confidenza.

Le future nascite non compenseranno i futuri decessi

Già da diversi anni, precisamente dal 2007, l'Italia sta affrontando la realtà di un ricambio naturale negativo, fattore alla base del processo di riduzione della popolazione, nonostante la parziale contropartita di dinamiche migratorie con l'estero di segno positivo.

Gli scenari previsivi di nascite e decessi ingigantiscono tale processo, misurando la tendenza a registrare annualmente saldi negativi per il movimento naturale della popolazione. Nemmeno negli scenari contrapposti più favorevoli (regimi di alta natalità incrociati con regimi di bassa mortalità) il numero proiettato di nascite arriverebbe a compensare quello dei decessi. Ad esempio, il limite superiore dell'intervallo di confidenza al 90% per le nascite (scenario nel quale il numero medio di figli per donna cresce da 1,24 nell'anno base a 1,88 nel 2070) determinerebbe pur sempre un quantitativo di nascite più basso dei decessi riscontrabile lungo il limite di confidenza inferiore (Figura 1).

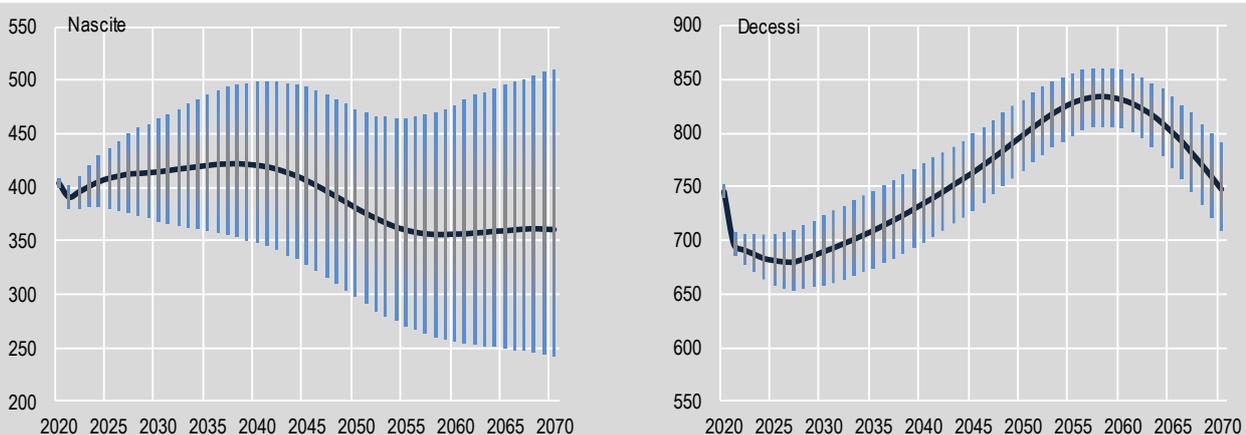
Nello scenario mediano, superato lo shock di breve termine imposto dalla pandemia, le nascite dovrebbero intraprendere un trend di lieve recupero, fino a 414mila nel 2030 e a un massimo di 422mila entro il 2038. Un periodo, questo, nel quale la fecondità viene prevista in rialzo, da 1,24 figli per donna a 1,44. Trend che in seguito, pur prospettato in prosecuzione (fino a 1,55 figli per donna entro il 2070), non produce un rialzo ulteriore delle nascite dopo il 2038, per una motivazione intrinseca alla composizione della popolazione femminile in età feconda. Le donne in età fertile, infatti, via via diminuiranno nonché a invecchiare in media, riducendo il potenziale riproduttivo del Paese.

A loro volta i decessi, riassorbito l'effetto perturbatore della pandemia, dovrebbero proseguire a esprimere la loro connaturata tendenza all'aumento. Tale risultato è in linea con quello che sarà il livello di invecchiamento della popolazione, pur in un contesto di buone aspettative sull'evoluzione della speranza di vita (86,5 e 89,5 anni quella prevista alla nascita nel 2070, rispettivamente per uomini e donne). Da circa 680mila decessi annui intorno al 2025 si transiterebbe così intorno a 800mila nel 2050 in modo pressoché lineare, con un picco di 835mila nel 2058, anno dopo il quale l'abbassamento del numero totale di decessi è in linea con la diminuzione assoluta della popolazione.

Il margine di incertezza legato alle previsioni di nascite e decessi, pur non modificando il quadro di fondo, aumenta nel tempo, soprattutto per le prime. Ciò si deve a due ragioni. L'intervallo di confidenza proiettato per la fecondità è oggettivamente alto, tra 1,23 e 1,88 figli per donna entro il 2070. Esso, quindi, oscilla tra una visione di fecondità simile a quella odierna e una che tende poco sotto il livello di sostituzione delle generazioni. Inoltre, man mano che ci si allontana nel futuro, i livelli di fecondità sono applicati a coorti di donne a loro volta previste.

Le previsioni contemplano, cioè, nati da madri che a loro volta devono ancora nascere. Per i decessi tali valutazioni valgono in parte. Gli intervalli di confidenza proiettati al 2070 per le probabilità di morte sono ampi (tra 84,3 e 88,5 anni la speranza di vita alla nascita degli uomini, tra 87,6 e 91,5 quella delle donne), ma queste impattano nel tempo nei confronti di individui che, al netto dei futuri movimenti migratori con l'estero, sono largamente già in vita.

FIGURA 1. NASCITE E DECESSI IN ITALIA, SCENARIO MEDIANO E INTERVALLO DI CONFIDENZA AL 90%. Anni 2020-2070, dati in migliaia.



Scenario migratorio positivo ma incerto

Superato lo shock pandemico, si ipotizza che le immigrazioni dall'estero possano recuperare i livelli mediamente rilevati nel quinquennio 2015-2019 a partire dall'anno 2023. Da tale momento, con una quota di immigrati intorno alle 280mila unità, lo scenario mediano contempla un sostanziale rientro alla normalità, in concomitanza della ripresa economica e della progressiva attuazione del PNRR (Figura 2). Nel medio e lungo termine si prevede, quindi, una graduale diminuzione degli ingressi fino al valore di 244mila nel 2070. Cumulato sull'intero periodo di previsione, lo scenario mediano prefigura, pertanto, un insediamento a carattere permanente di 13,3 milioni di immigrati.

Anche le emigrazioni per l'estero dovrebbero recuperare nel giro di pochi anni i valori espressi nel quinquennio antecedente l'avvento della pandemia. Nello scenario mediano si presuppone in seguito una loro evoluzione stabile, da circa 145mila uscite annue nel 2025 a 126mila nel 2070. In totale, sull'intero arco di proiezione gli emigrati dall'Italia sarebbero circa 6,9 milioni.

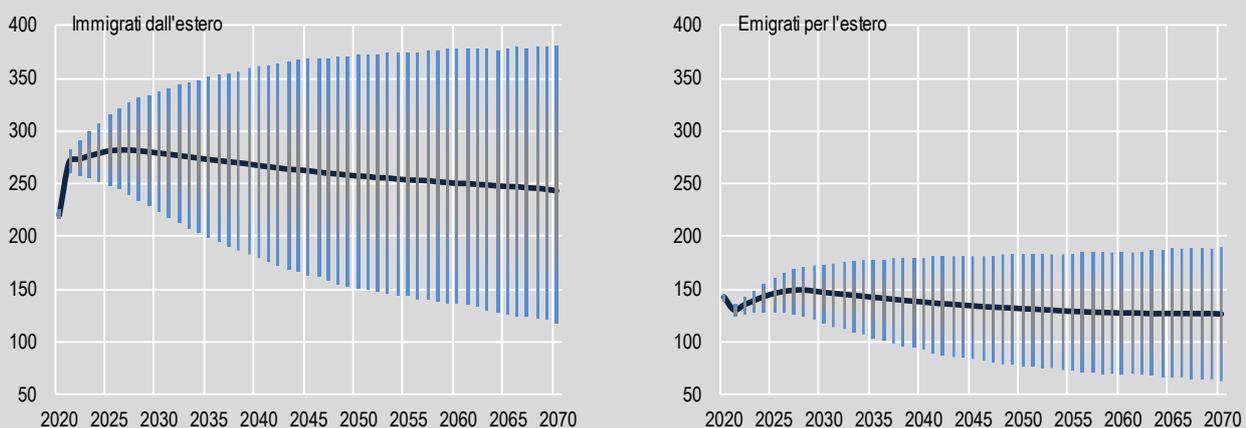
Il saldo migratorio con l'estero risultante in base allo scenario mediano è, pertanto, ampiamente positivo: superato il 2020 (+79mila unità), dal 2021 esso si riporta al valore di +141mila, cui segue una continua e regolare flessione che conduce l'indicatore al valore di +118mila nel 2070.

I flussi migratori con l'estero sono però contrassegnati da profonda incertezza. Le migrazioni internazionali sono governate, da una parte, da normative suscettibili di modifiche, dall'altra, da fattori socio-economici interni ed esterni al Paese di non facile interpretazione. Si pensi, ad esempio, alla pressione migratoria esercitata nei Paesi di origine, alle politiche di integrazione degli immigrati, alla modulazione del mercato del lavoro, all'emigrazione di cittadini residenti in Italia. Tutti questi fattori hanno la potenzialità di dare luogo a scenari migratori assai diversificati.

La quota di immigrati dall'estero, ad esempio, presenta al 2070 un intervallo di confidenza al 90% che varia tra 188mila e 380mila unità. Nel medesimo anno si prevede, inoltre, che il valore delle emigrazioni possa cadere in un intervallo compreso tra 62mila e 189mila unità.

Premesso che l'analisi di risultati a così lungo termine debba necessariamente accompagnarsi a una grande cautela, in realtà ci rivela due possibili fotografie del futuro tra loro molto diverse, persino alternative. Da un lato quella di un Paese molto attrattivo, dall'altro quella di un Paese che potrebbe radicalmente mutare la sua natura di accoglienza per tornare a essere un luogo da cui emigrare. Peraltro, nel quadro di tutte le simulazioni condotte, scaturisce che la probabilità che il Paese possa conseguire un saldo netto con l'estero di segno negativo è tutt'altro che scarsa, per quanto bassa. Tale prospettiva ha infatti una probabilità di realizzarsi che va dall'1,3% nel 2040 al 4,1% nel 2050% e all'8,5% nel 2070.

FIGURA 2. MOVIMENTO MIGRATORIO CON L'ESTERO, SCENARIO MEDIANO E INTERVALLO DI CONFIDENZA AL 90%. Anni 2020-2070, dati in migliaia.



Giovani e anziani in rapporto di uno a tre entro il 2050

Il protratto regime di bassa fecondità che ha caratterizzato l'Italia negli ultimi 40 anni, combinato con i traguardi raggiunti sul versante della sopravvivenza, hanno fatto sì che già oggi si possa parlare di Paese a elevato processo di invecchiamento. La popolazione di 65 anni e più rappresenta il 23,2% del totale, quella fino a 14 anni di età il 13%, quella nella fascia 15-64 anni il 63,8%, mentre l'età media si è avvicinata al traguardo dei 46 anni. Cosicché, è certo che le future prospettive saranno in gran parte governate dall'attuale articolazione per età della popolazione, e solo in parte minore dai cambiamenti immaginati circa l'evoluzione della fecondità, della mortalità e delle dinamiche migratorie. In base a un rapporto di importanza, all'incirca, di due terzi e un terzo rispettivamente.

Entro il 2050 le persone di 65 anni e più potrebbero rappresentare il 35% del totale secondo lo scenario mediano, mentre l'intervallo di confidenza al 90% presenta un campo di variazione compreso tra un minimo del 33,1% e un massimo del 36,9%. Comunque vada sarà pertanto necessario adattare ancor più le politiche di protezione sociale a una quota così crescente di popolazione anziana.

I giovani fino a 14 anni di età, sebbene nello scenario mediano si preveda una fecondità in recupero, potrebbero rappresentare entro il 2050 l'11,7% del totale, registrando quindi una lieve flessione. Rimane aperta, tuttavia, la questione che a tale data il rapporto tra ultrasessantacinquenni e ragazzi risulterà in misura di 3 a 1.

L'impatto dell'invecchiamento va evidenziato anche nei confronti della probabile evoluzione della popolazione in età lavorativa. Nei prossimi trent'anni, infatti, la popolazione di 15-64 anni è la componente che sarà più soggetta a repentina variazione (scenderebbe dal 63,8% al 53,3% in base allo scenario mediano), con una forchetta potenziale compresa tra il 51,9% e il 54,7%. Come nel caso della popolazione anziana, quindi, anche qui un quadro evolutivo certo del quale non solo vanno valutati gli effetti sul mercato del lavoro e sulla programmazione economica futura, ma anche la pressione che il Paese dovrà affrontare nel cercare di mantenere l'attuale livello di *welfare*.

Un parziale riequilibrio nella struttura della popolazione, anche se nel lungo termine, potrebbe aversi via via che le generazioni nate negli anni del *baby boom* (nati negli anni '60 e nella prima metà dei '70) tenderanno a estinguersi. Tali generazioni oggi occupano le tarde età adulte e si accingono a transitare integralmente tra gli over65 nel giro di venti anni. Secondo lo scenario mediano i 15-64enni potrebbero perciò riportarsi al 54,1% entro il 2070 mentre gli ultrasessantacinquenni ridiscendere al 34,3%. Stabile, invece, la popolazione giovanile con un livello dell'11,6%.

La trasformazione della struttura per età della popolazione caratterizzerà ogni area del Paese, sebbene il Mezzogiorno sia contraddistinto da un processo di invecchiamento progressivamente più marcato (Prospetto 2). In tale ripartizione, che oggi presenta ancora un profilo strutturale più giovane, l'età media transita da 44,6 anni nel 2020 a 50 anni nel 2040, sopravanzando a quel punto il Nord che raggiunge un'età media di 49,2 anni ma che nell'anno base parte da un livello più alto, ossia 46,3 anni.

PROSPETTO 2. ETÀ MEDIA DELLA POPOLAZIONE RESIDENTE PER RIPARTIZIONE GEOGRAFICA, SCENARIO MEDIANO E INTERVALLO DI CONFIDENZA AL 90%. Anni 2020-2070, 1° gennaio, in anni e decimi di anno (*).

Ripartizione geografica	2020	2030	2040	2050	2070
Nord	46,3	48,0	49,2	49,9	49,7
		[47,8 / 48,3]	[48,4 / 50,0]	[48,5 / 51,4]	[47,1 / 52,5]
Centro	46,4	48,6	50,3	51,3	51,1
		[48,4 / 48,8]	[49,5 / 51,0]	[49,8 / 52,7]	[48,5 / 53,9]
Mezzogiorno	44,6	47,6	50,0	51,6	52,1
		[47,3 / 47,8]	[49,3 / 50,7]	[50,2 / 53,0]	[49,4 / 54,9]
ITALIA	45,7	48,0	49,7	50,7	50,7
		[47,7 / 48,2]	[48,9 / 50,5]	[49,3 / 52,2]	[48,1 / 53,5]

(*) Tra parentesi quadre valori sottostanti gli intervalli di confidenza.

Inevitabile il declino demografico nel Mezzogiorno

La traiettoria di crescita dell'invecchiamento nella ripartizione si estende anche negli anni successivi, fino a completare dal 2050 il sorpasso nei confronti del Centro. A quella data, infatti, l'età media nel Mezzogiorno avrebbe raggiunto i 51,6 anni contro i 51,3 del Centro, sebbene anche quest'ultima ripartizione parta da un livello di invecchiamento oggi decisamente più elevato, ossia 46,4 anni. Nonostante la presenza di un significativo margine di incertezza, che al 2070 comincia a essere importante per un indicatore come l'età media, il Mezzogiorno rallenterebbe ma non fermerebbe il suo percorso raggiungendo un'età media della popolazione superiore ai 52 anni. A quel punto, invece, sia il Nord (49,7 anni) sia il Centro (51,1) avrebbero già avviato da tempo il percorso contrario, ossia quello verso una struttura per età almeno in piccola parte ringiovanita.

Entro il 2030 la maggior parte dei Comuni in calo demografico ⁱ⁾

Entro 10 anni (periodo per il quale l'Istat definisce per la prima volta nella sua storia un'apposita metodologia previsiva, coerente col livello regionale e nazionale sopra descritto) un numero crescente di Comuni andrà incontro a un calo demografico. Secondo lo scenario mediano, l'81% dovrebbe trovarsi in tali condizioni entro il 2030. Ciò si deve alla bassa fecondità, che colpisce uniformemente alla base la struttura per età delle popolazioni, ma anche a livelli migratori sfavorevoli per alcune realtà territoriali, laddove è più forte sia l'emigrazione per l'estero sia quella per l'interno.

A livello nazionale si valuta che tra il 2020 e il 2030 i Comuni delle zone rurali possano nel complesso contrarre una riduzione della popolazione pari al 6%, passando da 10,2 a 9,6 milioni di residenti (Prospetto 3). In tali aree i Comuni con saldo negativo della popolazione sono l'87% del totale. La questione investe soprattutto le aree del Mezzogiorno, dove i Comuni delle zone rurali con bilancio negativo sono il 93% del totale e dove si riscontra una riduzione della popolazione pari all'8,9%.

Per i 1.060 Comuni che ricadono in *Aree interne*, particolari zone del territorio nazionale che si contraddistinguono per la distanza fisica dall'offerta di servizi essenziali, la condizione demografica appare ancor più sfavorevole. Qui, infatti, la quota di Comuni con saldo negativo della popolazione nel decennio sale al 95%, facendo nel complesso registrare una riduzione della popolazione pari al 9,6% (del 10,4% considerando il solo Mezzogiorno).

In una situazione relativamente migliore si collocano i Comuni a densità intermedia (piccole città e sobborghi), dove il calo demografico atteso è del 2,2% (la popolazione transita nel decennio da 28,4 a 27,8 milioni). Minore è anche la quota di Comuni interessati al calo demografico, il 72% del totale, che tuttavia sale all'85% nel solo Mezzogiorno. Infine, sebbene a un livello minore, anche Città e Zone densamente popolate saranno interessate dal tema dello spopolamento. La capacità attrattiva delle aree a più forte urbanizzazione farà sì che nel decennio il calo complessivo della popolazione sia solo del 2,1%, con il 67% dei Comuni destinati a subire un saldo negativo tra i propri residenti.

PROSPETTO 3. POPOLAZIONE RESIDENTE PER GRADO DI URBANIZZAZIONE / AREA INTERNA E RIPARTIZIONE GEOGRAFICA. Anni 2020 e 2030 1° gennaio, scenario mediano, dati in migliaia *

Ripartizione geografica	Anno	Città o Zone densamente popolate	Piccole città e sobborghi, Zone a densità intermedia	Zone rurali o scarsamente popolate	Aree interne	Non aree interne
Nord	2020	9.392	13.370	4.854	501	27.115
	2030	9.369	13.231	4.656	462	26.794
Centro	2020	4.514	5.388	1.929	382	11.449
	2030	4.448	5.303	1.817	346	11.222
Mezzogiorno	2020	7.141	9.631	3.423	1.092	19.102
	2030	6.796	9.217	3.118	978	18.153
ITALIA	2020	21.047	28.388	10.206	1.975	57.667
	2030	20.613	27.752	9.590	1.786	56.169

(*) Il grado di urbanizzazione risponde alla classificazione dei Comuni *Degurba/Statistiche sulla popolazione per griglia regolare* (www.istat.it/it/archivio/155162). Le aree interne, definite dall'Agenzia per la Coesione Sociale, sono zone del territorio italiano caratterizzate da significativa distanza dai centri di offerta di servizi essenziali. Comprendono 1.060 Comuni (www.agenziacoesione.gov.it).

Le migrazioni interne una delle chiavi di lettura dello spopolamento

Con riferimento alle migrazioni interne nel periodo 2020-2029, cumulativamente, si prevede che i trasferimenti di residenza possano interessare 13,1 milioni di cittadini, il 24% dei quali a carattere interregionale mentre per il restante 76% tra Comuni all'interno di una medesima Regione.

Città e zone densamente popolate saranno favorite da saldi migratori interregionali positivi (+108mila unità nel decennio) a danno delle zone a densità intermedia (-59mila) e delle zone rurali (-49mila). Lo schema, tuttavia, ha valenza solo per i Comuni nel loro insieme. Infatti, caratteristica delle migrazioni a lungo raggio è vedere favorite le aree del Centro-nord, dove ad avere bilanci migratori positivi sono soprattutto Città e zone densamente popolate (+312mila) ma anche zone a densità intermedia (+166mila) e persino zone rurali (+37mila). Ciò avverrebbe a danno delle aree del Mezzogiorno, dove le perdite di popolazione ammonterebbero, rispettivamente, a 204mila, 226mila e 86mila unità.

Il tema dello spopolamento delle zone rurali va visto nella logica delle migrazioni a breve raggio (intraregionali). Da questo punto di vista emerge la forza attrattiva dei piccoli centri che mostrano saldi migratori positivi (+203mila nel decennio) rispetto alle zone rurali (-54mila) e alle grandi Città (-149mila).

La prospettiva dei prossimi anni sarà dunque vedere i grandi Centri urbani, soprattutto del Centro-nord, continuare a esercitare capacità attrattiva dalle zone rurali più remote, soprattutto se del Mezzogiorno. Guardando invece alle dinamiche migratorie di breve raggio, i grandi Centri urbani non costituiscono affatto la meta di destinazione preferita che, al contrario, si dimostra essere il piccolo centro, spesso alla periferia o nelle vicinanze dello stesso grande Centro urbano.

Famiglie in aumento ma frammentate

Prosegue l'aumento del numero di famiglie (quasi un milione di unità in più): da 25,7 milioni nel 2020 si stima una crescita fino a 26,6 milioni nel 2040 (+3,5 punti percentuali), ma cresce la loro frammentazione (Prospetto 4). Le famiglie senza nuclei passano da 9,2 a 11 milioni (+20%). Quelle con almeno un nucleo, contraddistinte dalla presenza di almeno una relazione di coppia o di tipo genitore-figlio, seguono invece una tendenza opposta, diminuendo da 16,6 a 15,6 milioni (-6%).

Il calo delle famiglie con nuclei si deve alle conseguenze delle dinamiche socio-demografiche di lungo periodo: l'invecchiamento della popolazione, l'incremento dell'instabilità coniugale, la bassa natalità. L'aumento della speranza di vita genera un maggior numero di persone sole. Il calo della natalità incrementa le persone senza figli mentre l'aumento dell'instabilità coniugale accresce il numero di persone che vivono sole o di genitori soli in seguito allo scioglimento di un legame di coppia.

Si assottiglia la dimensione media familiare, che si prevede possa scendere da 2,3 componenti nel 2020 a 2,1 nel 2040. Prendendo in esame le sole famiglie con presenza di nuclei la dimensione media risulterebbe variata nel medesimo termine temporale da 3 a 2,8 componenti.

PROSPETTO 4. NUMERO DI FAMIGLIE PER TIPOLOGIA E NUMERO MEDIO DI COMPONENTI FAMILIARI. Anni 2020*, 2030, 2040, scenario mediano, valori in migliaia e medi.

TIPOLOGIA	2020	2030	2040	TIPOLOGIA/INDICATORE	2020	2030	2040
Numero totale di famiglie	25.737	26.224	26.628	Coppie senza figli	5.101	5.498	5.739
Famiglie con nuclei	16.569	16.182	15.585	Genitore solo maschio	587	759	892
Famiglie senza nuclei	9.169	10.042	11.044	Genitore solo femmina	2.204	2.202	2.211
Persone sole maschi	3.619	3.892	4.252	Altro tipo di famiglia **	1.007	1.058	1.093
Persone sole femmine	4.953	5.489	6.079	Numero medio di componenti	2,3	2,2	2,1
Coppie con figli	8.267	7.325	6.362	Num. medio di componenti per famiglie con almeno un nucleo	3,0	2,9	2,8

(*) I dati ufficiali di indagine sugli Aspetti della vita quotidiana sono normalmente diffusi in media biennale. Qui, invece, i dati si riferiscono al 1° gennaio dell'anno indicato. Per il 2020 ciò può dare luogo a delle differenze.

(**) famiglie multipersonali (formate da più persone che non costituiscono nucleo) e famiglie con due o più nuclei.

Famiglie più piccole e nuovi ruoli familiari

La dinamica demografica, sia del passato sia in chiave prospettica, comporta una riduzione delle nuove generazioni, sia in termini assoluti che relativi. La struttura per età della popolazione evidenzia, già oggi, un elevato squilibrio a favore delle generazioni più anziane e non ci sono al momento fattori che possano far pensare a inversioni di rotta. Dalle previsioni demografiche appare poco probabile una svolta nel numero delle nascite negli anni a venire, pur a fronte di ipotesi favorevoli nei confronti della propensione media di riproduttività da parte delle coppie. Ciò in quanto la prospettiva di avere a che fare con un numero decrescente di donne in età fertile, da un lato, e la tendenza a posticipare la genitorialità dall'altro, sembrano assumere un peso crescente.

L'analisi della popolazione prevista al 2040 in base al ruolo in famiglia evidenzia in modalità congiunta il processo di invecchiamento e i cambiamenti nelle posizioni familiari. In particolare, essa mostra la diminuzione delle persone in coppia con figli, l'aumento di quelle senza figli e delle persone che vivono sole, queste ultime soprattutto se anziane (Figura 3). Le classi di età più giovani si assottigliano in consistenza ma la posizione familiare in qualità di figlio/a rimane prevalente fino ai 30 anni, a causa della prolungata permanenza dei giovani nella famiglia di origine.

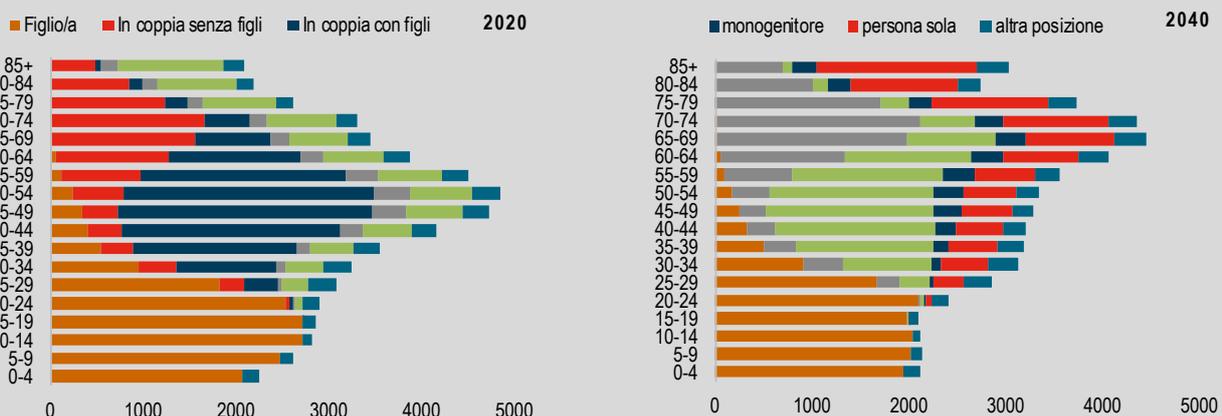
Entro il 2040 oltre 10 milioni di persone sole

Tra le molteplici dinamiche che conducono alla trasformazione delle strutture familiari, vi è quella che porta all'incremento delle persone sole, delle vere e proprie micro-famiglie cui si deve principalmente la crescita assoluta del numero totale di famiglie.

Gli uomini che vivono da soli passeranno da 3,6 milioni nel 2020 a 4,3 milioni nel 2040 (+17%). Le donne sole sarebbero invece destinate ad aumentare da 5 a 6,1 milioni (+23%). Questo tipo di famiglie, in ragione soprattutto della loro composizione per età, ha una ricaduta sociale importante: è, infatti, principalmente nelle età avanzate che aumentano molto le persone sole.

Fra gli individui di 65 anni e più, nel ventennio in esame si riscontrerebbe un aumento di ben 640mila uomini soli e di 1,2 milioni di donne sole. L'aumento della sopravvivenza delle persone anziane e, tra queste, di quelle sole, potrebbe comportare un futuro aumento dei fabbisogni di assistenza. Ma l'aumento delle persone sole anziane comporta anche risvolti positivi; l'aumento della sopravvivenza in quantità di anni di vita vissuti e, si presuppone, anche in qualità della vita, potrebbe consentire a queste persone di svolgere un ruolo attivo nella società. Ad esempio, come già accade oggi e più verosimilmente un domani, supportando le famiglie dei propri figli nella cura dei nipoti e garantendo loro sostegno economico, partecipando al ciclo economico non solo come consumatori di servizi assistenziali ma anche come investitori di capitali.

FIGURA 3. POPOLAZIONE PER POSIZIONE FAMILIARE E CLASSI QUINQUENNALI DI ETÀ. Anni 2020 e 2040, scenario mediano, valori in migliaia



In calo le coppie con figli

Per via dei livelli di fecondità riscontrati negli ultimi anni, nonché sulle base delle ipotesi prodotte nello scenario mediano su questo tema, si prevede una consistente diminuzione delle coppie con figli. Tra il 2020 e il 2040 la loro consistenza diminuirebbe di ben il 23%, ossia da 8,3 milioni fino a 6,4 milioni. Nello stesso frangente temporale, le coppie senza figli aumenteranno da 5,1 a 5,7 milioni, per un incremento del 13%. Se tale andamento dovesse procedere con la stessa intensità prevista fino al 2040, soprattutto per quel che riguarda il ritmo di discesa delle coppie con figli, il sorpasso ai danni di queste ultime da parte delle coppie senza figli potrebbe avvenire già entro il 2045.

In crescita i padri soli, stabili le madri sole

L'instabilità coniugale, sempre più diffusa nel Paese, vedrà aumentare le famiglie composte da un genitore solo, maschio o femmina, con uno o più figli. Nel 2020 i monogenitori sono in totale 2,8 milioni, per lo più madri (2,2 milioni) che padri (meno di 600mila) e rispettivamente rappresentano l'8,6% e il 2,3% del totale delle famiglie.

Se in passato, a seguito di uno scioglimento della coppia, i figli (soprattutto se piccoli) venivano generalmente affidati alle madri, dalla promulgazione della legge sull'affido congiunto del 2006 questa prevalenza è andata diminuendo. Ciò ha determinato una sempre maggiore diffusione di padri in qualità di genitori affidatari nelle sentenze di separazione o divorzio.

È in questa chiave che va letta l'evoluzione della curva dei padri soli i quali, pur rimanendo minoritari rispetto alle madri sole, saranno circa 900mila entro il 2040 (il 3,4% del totale delle famiglie). In tale anno le madri sole risulterebbero invece numericamente invariate a 2,2 milioni (8,3% del totale), cosicché il totale di monogenitori sarebbe pari a 3,1 milioni.

La struttura familiare del Mezzogiorno converge a quella del Nord

Le tipologie familiari rispondono a dinamiche demografiche e comportamenti sociali ben precisi e differenti tra le diverse zone del Paese, in particolar modo tra Nord e Mezzogiorno. Al Nord, nel 2020, la quota di famiglie con almeno un nucleo è decisamente più bassa, precisamente del 62,8% contro il 67,7% del Mezzogiorno (Prospetto 5). Tuttavia, sotto questo profilo le previsioni evidenziano una parziale tendenza alla convergenza delle due aree.

Nel Mezzogiorno si prevede un cambiamento più consistente di questo tipo di famiglie: nel 2040 potrebbero costituire il 61% delle famiglie totali, ovvero registrerebbero una riduzione di circa 7 punti percentuali rispetto al 2020. Nel Nord le famiglie con almeno un nucleo avrebbero una riduzione relativa inferiore, pari a 5 punti percentuali, arrivando a costituire il 57,5% delle famiglie totali.

PROSPETTO 5. FAMIGLIE PER TIPOLOGIA E RIPARTIZIONE GEOGRAFICA. Anni 2020, 2030, 2040, scenario mediano, valori percentuali.

Tipologia familiare	Nord			Centro			Mezzogiorno			Italia		
	2020	2030	2040	2020	2030	2040	2020	2030	2040	2020	2030	2040
Persone sole maschi	14,7	15,8	17,0	15,0	15,6	16,7	12,5	13,0	13,9	14,1	14,8	16,0
Persone sole femmine	20,3	21,5	22,9	19,9	21,7	23,4	17,3	19,6	22,4	19,2	20,9	22,8
Coppie senza figli	21,8	22,7	23,1	18,2	19,2	19,6	17,9	19,5	20,4	19,8	21,0	21,6
Coppie con figli	30,0	26,2	22,8	29,7	25,7	21,9	36,9	32,0	26,9	32,1	27,9	23,9
Genitori soli maschi	2,2	2,8	3,3	2,6	3,3	3,7	2,2	2,8	3,2	2,3	2,9	3,4
Genitori soli femmine	7,5	7,3	7,1	10,3	10,1	10,3	9,0	8,9	8,9	8,6	8,4	8,3
Altro tipo di famiglia	3,5	3,7	3,8	4,3	4,5	4,6	4,3	4,3	4,3	3,9	4,0	4,1
Famiglie con nuclei	62,8	60,3	57,5	62,9	60,3	57,2	67,7	64,8	61,0	64,4	61,7	58,5
Famiglie senza nuclei	37,2	39,7	42,5	37,1	39,7	42,8	32,3	35,2	39,0	35,6	38,3	41,5
Totale	100,0											

In calo le coppie con figli soprattutto nel Mezzogiorno

La tipologia familiare che si prevede possa subire il cambiamento più evidente nei prossimi vent'anni è quella delle coppie con figli. Se queste oggi costituiscono il 32,1% delle famiglie totali, nel 2040 passerebbero al 23,9%. La diminuzione delle coppie con figli sarebbe più accentuata nel Mezzogiorno, pari a 10 punti percentuali (dal 36,9% al 26,9%).

Il *gap* di genere nella sopravvivenza si rispecchia nell'aumento di donne che vivono sole. Per l'Italia in complesso si prevede che questa tipologia familiare possa andare a costituire il 22,8% delle famiglie totali entro il 2040, da un valore attuale del 19,2%, per una variazione di circa quattro punti percentuali. Una variazione che sale a oltre cinque punti percentuali per il Mezzogiorno (dal 17,3% al 22,4%).

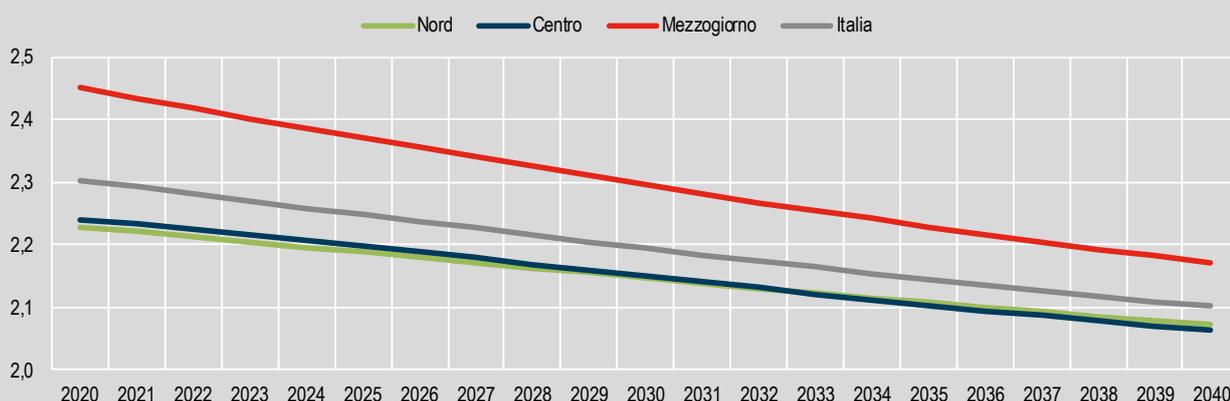
Nel caso degli uomini, la diversità territoriale in termini di speranza di vita, ma anche una loro più evidente predisposizione alle "seconde nozze", fanno sì che anche l'aumento di persone sole abbia delle specificità territoriali. Nel 2040 questa tipologia familiare costituirebbe il 16% di tutte le famiglie a livello nazionale, il 17% nel Nord.

La tipologia familiare della coppia senza figli, che si prevede possa costituire il 21,6% delle famiglie totali nel 2040, continuerà a essere più diffusa al Nord (23,1%), pur conseguendo un incremento più contenuto. Il cambiamento più forte è infatti previsto nel Mezzogiorno, dove, a fronte di una situazione iniziale meno diffusa, le coppie senza figli incrementerebbero in venti anni di circa tre punti percentuali (dal 17,9% al 20,4%).

Il combinato disposto delle trasformazioni familiari in atto e in prospettiva, fa sì che la dimensione familiare media continuerà a scendere, non solo sul piano nazionale (da 2,3 a 2,1 componenti), ma anche seguendo le specificità demografiche e sociali del territorio.

Il Nord e il Centro, con valori attuali e traiettorie future assai simili, perverranno a un valore medio di componenti poco al di sotto del dato nazionale. Il Mezzogiorno, grazie a tassi di fecondità più elevati nel non più recente passato, aveva sempre evidenziato famiglie mediamente più numerose rispetto al Nord. Oggi, con livelli riproduttivi più contenuti anche nel Mezzogiorno, questo primato (2,5 componenti) tende a farsi meno netto. In futuro, per quanto si preveda possa mantenerlo fino al 2040, l'aspettativa è per un'ulteriore diminuzione fino a 2,2 componenti (Figura 4).

FIGURA 4. NUMERO MEDIO DI COMPONENTI FAMILIARI PER RIPARTIZIONE GEOGRAFICA. Anni 2020-2040
scenario mediano



Glossario

Campo di variazione (*range*): misura della variabilità di un fenomeno quantitativo definita dalla differenza tra il valore massimo e il valore minimo osservato.

Coppia: due persone legate da relazione affettiva e sentimentale. Può essere formata da persone di sesso opposto o dello stesso sesso. I vincoli tra persone in coppia possono essere formali (coppia *de iure*: coniugati, uniti civilmente o conviventi di fatto ai sensi della legge 76/2016) o informali (coppia *de facto*).

Crescita naturale (tasso di): differenza tra il tasso di natalità e il tasso di mortalità.

Crescita totale (tasso di): somma del tasso migratorio netto totale e del tasso di crescita naturale.

Coorti componenti (modello): algoritmo di calcolo continuo che in modalità iterativa simula l'evoluzione dell'equazione fondamentale della popolazione per classi di età, consentendo di determinare le poste demografiche oggetto di previsione e di ottenere la popolazione superstite alla fine di ogni anno.

Decesso: cessazione di ogni segno di vita in un qualsiasi momento successivo alla nascita vitale.

Dipendenza anziani (indice di): rapporto tra la popolazione di 65 anni e più e la popolazione di 15-64 anni, moltiplicato per 100.

Dipendenza strutturale (indice di): rapporto tra la popolazione in età non attiva (0-14 anni e 65 anni e più) e la popolazione in età attiva (15-64 anni), moltiplicato per 100.

Emigratorio interno (tasso): rapporto tra il numero di cancellati per l'interno e l'ammontare medio della popolazione residente, moltiplicato per 1.000.

Emigratorio per l'estero (tasso): rapporto tra il numero di cancellati per l'estero e l'ammontare medio della popolazione residente, moltiplicato per 1.000.

Età media: età media della popolazione detenuta a una certa data espressa in anni e decimi di anno.

Età media al parto: età media al parto delle madri espressa in anni e decimi di anno, calcolata considerando i soli nati vivi.

Famiglia: insieme di persone legate da vincoli di matrimonio, parentela, affinità, adozione, tutela, affettivi, coabitanti e aventi dimora abituale nello stesso Comune. Può essere costituita anche da una sola persona.

Famiglie con nuclei: comprende le coppie con figli, le coppie senza figli, i genitori soli, le famiglie con due o più nuclei.

Famiglie senza nuclei: comprende famiglie formate da una sola persona e le famiglie multipersonali, ossia che non costituiscono un nucleo familiare anche se composte da più persone.

Fecondità per età (tasso specifico di): il rapporto tra il numero di nati vivi da donne di età compresa tra x e $x+1$ e il numero medio di donne residenti di tali età in un dato anno.

Immigratorio dall'estero (tasso): rapporto tra il numero di iscritti dall'estero e l'ammontare medio della popolazione residente, moltiplicato per 1.000.

Immigratorio interno (tasso): rapporto tra il numero di iscritti dall'interno e l'ammontare medio della popolazione residente, moltiplicato per 1.000.

Intervallo predittivo (o di confidenza): intervallo associato a una variabile casuale ancora da osservare, con una specifica probabilità che la variabile casuale vi ricada all'interno.

Iscrizione e cancellazione anagrafica per trasferimento di residenza: iscrizione riguarda le persone trasferitesi in un Comune da altri Comuni o dall'estero; la cancellazione riguarda le persone trasferitesi in altro Comune o all'estero.

Migratorio netto con l'estero (tasso): differenza tra il tasso immigratorio dall'estero e il tasso emigratorio con l'estero.

Migratorio netto con l'interno (tasso): differenza tra il tasso immigratorio dall'interno e il tasso emigratorio per l'interno.

Migratorio netto totale (tasso): somma del tasso migratorio netto con l'interno e del tasso migratorio netto con l'estero.

Mortalità (tasso di): rapporto tra il numero dei decessi nell'anno e l'ammontare medio della popolazione residente, moltiplicato per 1.000.

Nato vivo: prodotto del concepimento che, una volta espulso o completamente estratto dal corpo materno, indipendentemente dalla durata della gestazione, respiri o manifesti altro segno di vita.

Natalità (tasso di): rapporto tra il numero dei nati vivi nell'anno e l'ammontare medio della popolazione residente, moltiplicato per 1.000.

Nucleo familiare: insieme delle persone che formano una relazione di coppia o di tipo genitore-figlio. Si intende la coppia coniugata, unita civilmente o convivente, senza figli o con figli celibi o nubili, o anche un solo genitore assieme ad uno o più figli mai sposati. Nell'ambito di una famiglia possono esistere uno o più nuclei familiari, ma può anche non esservene nessuno, come nel caso delle famiglie formate da un membro isolato (famiglie monocomponenti) o più membri isolati (altre persone residenti).

Numero medio di figli per donna: numero di figli che una donna metterebbe al mondo nel caso in cui, nel corso nella propria vita riproduttiva, fosse sottoposta al calendario di fecondità (sotto forma di tassi specifici di fecondità per età) dell'anno di osservazione.

Previsione: sviluppo atteso nel futuro.

Previsione demografica: elaborazione che mostra lo sviluppo futuro di una popolazione quando vengono assunte determinate ipotesi riguardo al futuro corso della mortalità, della fecondità e della migratorietà.

Previsione demografica deterministica: elaborazione sul futuro sviluppo di una popolazione, riassumibile in una singola serie di valori ottenuti da un singolo set di ipotesi demografiche, che non riporta alcuna misura riguardo all'incertezza usualmente associabile ai risultati.

Previsione demografica probabilistica: elaborazione sul futuro sviluppo di una popolazione, riassumibile in un set di valori o in una distribuzione di probabilità, in cui le variabili utilizzate sono di natura casuale che non possono essere previste con certezza e in cui non tutte le ipotesi sono egualmente probabili.

Popolazione residente: costituita in ciascun Comune (e analogamente per altre ripartizioni territoriali) delle persone aventi dimora abituale nel Comune stesso. Non cessano di appartenere alla popolazione residente le persone temporaneamente dimoranti, in altro Comune o all'estero, per l'esercizio di occupazioni stagionali o per causa di durata limitata.

Probabilità (prospettiva) di migrazione interregionale: probabilità che un individuo di età x (in anni compiuti al 1° gennaio) sposti residenza tra due regioni prima che si concluda l'anno.

Probabilità (classica) di morte: probabilità che un individuo di età precisa x muoia prima del compimento del compleanno $x+1$.

Probabilità (prospettiva) di morte: probabilità che un individuo di età x (in anni compiuti al 1° gennaio) non sopravviva entro l'anno.

Saldo migratorio con l'estero: differenza tra il numero degli iscritti per trasferimento di residenza dall'estero e il numero dei cancellati per trasferimento di residenza all'estero.

Saldo migratorio interno: differenza tra il numero degli iscritti per trasferimento di residenza da altro Comune e il numero dei cancellati per trasferimento di residenza in altro Comune.

Saldo migratorio totale: somma del saldo migratorio con l'estero e del saldo migratorio interno.

Saldo naturale (o dinamica naturale): differenza tra il numero di nascite e il numero di decessi.

Saldo totale: somma del saldo naturale e del saldo migratorio totale.

Scenario previsivo: descrizione del contesto, anche concettuale, nel quale la popolazione viene proiettata. In un approccio deterministico normalmente si riferisce all'ipotesi definita principale o centrale. In uno stocastico può riferirsi all'ipotesi identificata come media o mediana.

Simulazione: implementazione quantitativa di un singolo set di ipotesi demografiche da lanciare nel modello coorti-componenti al fine di ottenere un singolo set di previsioni demografiche.

Speranza di vita alla nascita (o vita media): numero medio di anni che una persona può contare di vivere dalla nascita nell'ipotesi in cui, nel corso della propria esistenza, fosse sottoposta ai rischi di mortalità per età dell'anno di osservazione.

Speranza di vita all'età "x": numero medio di anni che una persona di età compiuta "x" può contare di sopravvivere nell'ipotesi in cui, nel corso della successiva esistenza, fosse sottoposta ai rischi di mortalità per età (dall'età "x" in su) dell'anno di osservazione.

Tipologia familiare: classificazione in base alla presenza o meno di almeno un nucleo familiare e per tipo di nucleo.

Vecchiaia (indice di): rapporto tra la popolazione di 65 anni e più e la popolazione di età 0-14 anni, moltiplicato per 100.

Nota metodologica

1) Previsioni della popolazione residente per sesso, età e regione. Anni 2020-2070

Le previsioni demografiche regionali dell'Istat sono costruite con l'obiettivo di rappresentare il possibile andamento futuro della popolazione, sia in termini di numerosità totale sia in termini di struttura per età e sesso. Le informazioni prodotte rappresentano uno strumento importante a supporto delle decisioni nelle politiche economiche e sociali, come quelle relative ai sistemi pensionistici, sanitari, scolastici e abitativi. Le previsioni sono aggiornate periodicamente riformulando le ipotesi evolutive sottostanti la fecondità, la sopravvivenza, i movimenti migratori internazionali e quelli interni.

Le previsioni in base 1° gennaio 2020 sostituiscono quelle in base 2018 pubblicate dall'Istat nell'ottobre 2019. Titolare e responsabile della produzione e della diffusione delle previsioni è l'Istat, come documentato nel Programma statistico nazionale. L'impianto metodologico alla base delle correnti previsioni è il medesimo di quello implementato nel precedente ciclo triennale, che diede luogo al rilascio in successione delle previsioni in base 2016, 2017 e 2018. Tale metodologia venne definita, tra il 2009 ed il 2015, da un gruppo di lavoro che vide la collaborazione di ricercatori dell'Istat e dell'Università Luigi Bocconi di Milano.

L'approccio metodologico, attorno al quale ruota tutto il modello previsivo, è di tipo semi-probabilistico. La caratteristica fondamentale delle previsioni probabilistiche è quella di considerare l'incertezza associata ai valori previsti, determinando gli intervalli di confidenza delle variabili demografiche e dando la possibilità all'utente di poter scegliere autonomamente il grado di fiducia da assegnare ai risultati.

Rispetto all'approccio "deterministico", più largamente utilizzato su scala internazionale e in passato adottato anche dall'Istat (fino alle previsioni in base 2011), si tratta di un avanzamento metodologico significativo. Infatti, nel modello previsivo deterministico l'utente finale non dispone di misure di probabilità associate ai risultati. Cosicché, ulteriore vantaggio del metodo probabilistico è costituito dal fatto che l'utente può cessare di confidare acriticamente sul lavoro dei *projection makers*, che con le varianti "basso/alto" tipiche dell'approccio a scenari deterministici definiscono a priori i confini alternativi alla variante ritenuta "più probabile", generalmente identificata come "scenario centrale".

La quantificazione dell'incertezza non rappresenta peraltro l'unico vantaggio del modello probabilistico. Se ne individua anche un altro che si deve alla più efficace rappresentazione della potenziale evoluzione di una popolazione. Nel modello probabilistico, infatti, gli scenari definibili sono infiniti sul piano teorico (per quanto nella realtà, come si vedrà più avanti, se ne selezionano sempre un numero finito), per cui è possibile che ipotesi di bassa sopravvivenza si mescolino con ipotesi di alta fecondità o medio livello delle migrazioni, o il contrario. Invece, le ipotesi degli scenari alto/basso dell'approccio deterministico sono definite perseguendo una logica *output oriented*: lo scenario alto contempla ipotesi di massimo incremento della sopravvivenza, della fecondità e delle migrazioni, mentre, all'opposto, lo scenario basso contempla solo ipotesi di minimo. La costruzione di tali scenari contrapposti coglie, in effetti, l'obiettivo di determinare un futuro campo di variazione per la popolazione e le sue componenti strutturali, ma si fonda su ipotesi concomitanti che hanno una scarsa possibilità di verificarsi.

Le sezioni successive della nota, oltre a contenere informazioni di carattere generale, illustrano in maniera sintetica i passaggi che hanno reso possibile la costruzione delle previsioni. Tali sezioni includono informazioni sui seguenti aspetti:

- popolazione base
- tecnica di proiezione
- periodo di previsione
- panel di esperti
- questionario degli esperti e modello probabilistico
- relazione tra le previsioni nazionali e regionali
- dati di base
- componente correttiva di *nowcasting*
- intervalli di confidenza e scenario mediano
- previsioni regionali di fecondità
- previsioni regionali di mortalità
- previsioni regionali dei movimenti migratori internazionali
- previsioni regionali delle migrazioni interne
- confronto con le precedenti previsioni
- confronto con le previsioni prodotte da Eurostat e Nazioni Unite
- diffusione dati e termini di utilizzo
- informazioni di contatto e richieste dati personalizzate.

Popolazione base

La popolazione base è quella articolata per sesso, singola classe di età e regione al 1° gennaio 2020, così come identificata dal Censimento permanente della popolazione e delle abitazioni. La popolazione include tutte le persone usualmente residenti in Italia, di qualunque cittadinanza, mentre non include né cittadini italiani residenti all'estero, né cittadini illegalmente o irregolarmente presenti sul territorio nazionale che non risultino iscritti presso alcuna anagrafe.

Tecnica di proiezione

Le previsioni sono condotte con tecnica iterativa tra il 1° gennaio e il 31 dicembre di ogni anno, utilizzando il cosiddetto metodo per "coorti-componenti". In corrispondenza di ciascuna classe di età alla popolazione iniziale vengono sommate le immigrazioni (dall'estero o da altre regioni) mentre vengono sottratti decessi ed emigrazioni (per l'estero o per altre regioni), ottenendo così la popolazione in vita alla fine dell'anno. A ciò vanno aggiunti i nati nel corso dell'anno che, al netto dei decessi e dei movimenti migratori che li riguardano, risultano ancora in vita al 31 dicembre.

Per la popolazione (stock) l'età è definita in anni compiuti al 1° gennaio (da 0 a 110 anni e più). Lo stesso dicasi per i dati di flusso come quelli sui nati, sui morti e sui movimenti migratori. Ciò permette di identificare, sempre e comunque, gli eventi demografici per anno di nascita dei soggetti coinvolti, assicurando la coerenza richiesta all'interno dell'equazione della popolazione.

Si assume che gli eventi demografici possano occorrere linearmente in qualunque momento dell'anno. Tra l'evento di morte e quello di migrazione (per l'interno o con l'estero) si assume che sussista incompatibilità, ossia che non possano coinvolgere un medesimo individuo nello stesso anno.

I decessi si determinano moltiplicando la popolazione residente per classi di età al 1° gennaio per le rispettive probabilità (prospettive) di morte, quelle che cioè interessano soggetti appartenenti alla stessa coorte di nascita.

Le nascite di un dato anno si ottengono in tre passaggi. Nel primo si moltiplica il contingente medio di donne relativo a ogni età feconda (ottenuto come media delle popolazioni di tal età all'inizio e alla fine dell'anno) per il rispettivo tasso di fecondità. Nel secondo si effettua la somma dei nati per età della madre, ottenendo il totale dei nati nell'anno. Nel terzo si scompongono i nati per sesso utilizzando il rapporto fisso di 106 nati di sesso maschile ogni 100 nati di sesso femminile.

Le previsioni hanno un profilo territoriale e sono costruite nella logica del modello multi-regionale, modello che, con particolare riguardo ai flussi migratori interni, tratta simultaneamente e coerentemente le distinte unità territoriali di riferimento. Il modello previsivo sulle migrazioni interne parte dalla costruzione di una matrice multi-regionale di probabilità migratorie per regione di origine, regione di destinazione, sesso, ed età. Tale matrice, applicata alla popolazione a rischio di migrare, identifica in ogni anno di previsione una serie coerente di immigrati ed emigrati.

Periodo di previsione

Le previsioni coprono il periodo tra il 1° gennaio 2020 e il 1° gennaio 2070. Scopo principale è fornire indicazioni sul futuro sviluppo della popolazione nel breve termine (2030), quindi quello di fornirne nel medio (2050) e lungo termine (2070), precisando che quest'ultime vanno utilizzate con cautela dal momento che i risultati diventano tanto più incerti quanto più ci si allontana dall'anno base (2020). Tale rischio è tanto più concreto quanto più si cala l'attenzione sulle unità territoriali più piccole, come nel caso di alcune regioni italiane.

Panel di esperti

Un panel di esperti nazionali ha supportato l'Istat nella definizione delle ipotesi demografiche relative all'Italia. Le ipotesi relative alle regioni, invece, sono state curate dall'Istat sulla base di un'apposita metodologia "ponte" tra le ipotesi nazionali e quelle regionali. Gli esperti che hanno risposto al questionario (con tecnica CAWI), fornendo informazioni utili e complete a definire le ipotesi, sono stati 86. Essi sono stati reclutati volontariamente tra i partecipanti alla 13^a edizione delle Giornate di Studio della Popolazione organizzata dall'Associazione Italiana per gli Studi di Popolazione (AISP), che ha avuto luogo a Milano tra il 24 e il 26 gennaio 2019 presso l'Università Bocconi. In particolare, si tratta di 50 donne e 36 uomini, principalmente occupati nelle Università (21 del Nord Italia, 11 rispettivamente del Centro e del Mezzogiorno e 10 afferenti Università estere) o in altro Ente di Ricerca pubblico (24). L'età media dei rispondenti è pari a 44 anni mentre la loro esperienza lavorativa è pari in media a 16 anni.

In tutte le fasi che hanno riguardato la costruzione dell'impianto metodologico alla base delle previsioni, l'Istat si è avvalsa della concreta cooperazione di Francesco Billari e Rebecca Graziani dell'Università Bocconi di Milano.

Questionario degli esperti e modello probabilistico

Il metodo probabilistico adottato poggia sulle opinioni degli esperti (*expert-based model*) per definire l'evoluzione futura dei più importanti indicatori demografici e rientra nella classe più ampia dei modelli random scenario. Tale modello, utilizzato per la definizione degli scenari probabilistici a livello nazionale, si fonda sull'elicitazione di una serie di parametri da cui viene derivata la futura evoluzione stocastica di ciascuna componente demografica. Gli esperti sono chiamati a fornire dei valori a un dato anno "t" riguardo a una serie di indicatori demografici di sintesi, condizionatamente ai valori assunti dagli stessi indicatori in istanti di tempo precedenti l'anno "t" (Billari, Graziani e Melilli, 2012).

Il metodo ha il vantaggio di essere semplice e flessibile. Nel questionario, infatti, le componenti demografiche necessarie sono sintetizzate attraverso i seguenti indicatori: il numero medio di figli per donna; la speranza di vita alla nascita distinta per sesso; le immigrazioni e le emigrazioni con l'estero. Le altre informazioni necessarie alla produzione delle previsioni, come quelle riguardo alla distribuzione per età (cadenza) degli eventi demografici, sono tenute appositamente fuori e lavorate successivamente al fine di rendere parsimonioso il questionario e il modello previsivo medesimo.

PROSPETTO A1. VALORI MEDI, VARIANZE E CORRELAZIONI OTTENUTE SOTTO LE IPOTESI DEGLI ESPERTI PER TIPO DI INDICATORE. Anni 2019, 2050 e 2080

Indicatore	Numero medio di figli per donna	Speranza di vita alla nascita – Uomini	Speranza di vita alla nascita – Donne	Immigrazioni dall'estero (migliaia)	Emigrazioni l'estero (migliaia)
Anno 2019					
Valore osservato	1,27	81,1	85,4	333	180
Anno 2050					
Ipotesi media	1,51	84,7	88,1	256	131
Ipotesi alta	1,75	86,3	89,5	343	172
Varianza	0,034	1,441	1,309	4.593	1.042
Anno 2080					
Ipotesi media condizionata all'ipotesi media 2050	1,55	87,1	90,0	240	127
Ipotesi media condizionata all'ipotesi alta 2050	1,74	88,4	91,0	305	158
Ipotesi alta condizionata all'ipotesi media 2050	1,75	88,5	91,5	348	186
Varianza	0,044	2,180	2,002	7.523	2.675
Correlazione 2050-2080					
Coefficiente di correlazione	0,68	0,66	0,54	0,51	0,46

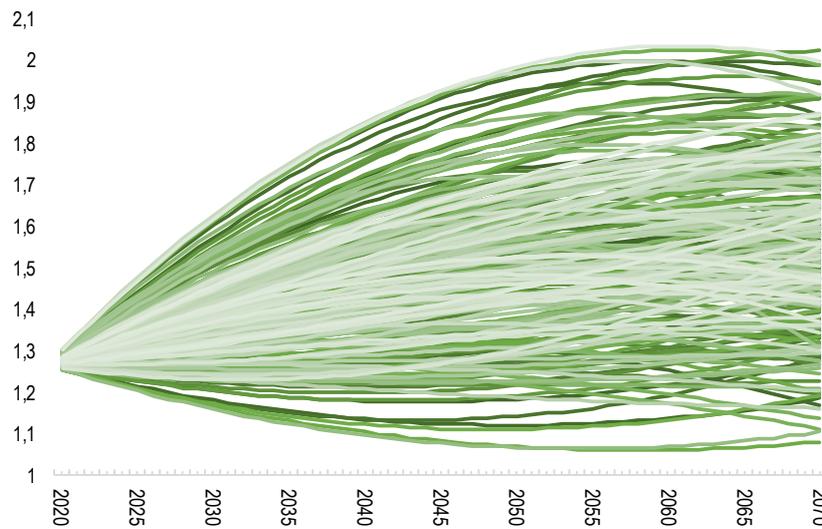
Per ogni indicatore demografico si prendono in considerazione due istanti temporali: un anno intermedio "t1" e un anno "t2" corrispondente all'ultimo anno di previsione. Nel questionario sottoposto agli esperti si considera "t0 = 2019", "t1 = 2050", "t2 = 2080", generando in questo modo due sotto-intervalli, 2019-2050 e 2050-2080. Indicare il valore della speranza di vita alla nascita nell'anno 2080, dato il valore previsto nel 2050, è un pratico esempio di come funziona il meccanismo.

Gli indicatori demografici sono assunti, per semplicità di ipotesi, tra loro indipendenti (ad esempio, la propensione media di fecondità espressa dalle donne non è influenzata dal livello delle migrazioni e viceversa), benché il modello permetta nella sua versione generalizzata la possibilità di far interagire tra loro le componenti demografiche in gioco. Si assume, inoltre, che la coppia di elicitazioni al 2050 e al 2080 di un dato indicatore possenga una distribuzione normale bivariata.

Sotto tali condizioni, una volta ottenuti i valori medi di risposta forniti dagli esperti è possibile stimare la varianza associata in ciascuno dei due istanti temporali futuri nonché la correlazione tra il primo e il secondo (Prospetto A1). Sulla base delle corrispondenti distribuzioni normali bivariate sono state quindi effettuate 3000 simulazioni col metodo Markov Chain Monte Carlo¹.

¹ La scelta è caduta sugli anni 2050 e 2080 al fine di individuare due intervalli temporali di analoga lunghezza. Per quanto la disponibilità di informazioni raccolte presso gli esperti permetta di estendere l'orizzonte delle previsioni fino al 2080, si è optato di arrestare l'esercizio iterativo di simulazione fino al 2070 onde ricavare un periodo previsivo di 50 anni precisi.

FIGURA A1. EVOLUZIONE DEL NUMERO MEDIO DI FIGLI PER DONNA SULLA BASE DI 3.000 SIMULAZIONI STOCASTICHE OTTENUTE A PARTIRE DALLE VALUTAZIONI DEGLI ESPERTI. Anni 2020-2070



L'ultimo passo di stima è finalizzato al calcolo dei valori di ciascun parametro negli anni intermedi i due intervalli 2019-2050 e 2050-2080. Tale attività è realizzata, per ciascuna delle 3000 simulazioni, mediante interpolazione con curve di tipo quadratico, passanti per i punti noti in corrispondenza degli anni 2019, 2050 e 2080. Si è così pervenuti alla definizione di 3000 curve stocastiche per ogni indicatore demografico considerato a livello nazionale. A titolo di esempio, la Figura A1 descrive il fascio di curve relativo al numero di figli per donna, ricavato dalla procedura sopra descritta.

La scelta di considerare un numero di 3000 simulazioni è frutto di un compromesso tra due necessità, entrambe strategiche: quella di rappresentare fedelmente l'aleatorietà degli eventi demografici e quella di ottimizzare i tempi macchina di elaborazione delle previsioni. Questi ultimi, nonostante l'odierna disponibilità di strumenti hardware/software sempre più potenti e sofisticati, costituiscono a loro volta un aspetto tecnico tutt'altro che secondario considerando la notevole mole di dati trattati.

Relazione tra le previsioni nazionali e regionali

Il modello probabilistico fornisce un set di 3000 simulazioni nazionali per ciascun indicatore demografico di sintesi. Poiché l'obiettivo delle previsioni Istat è anche quello di fornire indicazioni a livello territoriale, proseguendo la tradizione del modello multiregionale, è stata implementata una procedura "ponte" tra la definizione degli input nazionali e quelli regionali. L'approccio è di natura *top-down* dal lato della costruzione delle ipotesi mentre, si vedrà più avanti, è di natura *bottom-up* dal lato della produzione degli output finali.

L'azione principale è quella di derivare 3000 scenari regionali stocastici dai 3000 scenari nazionali frutto della consultazione tra gli esperti. La prima operazione in tal senso è quella di elaborare una previsione deterministica intermedia, applicando il modello multi-regionale a coorti componenti, includendovi anche la modellazione delle migrazioni interregionali. Da tale previsione, ottenuta estrapolando le tendenze regionali ritenute più probabili per ciascuna componente (cfr. paragrafi successivi), vengono ricavati gli stessi indicatori di sintesi oggetto del modello stocastico precedentemente descritto, ossia numero medio di figli per donna, speranza di vita alla nascita maschile e femminile, movimenti migratori con l'estero. Questa prima previsione intermedia, unica e deterministica, somiglia in sostanza a quella che in un approccio di tipo deterministico verrebbe etichettata col termine di "scenario centrale".

A questo punto, la transizione dal modello regionale deterministico al modello regionale stocastico si realizza moltiplicando, e reiterando 3000 volte il procedimento, la previsione dell'indicatore di sintesi deterministico regionale per il rapporto intercorrente tra la previsione nazionale stocastica e quella deterministica. In formula, indicando con "n" la generica simulazione ($n=1, \dots, 3000$), con "j" il codice territoriale regionale, con DR la previsione regionale deterministica, con SR quella stocastica, con DN e SN, rispettivamente, la previsione nazionale deterministica e stocastica, si ha:

$$SR_{t,n}^j = DR_t^j \times \frac{SN_{t,n}}{DN_t}$$

agganciando così, a ogni simulazione, il vettore dei valori regionali al valore nazionale stocastico di riferimento.

Si noti che per quanto concerne gli indicatori sintetici delle immigrazioni e delle emigrazioni con l'estero si ha che:

$$DN_t = \sum_j DR_t^j$$

$$SN_{t,n} = \sum_j SR_{t,n}^j$$

Ottenuti gli indicatori sintetici stocastici a livello regionale si passa alla costruzione degli input necessari all'applicazione del metodo per coorti-componenti, ossia le probabilità prospettive di morte per sesso ed età, i tassi specifici di fecondità per età della madre e la distribuzione degli immigrati/emigrati con l'estero per sesso ed età. La procedura associa quindi a ogni indicatore di sintesi la relativa cadenza del fenomeno (sesso ed età). Quest'ultima, non trattata in via stocastica, è quella che deriva dal modello deterministico regionale e, di simulazione in simulazione, riadattata allo specifico indicatore sintetico stocastico.

L'accoppiamento dei 3000 vettori di probabilità di morte (ciascun vettore sviluppa un numero di elementi pari a "numero di regioni X classi di età X sesso X anni di previsione") con i 3000 vettori di fecondità, e gli altrettanti sulle immigrazioni e le emigrazioni con l'estero e, infine, con le 3000 matrici O/D di probabilità di migrazione interna, è di tipo casuale.

Dopo aver introdotto una componente correttiva di *nowcasting* (cfr. paragrafo successivo) relativa ai primissimi anni di previsione, il modello coorti componenti viene fatto a quel punto girare 3000 volte, ottenendo così gli output richiesti: popolazione per sesso ed età, flussi demografici per sesso ed età, più la serie di indicatori demografici di supporto all'analisi che va dai tassi generici (di natalità, di mortalità ecc..) agli indicatori strutturali (età media, indice di dipendenza, di vecchiaia ecc..).

I risultati a livello nazionale (nonché di ripartizione geografica) nell'ambito di ciascuna simulazione regionale si ottengono per somma (approccio *bottom-up*). Pertanto, l'ammontare della popolazione prevista, dei decessi, e delle migrazioni, classificate per sesso ed età, e delle nascite per età della madre che si determinano a livello nazionale (o di ripartizione) sono la sommatoria delle traiettorie previsive regionali. Anche gli assunti livelli nazionali (e di ripartizione) relativi agli indicatori di sintesi posti in diffusione, riguardo ad esempio la speranza di vita o il numero medio di figli per donna, sono ricalcolati ex-post sulla base di tali riepiloghi regionali.

È opportuno segnalare che la stocasticità introdotta a livello regionale, mutuata *top-down* da quella nazionale e limitata ai soli indicatori di sintesi, potrebbe risultare non sempre sufficiente a riprodurre l'aleatorietà dei vari eventi demografici. Ciò è particolarmente vero nelle piccole realtà territoriali dove l'incertezza tende a essere relativamente maggiore. Per questa ragione è più opportuno parlare di approccio semi-stocastico quando ci si riferisce alle previsioni regionali, per quanto la numerosità delle simulazioni condotte offra comunque ampia garanzia di rappresentatività della variabilità su scala regionale.

Una seconda osservazione riguarda il fatto che nel modello Istat è esclusa una trattazione statistica generalizzata della covarianza tra le Regioni (ad esempio: la previsione di incremento/decremento della fecondità in una data regione quanto condiziona o quanto è a sua volta condizionata dalla previsione di incremento/decremento in un'altra). A questa strada, esclusa anche per ragioni di parsimonia del modello statistico, se ne è preferita un'altra, quella della convergenza territoriale. Infatti, il modello regionale deterministico iniziale, successivamente trasformato in modello stocastico attraverso la procedura sopra descritta, è costruito su ipotesi di convergenza a lunghissimo termine (2120, ben oltre l'ultimo anno delle previsioni) tra le regioni per ciascuna componente demografica fondamentale. Ciò comporta che i 3000 scenari stocastici regionali raffigurano 3000 diverse ipotesi di convergenza dei comportamenti demografici sul territorio.

L'ipotesi principale alla base della convergenza è che le differenze socio-economiche e culturali attualmente in essere tra le regioni costituenti il territorio italiano siano destinate a scomparire nel lungo termine. Perciò, il loro progressivo annullamento comporterebbe anche un generalizzato riavvicinamento dei comportamenti demografici. L'idea di convergenza non è nuova in demografia ed esistono molti esempi di previsioni demografiche che la seguono (Eurostat e ONU, in particolare), tra cui anche quelle passate dell'Istat. Nelle previsioni Istat la convergenza è intesa come spostamento progressivo di un dato comportamento demografico verso un punto molto distante nel futuro che rappresenta l'istante di piena convergenza regionale (nel senso che a quel punto i valori risulterebbero identici per le diverse regioni), ma che in realtà è lungi dall'essere raggiunto all'interno dell'orizzonte previsivo considerato (2020-2070). Di fatto, risulta corretto in tale circostanza parlare più di modello di semi-convergenza che di modello di convergenza piena.

Dati di base

Le ipotesi definite a livello regionale nel modello deterministico preliminare, prima del passaggio al modello stocastico, sono state ottenute estrapolando i futuri trend dall'analisi delle serie storiche osservate. In particolare tali ipotesi sono state definite ricorrendo alle seguenti serie di dati:

- per la fecondità, i tassi specifici per età della madre del periodo 1977-2019;
- per la mortalità, le probabilità di morte per sesso ed età del periodo 1974-2019;
- per le migrazioni interne e internazionali, i trasferimenti di residenza per sesso ed età del 2015-2019.

Componente correttiva di nowcasting

Prima di essere lanciate a pieno regime lungo l'orizzonte previsivo col metodo coorti-componenti, le previsioni incorporano un fattore correttivo di *nowcasting* (dal termine *nowcast*=previsione del presente, contrapposto al più noto termine *forecast*=previsione del futuro). Con tale operazione si intende fare in modo che la previsione relativa ai primissimi anni risulti quanto più in linea alla tendenza emersa nell'ultimo periodo o nell'ultimo anno storico (effetto *jump-off*), pur preservando quella che è la potenziale variabilità del fenomeno studiato. Questo genere di operazione si rivela particolarmente idonea in anni contraddistinti da improvvisi, e in quanto tali imprevedibili, cambiamenti della congiuntura demografica. È questo il caso, come avvenuto nel 2020 e in parte minore nel 2021, degli effetti provocati dalla pandemia Covid-19 su tutte le componenti del ricambio demografico. Non soltanto, per quanto prioritariamente, sulla mortalità, ma anche sulla natalità e sui trasferimenti di residenza interni e internazionali. Nessun modello previsivo applicato alle serie storiche menzionate al paragrafo precedente, e di cui si dà documentazione in quelli successivi, avrebbe potuto essere in grado di prevedere con precisione lo shock determinato dalla pandemia. Non certamente gli oltre 746 mila decessi riscontrati per tutte le cause nel 2020 (circa 100 mila in più di quanto atteso alla luce delle prime valutazioni effettuate), ma nemmeno il calo ulteriore delle nascite (404 mila) nel contesto di un quadro complessivo già di per sé compromesso dalla ben nota contrazione dei comportamenti riproduttivi. Né, infine, poteva risultare prevedibile la forte contrazione registrata nelle migrazioni, conseguenti le misure intraprese a livello nazionale per il contenimento della diffusione del virus (*lockdown*).

Dal momento che la popolazione base delle previsioni è quella rilevata al 1° gennaio 2020, si sono rese necessarie operazioni di correzione a breve termine degli input previsivi che hanno interessato gli anni 2020 e 2021, direttamente, e quelli fino al 2024 indirettamente. Con ciò si vuole, infatti, non solo tenere conto degli eventi eccezionali che hanno caratterizzato i primi due anni, ma anche quelli successivi entro i quali si presuppone che gli effetti pandemici possano andare ad esaurirsi e a far sì di raccordare la serie prevista degli input di breve termine con quelli di medio e lungo².

Dal punto di vista computazionale la rivisitazione delle ipotesi di breve termine si realizza applicando dei fattori di correzione. Sia, ad esempio, E_b^j il numero di eventi demografici previsti nel primo anno sulla base dello scenario mediano nella regione j . E sia, invece, \hat{E}_b^j il valore osservato di tali eventi oppure, in assenza del valore realmente osservato, la migliore stima che si possa ottenere (ad esempio, utilizzando procedure di *nowcasting* o analoghi modelli statistici). Il rapporto:

$$r_b^j = \hat{E}_b^j / E_b^j$$

rappresenta il fattore di correzione da applicare alle misure statistiche che danno luogo agli eventi di tipo "E" nell'anno "b" per la regione j . Se ad esempio tali eventi fossero il numero totale di nascite allora la quantità:

$$\hat{f}_{b,x}^{n,j} = r_b^j \cdot f_{b,x}^{n,j} \quad \text{con } x=14, \dots, 50 \text{ e } n=1, \dots, 3000$$

rappresenta la serie dei tassi specifici di fecondità per età della madre (n -esima simulazione) corretta per l'anno "b". Analoghe considerazioni valgono per la determinazione dei coefficienti di correzione inerenti la mortalità e i movimenti migratori. Per quanto concerne il 2020 i fattori di correzione sono stati costruiti rapportando i dati del bilancio demografico provvisorio di ciascuna regione, rilasciati nel marzo 2021 dall'Istat³, alle previsioni preliminarmente prodotte per tale anno. Per quanto riguarda, invece, il 2021 si è fatto ricorso alla disponibilità dei dati preliminari di bilancio relativi ai primi 7 mesi (gennaio-luglio), completando la serie annuale (agosto-dicembre) applicando il modello previsivo di Holt-Winters, un particolare modello appartenente alla famiglia degli Exponential smoothing che ben si adatta alle serie storiche mensili qui considerate⁴. Le stime annuali così ottenute sono state

² Tenendo peraltro presente il meccanismo iterativo di calcolo offerto dal metodo coorti componenti, ovvero un meccanismo di continua interazione stock-flussi nel tempo, la correzione imposta nei primi anni ha effetti anche sui risultati di tutti quelli successivi, fino al 2070 in questo caso.

³ Cfr.: Istat, la dinamica demografica durante la pandemia covid-19 – anno 2020, <https://www.istat.it/it/archivio/255802>.

⁴ Si veda anche http://schedefontidati.istat.it/index.php/Nowcast_per_indicatori_demografici per approfondimenti a carattere metodologico sulle stime anticipate degli indicatori demografici.

quindi poste a numeratore per la costruzione dei relativi rapporti di correzione con i quali ricalibrare, per il 2021, le previsioni preliminari regionali sugli effetti di congiuntura⁵.

Per gli anni successivi al 2021 i fattori di correzione vengono applicati per un periodo limitato dell'orizzonte previsivo, elaborando pesi che tendono progressivamente a uno. In particolare, il numero di anni per cui il fattore di correzione si applica alla serie di interesse si desume da:

$$Y^j = \text{abs}(1 - r_b^j) \cdot \epsilon$$

con ϵ quantità arbitraria, scelta opportunamente per fare in modo di assicurare che in media regionale il numero di anni per garantire il rientro dalle previsioni di breve a quelle di medio-lungo termine non superi i 5 anni. A questo punto, i livelli dei fattori di correzione per gli anni successivi a "b", per un totale di "Y" anni, sono dati da:

$$r_t^j = \frac{r_b^j \cdot (b + Y^j - t) + (t - b)}{Y^j} \quad \text{con } t = b, b + 1, \dots, b + Y^j - 1$$

Intervalli di confidenza e scenario mediano

Una volta lanciata la procedura di calcolo inerente le 3000 simulazioni regionali, il margine di incertezza viene calcolato per tutti i possibili livelli informativi, dalla popolazione prevista ai dati di flusso, sulla base delle componenti strutturali del sesso e dell'età. Tali margini di incertezza dipendono a loro volta dall'incertezza insita nei futuri livelli di mortalità, fecondità e migratorietà che sono anch'essi resi disponibili. La diffusione dei risultati contempla il rilascio dei soli intervalli di confidenza del 90%, dell'80% e del 50% ma è possibile definire intervalli su qualunque scala di interesse. L'intervallo di confidenza restituisce l'informazione su quanto sia probabile che un determinato indicatore demografico ricada entro prefissati limiti. Sotto questo punto di vista è opportuno ricordare che tale probabilità rappresenta essa stessa una previsione, in quanto fondata su ipotesi la cui validità è incerta. In nessun caso, inoltre, gli estremi dell'intervallo di confidenza vanno interpretati come estremi limiti, superiori o inferiori, del futuro comportamento demografico.

La costruzione di un intervallo di confidenza è qui fondata sulla determinazione dei percentili nella distribuzione delle 3000 simulazioni. Ad esempio, l'intervallo di confidenza al 90% per un dato indicatore è determinato considerando i valori della distribuzione che ricadono tra il 5° e il 95° percentile. Si ricorda, inoltre, che l'incertezza si riferisce sempre e comunque al dominio dello specifico parametro stimato. I limiti dell'intervallo di confidenza per un dato livello gerarchico sono stimati in proprio, e non costruiti per sommatoria di limiti ottenuti a un livello di disaggregazione gerarchicamente inferiore. Il criterio trova applicazione anche in contesti gerarchici non territoriali; ad esempio nella composizione per età della popolazione o in quella per sesso.

Con lo scopo di definire una previsione "puntuale" che possa essere adottata quale riferimento più probabile dell'evoluzione demografica futura è stato definito lo "scenario mediano". Tale scenario corrisponde a una 3001-esima simulazione, ottenuta per costruzione, ma che di fatto non è stata rilevata nel campo di osservazione delle 3000 simulazioni. Il set di ipotesi viene identificato prendendo a riferimento il valore mediano tra tutte le simulazioni a livello delle singole componenti demografiche (fecondità, mortalità, migrazioni) nell'ambito delle possibili combinazioni delle covariate età, regione e anno di previsione. Ad esempio, il tasso specifico di fecondità all'età di 32 anni della regione Toscana, nell'anno 2040, dello scenario mediano è identificato come il valore mediano con tali caratteristiche individuato tra tutte le simulazioni. Il medesimo tasso specifico ma all'età successiva, o nell'anno successivo, è individuato con lo stesso procedimento ma esso scaturisce, verosimilmente, da una simulazione diversa. Per l'identificazione dello scenario mediano sulla mortalità e sulle migrazioni la procedura è identica ma con l'ulteriore covariata del sesso. Per le migrazioni interne, inoltre, le covariate territoriali riguardano la regione di origine e di destinazione.

Lo scenario è pertanto "mediano" dal lato degli input fondamentali. Dal punto di vista degli output (popolazione e flussi previsti) che tale scenario genera una volta lanciata la procedura per coorti-componenti, per le proprietà tipiche della mediana esso restituisce valori molto prossimi a quelli mediani.

Previsioni regionali di fecondità

Per la fecondità regionale le previsioni hanno riguardato i classici parametri di intensità e cadenza, ossia il numero medio di figli per donna e la distribuzione dei tassi specifici di fecondità per età della madre.

Il numero medio di figli per donna è stato rappresentato ricorrendo a modelli di tipo ARIMA(n,p,k) ricercando, distintamente per singola regione, quello più idoneo a prevedere l'intensità futura del comportamento riproduttivo, sulla base della serie storica 1977-2019. Il modello prevalentemente utilizzato è di tipo ARIMA(2,0,0) con intercetta.

⁵ In particolare tali stime annuali hanno dato luogo, su scala nazionale, a 391mila nascite, 705mila decessi, 271mila iscrizioni dall'estero e a 130mila cancellazioni per l'estero.

Il profilo per età della fecondità è stato modellato ricorrendo a un sistema di funzioni *quadratic splines* (Schmertmann, 2003). Tale modello descrive funzionalmente la curva dei tassi specifici di fecondità standardizzati in funzione di tre parametri: l'età di inizio dell'età fertile α ; l'età P in cui la fecondità raggiunge il suo livello massimo; l'età H, successiva a P, nella quale la fecondità si dimezza rispetto al livello massimo. Per tasso specifico di fecondità standardizzato si intende il tasso specifico di fecondità normalizzato all'unità in corrispondenza del valore massimo individuato all'interno della sua distribuzione per età.

Il modello di *quadratic splines* adatta cinque polinomi di secondo grado alle curve di fecondità. La funzione finale risulta continua con la derivata prima anch'essa continua. Inoltre, grazie ad opportune restrizioni matematiche essa è univocamente determinata dai tre parametri $[\alpha, P, H]$ sopra menzionati.

In pratica, la previsione del tasso specifico di fecondità si trasforma nella previsione dei tre parametri (attraverso modelli ARIMA) che lo esprimono funzionalmente, una volta stimatane la serie nel periodo 1977-2019. Per fare questo è stata adottata un'ipotesi di convergenza tra le regioni italiane, assumendo che le differenze territoriali in termini di comportamento riproduttivo tendano a diminuire nel lungo periodo. Dal punto di vista operativo la piena convergenza è stata fissata nel 2120. In particolare il vincolo di convergenza prevede che, dal 2020 al 2120, i parametri del vettore regionale $[\alpha, P, H]$ convergano linearmente ai valori di un ipotetico vettore nazionale, appositamente disegnato per l'operazione.

Previsioni regionali di mortalità

Le previsioni di mortalità sono state prodotte ricorrendo al modello di Lee-Carter (1992) nella variante proposta da Lee-Miller (2001), modello nel quale la procedura di *adjustment* riconduce le probabilità teoriche di morte a riprodurre precisamente il livello osservato della speranza di vita alla nascita, anziché il totale dei decessi osservati come nella versione originale. Inoltre, qui il modello viene applicato alla distribuzione per età delle probabilità di morte anziché a quella dei tassi specifici di mortalità della formulazione originale.

Il modello approssima la forma logaritmica delle probabilità di morte utilizzando tre parametri sintetici, di cui uno legato al trend $[k(t)]$ e due legati alla distribuzione per età $[(a(x), b(x))]$.

Come per la fecondità, anche per la mortalità la costruzione del modello origina dalla definizione di uno scenario di riferimento provvisorio a livello nazionale. La previsione si determina proiettando nel futuro il solo parametro nazionale di trend $k(t)$, la cui serie è individuata sul periodo 1974-2019, mentre i parametri $a(x)$ e $b(x)$ rimangono in questa fase invariati nel tempo. In particolare, per via della sua sostanziale linearità il parametro $k(t)$ è stato proiettato al 2070 con la tecnica del *random walk with drift*.

Le ipotesi a livello regionale vengono fatte discendere dal provvisorio scenario di riferimento nazionale, stimando in primo luogo i valori regionali dei tre parametri nel 1974-2019 con la stessa metodologia e, successivamente, facendo convergere ogni parametro regionale al corrispondente parametro nazionale al 2120. Pertanto, come conseguenza del processo di convergenza e diversamente dall'impostazione classica del modello di Lee-Carter, qui si fanno variare nel tempo anche i parametri regionali $a(x)$ e $b(x)$.

Previsioni regionali dei movimenti migratori internazionali

Al fine di catturare le tendenze più recenti, le previsioni dei flussi migratori con l'estero concentrano l'analisi solo sugli ultimi cinque anni, vale a dire sul 2015-2019. Questa necessità, considerando la complessità di prevedere flussi migratori internazionali ricorrendo ad analisi di lunghe serie storiche, porta all'uso di un modello molto semplificato. Senza dimenticare che a questo livello delle operazioni si tratta di strutturare un modello deterministico intermedio, i cui valori sono successivamente calibrati sulle intensità prodotte dal modello stocastico *expert-based*.

Nel primo anno di proiezione i valori totali delle immigrazioni e delle emigrazioni con l'estero sono posti pari al valore medio osservato nel corso degli ultimi cinque anni. In conformità con il quadro generale di convergenza del modello deterministico, si suppone quindi che in ciascuna regione ingressi e uscite convergano linearmente nel lungo periodo (2120) al medesimo livello, ossia alla semisomma iniziale dei due valori.

Una volta determinati i totali dei flussi in ingresso e in uscita fino al 2070, le distribuzioni per sesso ed età associate vengono derivate applicando il modello Castro-Rogers (Rogers and Castro, 1981) alla serie 2015-2019. Con tale modello si dimostra che il caratteristico profilo per età delle migrazioni (qualunque esse siano, in ingresso o in uscita, con l'estero o con l'interno) può essere descritto, indipendentemente dall'intensità del fenomeno, da una funzione matematica composta di quattro componenti additive e fino a 11 parametri predittivi. Tali parametri, la cui stima nel periodo osservato viene prodotta grazie a una procedura generalizzata per modelli non lineari (categoria nella quale la funzione Castro-Rogers ricade in pieno), vengono tenuti costanti nel periodo di previsione. Il risultato conclusivo è dunque che l'intensità globale dei flussi migratori con l'estero possa variare nel tempo ma sulla base di una composizione per età fissa.

Previsioni regionali delle migrazioni interne

Le migrazioni interregionali sono trattate secondo un approccio multidimensionale, che permette di considerare simultaneamente le aree di origine e destinazione dei flussi migratori, e di definire gli ingressi in una determinata

area come somma delle uscite con quella destinazione da tutte le altre aree del sistema. Il sistema è per costruzione coerente per tutti gli anni di previsione in quanto la riga e la colonna marginali della matrice O/D, corrispondenti rispettivamente ai flussi in entrata e in uscita in/da ciascuna regione, danno la stessa somma, corrispondente all'ammontare complessivo dei movimenti interni al territorio nazionale.

La probabilità di migrazione specifica per età (110), sesso (2), regione di origine (21) e di destinazione (21) costituisce la componente elementare della matrice O/D composta di $110 \times 2 \times 21 \times 21 = 97020$ celle per ciascun anno di calendario. Le probabilità sono stimate sulla base dei livelli osservati nelle singole annualità del periodo 2015-2019. I vettori di probabilità così ottenuti, a livello di ciascuna annualità, sono successivamente perequati utilizzando la funzione Castro-Rogers.

Pertanto, indicando con

$$m_{x,s,t}^{i,j}$$

la generica probabilità prospettiva di migrare per un individuo di età "x" e sesso "s" tra la regione "i" e la regione "j" relativa all'annualità "t" (t=2015, ..., 2019), si assume che questa rappresenti una variabile casuale di tipo normale con media pari al valore medio del quinquennio e varianza pari alla varianza rilevata nel quinquennio:

$$\mu_{x,s}^{i,j} = E(m_{x,s,t}^{i,j})$$

$$\sigma_{x,s}^{i,j} = E(m_{x,s,t}^{i,j} - \mu_{x,s}^{i,j})^2$$

Dalle suddette variabili casuali vengono estratti casualmente 3000 valori per ciascuno dei 97020 elementi della matrice di O/D, dando così luogo alla creazione casuale di 3000 matrici tra loro diverse. La matrice O/D relativa allo scenario stocastico mediano viene identificata prendendo a riferimento il valore mediano tra tutte le simulazioni nell'ambito delle possibili combinazioni delle covariate sesso, età, regione di origine e regione di destinazione. Tale matrice mediana è quella anche usata con lo scopo preliminare di produrre la previsione deterministica della popolazione, antecedente la transizione al modello stocastico vero e proprio (cfr. precedente paragrafo sulla relazione tra previsioni nazionali e regionali).

Si noti che nell'ambito di ciascuna simulazione (compresa quella relativa allo scenario mediano) la matrice O/D è supposta invariante nel tempo. L'ipotesi alla base del modello si fonda, infatti, sul mantenimento per tutto il periodo previsivo di una propensione alla mobilità che rimanga costante. Ciò comporta che i flussi migratori interni evolvano nel tempo solo in virtù delle variazioni che interessano livello e struttura per età della popolazione esposta al rischio di migrare.

Confronto con le precedenti previsioni Istat

Il confronto tra le correnti previsioni e quelle precedenti è in parte improprio per almeno tre ordini di motivi. In primo luogo la distanza temporale, pari a due anni, rispetto all'edizione più prossima (base 2018). In secondo luogo, il fatto che le precedenti previsioni adottano una popolazione base pre-censuaria (ossia non allineata al Censimento 2020 della popolazione ma a quello precedente del 2011). Infine, il fatto che l'impostazione metodologica delle attuali previsioni risulta rinnovata per quanto concerne le opinioni degli esperti.

Ciò premesso, una parziale valutazione del cambiamento occorso tra i due ultimi round previsivi può essere effettuata confrontando gli scenari mediani delle previsioni in base 2018 e 2020.

In primo luogo va rilevata una differenza piuttosto importante tra la popolazione totale base 2018 (60 milioni 484mila) e quella 2020 (59 milioni 641mila). Secondo la ricostruzione intercensuaria della popolazione, recentemente prodotta dall'Istat⁶, al 1° gennaio 2018 si stimavano 59 milioni 938mila residenti, vale a dire 546mila in meno di quanto preso a riferimento nelle previsioni in base 2018.

Dal lato dei flussi previsti nel tratto di comune proiezione (2020-2065) si coglie una valutazione molto più espansiva nelle previsioni in base 2018, dove ad esempio si prevedevano 20,3 milioni di nascite contro le attuali 18,1 milioni. Importante la differenza anche riguardo ai flussi con l'estero, dovuta a più immigrati e meno emigrati nelle previsioni in base 2018: rispettivamente 13,7 e 6,1 milioni contro 12,1 e 6,2 milioni. Più modesta è invece la differenza rispetto al volume dei decessi: 34,8 milioni nello scenario 2018 contro 34,7 milioni in quello 2020. Quale risultante di tali differenti dinamiche demografiche, la popolazione prevista a fine 2065 nello scenario in base 2018 ammontava a 53,5 milioni contro i 48,8 milioni nello scenario corrente.

La differenza tra le due popolazioni finali dei due distinti esercizi previsivi è dunque ascrivibile in parte alle differenti popolazioni di partenza, ma anche alla rivisitazione delle ipotesi evolutive, sia quelle relative agli indicatori di sintesi sia quelle relative all'aggiornamento della cadenza (distribuzione per età) degli eventi demografici oggetto di previsione. Il prospetto A2 mette in risalto come il processo di rivisitazione delle ipotesi per tutte le componenti demografiche abbia interessato soprattutto i primi anni di previsione e, in parte minore, quelli a medio-lungo termine.

⁶ Ricostruzione della popolazione residente, per sesso, età e Comune - Anni 2002-2018, <https://www.istat.it/it/archivio/255173>.

PROSPETTO A2. CONFRONTO TRA LE IPOTESI SUI PRINCIPALI INDICATORI SINTETICI DELLO SCENARIO MEDIANO IN BASE 2018 E DELLO SCENARIO MEDIANO IN BASE 2020. Anni 2020, 2030, 2050 e 2065.

Scenario mediano	Numero medio di figli per donna	Speranza di vita alla nascita – Uomini	Speranza di vita alla nascita – Donne	Immigrazioni dall'estero (migliaia)	Emigrazioni per l'estero (migliaia)
Anno 2020					
Base 2018	1,34	81,2	85,7	331	152
Base 2020	1,24	79,3	84,1	221	142
Anno 2030					
Base 2018	1,47	82,6	86,9	316	136
Base 2020	1,37	82,2	86,2	279	146
Anno 2050					
Base 2018	1,55	84,8	88,9	287	129
Base 2020	1,50	84,7	88,1	258	131
Anno 2065					
Base 2018	1,59	86,1	90,2	271	132
Base 2020	1,54	86,1	89,2	248	126

Confronto con le previsioni realizzate da Eurostat e Nazioni Unite

Per confrontare le previsioni prodotte dall'Istat con quelle di altri Enti ha senso prendere a riferimento le previsioni rilasciate dall'Eurostat e dalla *United Nations Population Division* (UNPD). Da anni, l'istituto statistico dell'Unione europea assolve il compito di produrre previsioni demografiche con cadenza regolare per tutti i Paesi membri. Le ultime rilasciate sono in base 2019, il cui principale scenario di riferimento è il cosiddetto *baseline scenario*. Anche la UNPD, a sua volta, produce previsioni demografiche con regolarità attraverso i *World Population Prospects*, all'interno dei quali sono contemplati tutti i Paesi del globo. Anche in questo caso le ultime previsioni disponibili sono in base 2019 e il principale scenario di riferimento è la *medium variant*.

Occorre evidenziare in premessa che, nonostante la comparabilità sul piano della tecnica proiettiva, gli esercizi prodotti dai due organismi internazionali presentano alcune differenze metodologiche rispetto a quello italiano. Tra queste, in primo luogo, il fatto che le previsioni di Eurostat e UNPD hanno come base di riferimento il 1° gennaio 2019, ovvero proiettano in partenza una popolazione non allineata alle risultanze del censimento 2020 e, in particolare, rispetto a queste ultime una popolazione significativamente più alta. In secondo luogo, va menzionato che i due modelli internazionali qui presi in esame sono di tipo uni-nazionale, ossia proiettano la popolazione residente in Italia nel suo insieme non tenendo conto dello sviluppo delle regioni.

Il prospetto A3 presenta le principali ipotesi di scenario messe a confronto. Per quanto attiene i flussi migratori il confronto è limitato al saldo migratorio con l'estero in quanto sia Eurostat sia UNPD costruiscono le ipotesi direttamente su tale indicatore (senza distinzione tra immigrati ed emigrati).

Per tutte le componenti demografiche in gioco le ipotesi adottate sono in partenza molto differenziate tra Istat e Eurostat/UNPD. Ciò si deve al fatto che, a differenza dello scenario Istat, le previsioni Eurostat e UNPD non scontano lo shock demografico del 2020 prodotto dalla pandemia Covid-19. Superato tale periodo, tuttavia, le ipotesi continuano a risultare piuttosto differenziate tra i vari enti produttori. In particolare, nei confronti dei movimenti migratori, laddove rispetto a una UNPD che si dimostra piuttosto cauta sull'Italia, si contrappone Eurostat con una visione molto più ottimistica. Tale evidenza si deve in parte alla metodologia di Eurostat, la quale, oltre a prevedere l'evoluzione di fondo delle migrazioni nette, incorpora nel modello una componente additiva di *replacement-migration*⁷.

Le ipotesi sulla fecondità sono abbastanza simili, per quanto nel medio-lungo termine Eurostat produca previsioni meno favorevoli rispetto sia a Istat sia a UNPD. Le ipotesi sulla sopravvivenza sono anch'esse non particolarmente distanti, tuttavia Eurostat e soprattutto UNPD evidenziano aspettative molto favorevoli circa l'allungamento della vita media, che nel modello Istat si intravedono solo in parte.

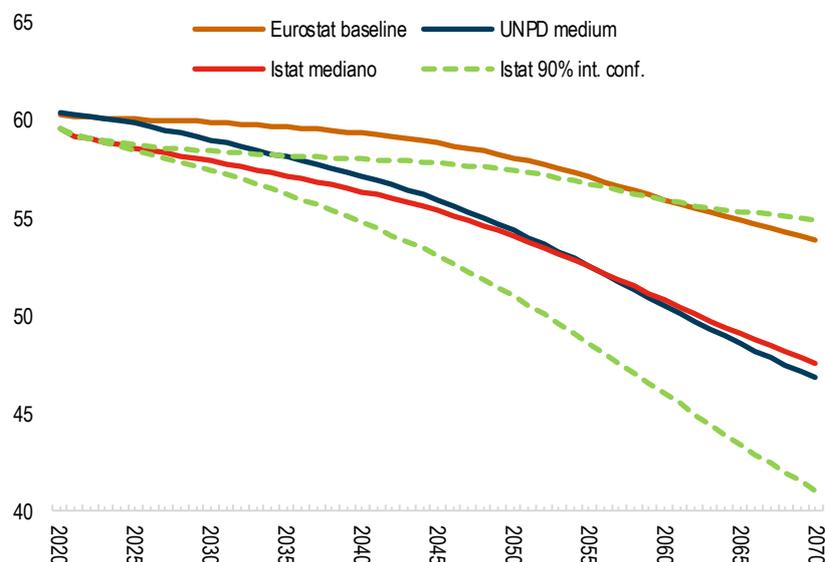
Lo sviluppo delle diverse ipotesi demografiche dà quindi luogo a delle differenze in termini di risultati attesi che, per quanto riguarda l'evoluzione della popolazione totale, è possibile apprezzare nella figura A2.

⁷ Tale componente assegna in ogni anno di previsione una quota aggiuntiva di migranti netti nella misura pari al 10% della riduzione riscontrata nella popolazione in età attiva (15-64 anni).

PROSPETTO A3. CONFRONTO TRA LE IPOTESI DEI PRINCIPALI INDICATORI SINTETICI TRA LO SCENARIO MEDIANO ISTAT IN BASE 2020, LO SCENARIO EUROSTAT IN BASE 2019 E LO SCENARIO UNPD IN BASE 2019. Anni 2020, 2030, 2050 e 2070.

Scenario	Numero medio di figli per donna	Speranza di vita alla nascita – Uomini	Speranza di vita alla nascita – Donne	Saldo migratorio estero (migliaia)
Anno 2020				
Istat Mediano	1,24	79,3	84,1	79
Eurostat Baseline	1,33	81,3	85,7	161
UNPD Medium	1,30	81,5	85,7	117
Anno 2030				
Istat Mediano	1,37	82,2	86,2	133
Eurostat Baseline	1,37	82,6	86,9	224
UNPD Medium	1,35	82,9	86,9	93
Anno 2050				
Istat Mediano	1,50	84,7	88,1	127
Eurostat Baseline	1,45	84,9	89,0	214
UNPD Medium	1,50	85,4	89,2	57
Anno 2070				
Istat Mediano	1,55	86,5	89,5	118
Eurostat Baseline	1,52	87,0	90,9	207
UNPD Medium	1,58	87,6	91,5	59

FIGURA A2. POPOLAZIONE TOTALE PREVISTA SECONDO GLI SCENARI ISTAT, EUROSTAT E UNPD. Anni 2020-2070, milioni di residenti



Purtroppo, lo scarto iniziale dovuto alle differenti basi di popolazione adottate nelle previsioni Eurostat/UNPD rendono piuttosto spurio il confronto con lo scenario Istat. Ciononostante, la traiettoria evolutiva della popolazione risulta coerente tra i tre scenari. Tutti, infatti, prevedono un progressivo declino della popolazione che tende ad accentuarsi nel medio-lungo termine. Lo scenario Eurostat, stante il significativo impatto di una previsione riguardo al saldo migratorio con l'estero più sostenuta, risulta particolarmente ottimista. Fino al punto da mantenere una popolazione più ampia persino del limite superiore dell'intervallo di confidenza al 90% delle previsioni Istat per gran parte dell'orizzonte previsivo. Lo scenario UNPD, al contrario, una volta assorbita la differenza rispetto alla popolazione base delle previsioni Istat, tende ad approssimarne lo scenario mediano.

Diffusione dati e termini di utilizzo

Il quadro dettagliato delle ipotesi sottostanti le previsioni e dei principali risultati è consultabile sia sul sito generalista di Istituto dati.istat.it (tema: Popolazione e famiglie > Previsioni demografiche) sia sul sito tematico demo.istat.it.

La diffusione dei dati è articolata in tre sezioni comprendenti tabelle elaborabili sulla struttura per sesso e singola classe d'età della popolazione, sulle componenti del bilancio demografico e sui principali indicatori demografici. Ciascuna tabella riporta i valori dello scenario mediano e i limiti, inferiore e superiore, degli intervalli di confidenza al 90%, all'80% e al 50%.

Tra le componenti del bilancio sono inclusi:

- popolazione a inizio e fine anno, saldo totale
- nascite e decessi, saldo naturale
- iscritti e cancellati con l'estero, saldo migratorio estero
- iscritti e cancellati con l'interno, saldo migratorio interno.

I dati sopra descritti e quelli relativi alla distribuzione per età della popolazione sono arrotondati all'unità.

Per quanto riguarda gli indicatori demografici le tabelle comprendono:

- tassi di natalità, mortalità e crescita naturale
- tasso immigratorio dall'estero, emigratorio per l'estero e tasso migratorio netto con l'estero
- tasso immigratorio dall'interno, emigratorio per l'interno e tasso migratorio netto con l'interno
- tasso migratorio netto totale e tasso di crescita totale
- età media della popolazione
- % di popolazione 0-14 anni, 15-64 anni, 65 anni e più, 85 anni e più
- indici di dipendenza strutturale, di dipendenza degli anziani e di vecchiaia
- numero medio di figli per donna
- speranza di vita alla nascita e a 65 anni di età per sesso.

La riproduzione delle informazioni contenute nella presente nota e nelle banche dati dati.istat.it e demo.istat.it è lasciata libera, a condizione che venga citata la fonte Istat.

L'Istat produce periodicamente le previsioni demografiche nel quadro della linea di attività "Sistema di stime e previsioni della popolazione", in conformità a quanto indicato nel Programma Statistico Nazionale, progetto "Previsioni demografiche" (cod. PSN IST-01448).

2) Previsioni delle famiglie per regione. Anni 2020-2040

Le previsioni delle famiglie mostrano l'andamento futuro del numero e della tipologia di famiglie che caratterizzeranno la popolazione in Italia dal 2020 al 2040. Si tratta di proiezioni derivanti dall'applicazione di un metodo statico, basato sui *Propensity rates*, applicati alla popolazione prevista. L'obiettivo è quello di offrire, insieme alle previsioni della popolazione, un sistema integrato di informazioni utili a numerosi utenti, pubblici e privati, che si occupano di beni e servizi destinati alle famiglie piuttosto che ai singoli individui. Data l'importanza del ruolo della famiglia, sia a livello protettivo sia per le scelte e i percorsi individuali, la domanda di informazioni su questo collettivo nasce dalle necessità di pianificazione in vari ambiti. Si pensi innanzi tutto alle decisioni da prendere nelle politiche di natura economica e sociale, come ad esempio quelle relative ai sistemi abitativi, sociali e assistenziali nei riguardi di giovani e anziani, ma anche alla produzione di beni di consumo durevoli per le famiglie e ai consumi di energia.

Dettaglio territoriale e periodo di previsione

Le previsioni delle famiglie sono diffuse per il totale Italia e con un dettaglio territoriale ripartizionale e regionale. Hanno base 1.1.2020 e si estendono per un arco temporale di 20 anni, dal 2020 al 2040.

Dati e popolazione base

I dati utilizzati sono: per la popolazione base, le Previsioni probabilistiche della popolazione residente per sesso, età e regione – base 1.1.2020 dello scenario mediano; per stimare la popolazione che vive in convivenza per sesso e singolo anno di età, il Registro base della popolazione – 1° gennaio 2019 e 2020; per quanto concerne le strutture familiari, l'Indagine Multiscopo "Aspetti della vita quotidiana", che offre la possibilità di utilizzare una lunga serie storica (dal 2002 al 2019). Per quanto concerne il concetto di famiglia si fa riferimento alla definizione di "famiglia di fatto" adottata in tutte le indagini Multiscopo dell'Istat, vale a dire quella che identifica come famiglia "l'insieme di persone legate da vincoli di matrimonio, parentela, affinità, adozione, tutela, o da vincoli affettivi, coabitanti e aventi dimora abituale nello stesso Comune".

La metodologia alla base delle previsioni

Il modello utilizzato per la previsione delle famiglie si basa su un adattamento al contesto italiano del metodo noto come “*Propensity model*”, a livello internazionale utilizzato negli ultimi anni dall'*Australian Bureau of Statistics* (ABS) per prevedere le famiglie in Australia e Nuova Zelanda (ABS, 2019)⁸. Si tratta di un metodo statico che va oltre il classico *Headship rate model*, superando il concetto di ‘capo famiglia’ e fornendo un set di informazioni molto più dettagliato. Come prodotto finale si ottengono previsioni del numero di famiglie future, la loro dimensione media e composizione. Alla base del metodo vi sono i *Propensity rates*, o tassi di propensione familiare, costruiti come la proporzione di persone di età x nella posizione familiare i al tempo t :

$$\text{Propensity Rate}_{x,i,t} = \frac{P_{x,i,t}}{P_{x,t}}$$

Ad esempio, la propensione per una persona con 30 anni di età a vivere in coppia con un partner sarà data dal numero di 30enni che vivono in coppia rapportati a tutta la popolazione dei 30enni.

I vantaggi del metodo sono molteplici: si lega facilmente alle previsioni della popolazione; non è necessario analizzare le transizioni tra le potenziali posizioni familiari, tipiche di un modello dinamico; è semplice da applicare e fornisce buoni dettagli nei risultati. Naturalmente vi sono anche alcune problematiche, che nascono principalmente dal fatto che si tratta di un metodo statico, e quindi non riproduce in modalità istantanea l'evoluzione dei fenomeni di formazione e scioglimento delle famiglie. Così l'applicazione dei tassi di propensione alla popolazione residente di riferimento può in alcuni casi determinare delle incoerenze, ad esempio tra sessi o per le posizioni familiari all'interno delle età, problematica che è quindi necessario aggiustare ex-post.

Il metodo si articola in 5 passi:

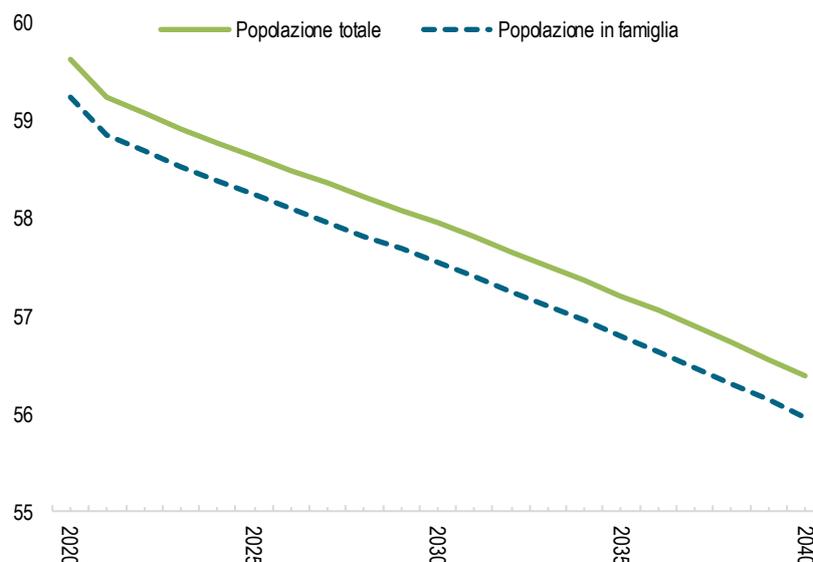
- Passo 1. Stimare la popolazione base prevista che vive in famiglia
- Passo 2. Calcolare le propensioni familiari (Propensity rates)
- Passo 3. Ipotizzare i trend futuri delle propensioni familiari
- Passo 4. Derivare la popolazione prevista nelle diverse posizioni familiari
- Passo 5. Calcolare il numero, la tipologia e la dimensione delle famiglie previste.

Di seguito si illustrano i vari passi in dettaglio.

Passo 1. Stimare la popolazione base e prevista che vive in famiglia

Poiché la popolazione base di riferimento è rappresentata dalla popolazione residente per sesso età e regione al 1° gennaio 2020 dell'ultimo Censimento permanente effettuato dall'Istat, è necessario effettuare preliminarmente una stima della popolazione che vive in famiglia, escludendo gli individui che risiedono in convivenza istituzionale (ospedali, caserme, carceri, case di riposo, strutture religiose, ecc.).

FIGURA A1. POPOLAZIONE RESIDENTE TOTALE E IN FAMIGLIA. Anni 2020-2040, scenario mediano, in milioni.



⁸ Le più recenti previsioni dell'ABS, rilasciate nel 2019, utilizzano un metodo di propensione sviluppato da McDonald e Kippen e pubblicato in “*Household Trends and Projections: Victoria, 1986-2011*”, riproposto dall'ABS in *Household and Family Projections, Australia, 2016 to 2041* - TECHNICAL NOTE – Method. Latest ISSUE Released: 14/03/2019, <https://www.abs.gov.au>.

Tale operazione va quindi ripetuta per tutti gli anni di previsione, defalcando dalle previsioni regionali (il cui riferimento è anche in questo caso la popolazione complessivamente residente) la quota parte di popolazione istituzionalizzata.

Ricorrendo ai dati del Registro base della popolazione all'1.1.2019 e all'1.1.2020, sono state calcolate le incidenze percentuali (medie) della popolazione che vive in istituzione per sesso, classe quinquennale di età e regione. Data la sostanziale stabilità nel tempo di questa popolazione, si è ipotizzato un trend costante per tutto il periodo di previsione.

Applicando alla popolazione totale, dal 2020 al 2040, il complemento a 100 di queste incidenze, si è ottenuta la popolazione che vive in famiglia per regione, sesso e classe di età (Figura A1).

Passo 2. Calcolare le propensioni familiari (*Propensity rates*)

Il secondo passo del modello consiste nel calcolare le propensioni a vivere in una determinata posizione familiare per sesso e classi di età quinquennali per le seguenti 8 posizioni familiari:

1. Persona sola
2. Persona in famiglia multipersonale (ad esempio 2 fratelli che convivono o un figlio divorziato che è tornato a casa da un genitore)
3. Persona in coppia senza figli
4. Persona in coppia con figli
5. Genitore solo con figli
6. Figlio (che vive con un genitore in coppia o con un solo genitore)
7. Altra persona che vive in un nucleo familiare
8. Persona in famiglia con 2 o più nuclei.

Le posizioni da 3 a 7 si riferiscono agli individui in famiglie composte da un solo nucleo familiare. Le persone che vivono in famiglie con 2 o più nuclei sono state considerate in una categoria a sé, costituendo tale tipologia familiare una quota esigua del totale delle famiglie (circa l'1,5%).

Come si è accennato, le propensioni (o *Propensity rates*) sono costruite come la proporzione di persone di età x nella categoria i . In questo contesto, la variabile età è stata considerata in classi quinquennali e i tassi sono stati disaggregati anche per sesso, variabile che risulta molto discriminante nei comportamenti familiari. Di seguito tali tassi sono denominati come *Living Arrangement Propensities* (LAP).

$$\text{Propensity Rate}_{x,i,s,t} = \frac{P_{x,i,s,t}}{P_{x,s,t}} = \text{LAP}_{x,i,s,t}$$

dove x = classe quinquennale di età 0-4, 5-9, ..., 80-84, 85+, i = posizione familiare, s =sesso, t =tempo.

Le LAP sono calcolate ricorrendo ai dati dell'indagine campionaria Istat Aspetti della vita Quotidiana, su tutta la serie storica 2002-2019. Poiché il dettaglio regionale per sesso e classi di età porta ad avere una esiguità di dati nelle regioni piccole, si è deciso di effettuare un raggruppamento di regioni in "macro-regioni".

Attraverso analisi statistiche multivariate, condotte sulla base di fattori di contesto sociodemografico⁹, si è arrivati a definire i seguenti 5 gruppi di regioni:

- Gruppo 1 - Nord-ovest (Piemonte, Valle d'Aosta, Lombardia, Liguria)
- Gruppo 2 - Est Adriatico (Veneto, Emilia-Romagna, Trentino-Alto Adige, Friuli-Venezia Giulia, Marche)
- Gruppo 3 - Tirrenico (Toscana, Lazio)
- Gruppo 4 - Sud (Campania, Puglia, Calabria, Sicilia)
- Gruppo 5 - Centrale (Umbria, Sardegna, Abruzzo, Molise, Basilicata).

Passo 3. Ipotizzare i trend futuri delle propensioni familiari

Occorre a questo punto fare delle ipotesi sull'evoluzione delle propensioni familiari dal 2020 al 2040. A tal fine, sono state inserite alcune modifiche al metodo dei *Propensity rates*. Il nuovo approccio si basa sull'introduzione di un nuovo indicatore sintetico, costruito come somma per età delle LAP, ponderate con gli anni vissuti alle varie età (L_x). Questo nuovo indicatore prende il nome di *Tasso di intensità Totale per ogni Posizione familiare* (TPT).

$$\text{TPT}_{i,s,t} = \sum_{x=0-4}^{85+} \text{LAP}_{x,i,s,t} * L_{x,s,t} = \sum_{x=0-4}^{85+} \frac{P_{x,i,s,t}}{P_{x,s,t}} * 100 * L_{x,s,t}$$

⁹ Allo scopo di individuare gruppi omogeni di regioni in base alle strutture familiari e alla loro evoluzione nel tempo, è stata effettuata un'analisi in componenti principali dinamica tramite la metodologia STATIS. L'analisi ha preso in esame le principali variabili socio-demografiche a livello regionale negli anni 2002-2019, tra cui: tassi di fecondità, età medie al parto, numero medio di componenti familiari, tassi di separazione e divorzio, tassi di occupazione femminile, tassi di migrazione interna ed estera, incidenza di alcune tipologie familiari (persone sole, coppie con e senza figli, genitori soli, ecc.). La procedura è stata ottimizzata eliminando le variabili con bassa variabilità latente spiegata dall'asse.

dove i = posizione familiare, s = sesso, x = classe quinquennale di età, t = tempo.

Gli $L_{x,s,t}$, che rappresentano gli anni vissuti nelle diverse classi di età x dal sesso s nell'anno t , sono desunti dalle tavole di mortalità proiettate dello scenario mediano.

Il TPT per una data posizione familiare rappresenta quanti anni mediamente si aspetta di vivere in quella posizione una generazione di individui, assumendo nel corso della vita i comportamenti familiari e le condizioni di mortalità osservate in un determinato anno di calendario. Si tratta, quindi, di una vera e propria aspettativa di vita in quello stato familiare, traslata dalla dimensione osservazionale trasversale a quella longitudinale. In altri termini, assume lo stesso significato che hanno indicatori trasversali più conosciuti, come ad esempio il numero medio di figli per donna, la speranza di vita alla nascita o il tasso di nuzialità totale.

Come può osservarsi nel Prospetto A1, se nel 2002 un uomo contava di vivere come persona sola mediamente 5,8 anni (su una speranza di vita totale di 77,2), nel 2019 il tempo atteso in questo stato sale a 9,4 anni (su un totale di 81). A seguito del calo della natalità, invece, nel 2002 le donne si aspettavano di vivere come persona in coppia con figli 22,7 anni (su un totale di 83), ma nel 2019 questo tempo atteso è sceso a 19,6 anni (pur su una aspettativa di vita totale nel frattempo salita a 85,3 anni). Per fare un ultimo esempio, il tempo nello stato di "figlio" è aumentato da 30,4 a 31 anni per i maschi e da 27,7 a 28,6 per le femmine, a causa della prolungata permanenza dei giovani all'interno della famiglia di origine.

PROSPETTO A1. TASSI DI INTENSITÀ TOTALE PER POSIZIONE FAMILIARE (TPT) E SESSO. Anni 2002-2019

Posizione familiare	MASCHI					FEMMINE				
	2002	2005	2010	2015	2019	2002	2005	2010	2015	2019
Persona sola	5,8	6,1	7,5	8,5	9,4	10,7	11,1	12,0	12,5	12,9
Persona in famiglia multipersonale	0,9	1,1	1,1	1,4	1,6	1,7	1,7	1,6	1,7	1,7
Persona in coppia senza figli	13,2	13,6	14,7	14,1	13,7	12,2	12,6	13,7	13,2	12,7
Persona in coppia con figli	23,2	22,6	21,6	20,4	19,7	22,7	22,0	21,1	20,2	19,6
Genitore solo	0,8	1,1	1,0	1,2	1,4	4,2	4,3	4,6	4,8	5,4
Figlio/a	30,4	30,7	30,2	30,9	31,0	27,7	28,1	27,8	28,3	28,6
Altra persona	0,8	0,6	0,8	0,9	0,9	1,5	1,4	1,1	0,9	1,0
Persona in famiglia con 2+ nuclei	2,1	2,3	2,4	2,7	3,3	2,3	2,3	2,4	3,0	3,4
Totale	77,2	78,1	79,3	80,1	81,0	83,0	83,5	84,3	84,6	85,3

Per ipotizzare i trend futuri delle propensioni, si è proceduto alla previsione dell'"intensità totale" del tempo trascorso in ogni posizione familiare, per poi stimarne la "cadenza" o distribuzione nelle varie età ($LAP_{x,i,s,t}$) in ogni anno di previsione. Prevedere in una prima fase l'intensità totale ha consentito, da un lato, di tradurre più facilmente le ipotesi previsionali sui comportamenti familiari, dall'altro, di mantenere insieme gli andamenti delle varie posizioni familiari che, se previste separatamente per fasce di età, risulterebbero di più difficile controllo con il rischio di ottenere risultati inattendibili (ad esempio, un tasso per la posizione familiare di figlio più alto alle età intermedie che in quelle giovanili).

L'obiettivo finale del passo 3, ossia la definizione delle LAP previste dal 2020 al 2040 per regione, è stato conseguito effettuando prima la previsione nei 5 gruppi territoriali stabiliti (passo 3.1), per poi arrivare da questi al dettaglio regionale (passo 3.2).

Passo 3.1 Previsione delle LAP nei 5 gruppi territoriali

L'intensità totale di ogni posizione familiare e sesso ($TPT_{i,s,t}$) è stata prevista attraverso l'estrapolazione dei trend nel periodo 2002-2019, utilizzando modelli di analisi delle serie storiche. Per ogni posizione familiare e sesso sono stati applicati i modelli descritti nel Prospetto A2. Tali modelli sono risultati validi per i 5 gruppi territoriali, a parità di posizione familiare e sesso.

Le ipotesi previsionali circa la futura evoluzione dei TPT sono rappresentate nella Figura A2, dove a titolo di esempio viene raffigurato il gruppo territoriale del Nord-ovest. In estrema sintesi, si ipotizzano delle variazioni del tempo trascorso nei vari ruoli familiari, che si traducono in:

- un aumento delle persone sole,
- la diminuzione dei partner con figli,
- un lieve aumento dei partner senza figli,
- un aumento delle persone nel ruolo di figlio,
- un leggero aumento dei genitori soli, soprattutto se padri,
- una sostanziale stabilità delle "altre persone" che vivono con famiglie con un nucleo,
- un lieve aumento delle persone in famiglie con 2+ nuclei.

PROSPETTO A2. MODELLI PREVISIVI DEI TASSI DI INTENSITÀ TOTALE PER POSIZIONE FAMILIARE E SESSO *

Posizione familiare	MASCHI	FEMMINE
Persona sola	RWD ARIMA(1,0,0)	RWD
Persona in famiglia multipersonale	RWD*	RWD
Persona in coppia senza figli	ARIMA(2,0,0)	RWD
Persona in coppia con figli	RWD ARIMA(2,1,0)	RWD ARIMA(2,1,0)
Genitore solo	RWD	RWD ARIMA(2,0,0)
Figlio/a	RWD	RWD
Altra persona	RWD	ARIMA(1,0,0)
Persona in famiglia con 2+ nuclei	ARIMA(1,1,0)	ARIMA(1,1,0)

*RWD=Random Walk with Drift model; ARIMA=AutoRegressive Integrated Moving Average model.

La stima della cadenza prevista, ossia delle $LAP_{x,i,s,t}$ dal 2020 al 2040, è stata ottenuta utilizzando i TPT previsti, gli anni vissuti previsti e le distribuzioni osservate dai dati dell'indagine AVQ nel triennio 2017-19. A tal fine, le distribuzioni medie 2017-19 di ogni ruolo in famiglia sono state ponderate con due coefficienti: uno che tenesse conto della previsione del TPT nell'anno t rispetto a quello del triennio 2017-19:

$$WP_{s,i,t} = \frac{TPT_{s,i,t}}{TPT_{s,i,2017-19}} \quad t = 2020, \dots, 2040$$

e un coefficiente che esprimesse i cambiamenti della mortalità nel tempo:

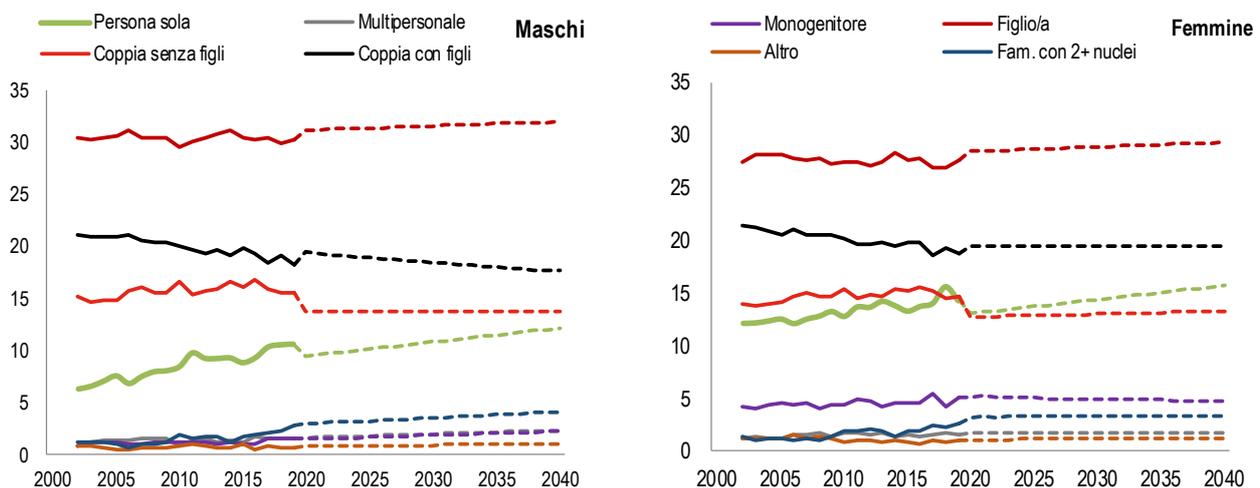
$$WL_{x,s,t} = \frac{L_{x,s,2017-19}}{L_{x,s,t}} \quad t = 2020, \dots, 2040$$

Pertanto le propensioni familiari per tutto l'orizzonte previsivo sono state calcolate con la formula:

$$LAP_{x,s,i,t} = LAP_{x,s,i,2017-19} * WP_{s,i,t} * WL_{x,s,t} \quad t = 2020, \dots, 2040$$

dove: x=classi di età 0-4, ... ,85+, s=sexo, i=ruolo in famiglia.

FIGURA A2. TASSI DI INTENSITÀ TOTALE PER POSIZIONE FAMILIARE E SESSO. Nord-ovest. Anni 2002-2040.



Con tale metodo di stima, non avendo prodotto ipotesi di variazione annuale circa la distribuzione per età delle LAP, si è implicitamente ipotizzato che i comportamenti in termini di scelte familiari mantengano in futuro una distribuzione per classi di età proporzionale a quella rilevata nell'indagine AVQ nel triennio 2017-2019. A seguito della procedura di stima, la somma delle LAP per posizione familiare in ogni classe di età approssima ma non sempre eguaglia il valore di 100, pertanto si rendono necessari degli aggiustamenti ex-post.

Passo 3.2. Previsione delle LAP nelle regioni

Per calare le previsioni delle famiglie a livello regionale è necessario fare in modo che ogni regione recuperi la propria specificità socio-demografica all'interno del gruppo previsivo di cui fa parte. A tale scopo si è definito un

fattore di correzione regionale da applicare alle LAP previste dei diversi gruppi territoriali in modo da ottenere quelle specifiche di ciascuna regione:

$$FC_{r,i} = \frac{TPT_{2017-19,i,r}}{TPT_{2017-19,i,G}}$$

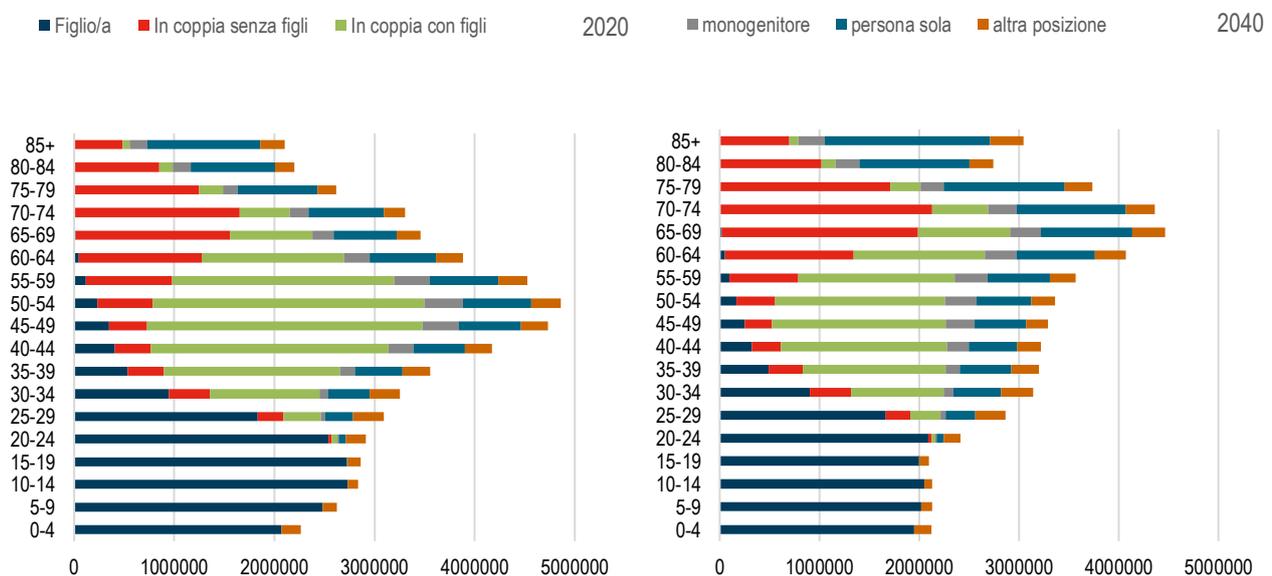
dove i=ruolo familiare, r=regione, G=gruppo cui appartiene la regione r.

Successivamente, le LAP previste per i gruppi territoriali vengono moltiplicate per il fattore di correzione regionale così calcolato, determinando la serie delle LAP regionali dal 2020 al 2040. Ad esempio, per la posizione familiare persona sola, il TPT rilevato in Piemonte è pari a 10,79 mentre nel gruppo 1 è pari a 10,42. Il fattore di correzione equivale dunque in questo caso a 1,04. Questo vuol dire che, poiché il Piemonte ha un TPT più alto di quello del gruppo cui appartiene, occorre fare un aggiustamento moltiplicando tutte le LAP alle diverse età e ai vari anni di previsione per 1,04, aumentandone leggermente il livello.

Passo 4. Derivare la popolazione prevista nelle diverse posizioni familiari

In questo passo, le propensioni regionali vengono applicate alla popolazione prevista che vive in famiglia ottenuta nel Passo 1. Si ricava così la popolazione prevista nei diversi stati familiari per sesso, classe di età e regione dal 2020 al 2040. La figura A3 mostra la struttura per età e posizione familiare nel 2020 e nel 2040 a livello Italia.

FIGURA 3. POPOLAZIONE PER POSIZIONE FAMILIARE. Anni 2020 e 2040, scenario mediano, valori assoluti.



Passo 5. Calcolare il numero di famiglie previste per tipologia

Il numero previsto di famiglie si ottiene direttamente dalla popolazione distinta per posizione familiare, sesso ed età, in quanto:

- ogni “persona sola” rappresenta 1 famiglia (coefficiente=1);
- le persone in coppia costituiscono lo 0,5 di una famiglia (coefficiente=0,5);
- ogni “genitore solo” rappresenta 1 famiglia (coefficiente=1);
- le famiglie “multipersonali” si ottengono dividendo il numero di persone che vivono in famiglie multipersonali per la dimensione media di questo tipo di famiglie, sostanzialmente stabile nel tempo e pari a circa 2,1 componenti (coefficiente=2,1);
- le famiglie “con 2 o più nuclei” si ricavano dividendo il numero di persone che vivono in famiglie con 2 o più nuclei per la dimensione media di questo tipo di famiglie, che assume valori storici di coefficiente tra 5 e 5,4 a seconda del gruppo territoriale di riferimento.

Applicando i coefficienti alla popolazione di sesso s ed età x si ottiene come prodotto finale il numero di famiglie distinto per le tipologie familiari di interesse.

Il numero medio di componenti familiari viene quindi calcolato rapportando la popolazione che vive in famiglia per il numero di famiglie. Può essere disaggregato per le famiglie totali e per quelle con almeno un nucleo familiare (escludendo le persone sole e le famiglie multipersonali).

Ai fini della diffusione, le famiglie multipersonali e quelle con 2 o più nuclei sono considerate insieme nella modalità “altro tipo di famiglia”. Nel Prospetto A3 si riporta il numero di famiglie previste per tipologia e il numero medio dei componenti familiari per gli anni 2020, 2025, 2030, 2035 e 2040.

PROSPETTO A3. FAMIGLIE PER TIPOLOGIA E NUMERO MEDIO DI COMPONENTI. Anni 2020-2040, scenario mediano.

	2020 *	2025	2030	2035	2040
Persone sole Maschi	3.618.643	3.727.303	3.892.183	4.078.456	4.251.616
Persone sole Femmine	4.953.422	5.195.014	5.489.075	5.792.028	6.079.054
Coppie senza figli	5.100.534	5.270.971	5.498.220	5.675.411	5.739.371
Coppie con figli	8.267.009	7.812.824	7.324.895	6.825.038	6.362.153
Genitore solo maschio	587.026	672.989	758.798	829.465	892.229
Genitore solo femmina	2.203.724	2.196.800	2.202.387	2.205.509	2.210.750
Altro tipo di famiglia	1.006.978	1.031.936	1.058.221	1.080.539	1.093.248
Famiglie con nuclei	16.568.501	16.358.225	16.182.266	15.925.626	15.584.663
Famiglie senza nuclei	9.168.835	9.549.612	10.041.513	10.560.821	11.043.759
Totale famiglie	25.737.335	25.907.838	26.223.779	26.486.446	26.628.422
Numero medio di componenti familiari	2,3	2,2	2,2	2,1	2,1
Numero medio di componenti famiglie con almeno un nucleo	3,0	2,9	2,9	2,9	2,8

(*) I dati ufficiali di indagine sugli Aspetti della vita quotidiana sono normalmente diffusi in media biennale. Qui, invece, i dati si riferiscono al 1° gennaio dell'anno indicato. Per il 2020 ciò può dare luogo a delle differenze.

Riferimenti bibliografici

1. ABS - Australian Bureau of Statistics (2019), *Household and Family Projections, Australia. Household and family projections (based on different assumptions of living arrangements) for Australia, states and territories and capital cities. Reference period: 2016 – 2041.* <https://www.abs.gov.au/>.
2. ABS - Australian Bureau of Statistics (2019), *Household and Family Projections, Australia methodology. Reference period: 2016 – 2041.* <https://www.abs.gov.au/methodologies/household-and-family-projections-australia-methodology/2016-2041>
3. Alho J. e Keilman N. (2010), *On future household structure.* Journal of the Royal Statistical Society Series A, 2010, vol. 173, Issue 1, 117-143.
4. Alho J.M. e Nikeer T. (2004), *Uncertain population of Europe - summary results from a stochastic forecast.* http://www.stat.fi/tup/euue/rp_reports_e_pub.html.
5. Alho J.M., Spencer B.D. (2005), *Statistical demography and forecasting.* New York: Springer.
6. Bernard A., Bell M. (2012), *A Comparison of Internal Migration Age Profile Smoothing Methods.* Working Paper 2012/01. Queensland Centre for Population Research, The University of Queensland.
7. Billari, F.C., Corsetti G., Graziani R., Marsili M. e Melilli E. (2014), *Towards stochastic forecasts of the Italian population: an experiment with conditional expert elicitations.* Proceedings of the 6th Eurostat/UNECE Work Session on Demographic Projections, pagg. 326-338, Istat, 2014.
8. Billari, F.C., Corsetti G., Graziani R., Marsili M. e Melilli E. (2014), *A stochastic multi-regional model for Italian population projections.* Budapest, 25-28 giugno 2014, [European Population Conference. http://epc2014.princeton.edu/papers/140361](http://epc2014.princeton.edu/papers/140361).
9. Billari, F.C., Graziani R. e Melilli E. (2012), *Stochastic population forecasts based on conditional expert opinions.* Journal of the Royal Statistical Society. Series A. 175(2): 491-511.
10. Blangiardo G., Barbiano di Belgiojoso E., Bonomi P. (2012), *Le previsioni demografiche delle famiglie.* In: Donati P. (a cura di), *La famiglia in Italia. Sfide sociali e innovazioni nei Servizi.* Osservatorio Nazionale sulla Famiglia. Rapporto biennale 2011-2012. Volume I Aspetti demografici, sociali e legislativi. Pagg 91-123.
11. Box G. E. P., Jenkins G. M., Reinsel G. C. and Ljung G. M. (2015), *Time Series Analysis: Forecasting and Control*, 5th Edition. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1.
12. Booth H. (2006), *Demographic forecasting: 1980 to 2005 in review,* International Journal of Forecasting, 22: 547-581.
13. Cooper J., Bell M. e Les M. (1995), *Household and Family forecasting Models: a review.* Review of IPC Long-Term Projections Model. Paper No. 3. Australian Housing and Urban Research institute in conjunction with Demographics Australia.

14. CBS (2011), *Key figures of the population forecasts 2010-2060*. Statline, Centraal Bureau voor der statistiek, <http://statline.cbs.nl/statweb/>.
15. Corsetti G., Marsili M. (2012), *A stochastic population projection from the perspective of a national statistical office*. European Population Conference. Stoccolma, 13-16 giugno 2012, EAPS. <http://epc2012.princeton.edu/papers/120635>.
16. Corsetti G., Marsili M. (2013), *Previsioni stocastiche della popolazione nell'ottica di un Istituto nazionale di statistica*. Rivista di statistica ufficiale, n. 2-3, p. 5-29, Istat.
17. Eurostat (2015), *People in the EU: who are we and how do we live?* - 2015 edition, Luxembourg: Publications Office of the European Union.
18. Eurostat (2020), *Methodology of the Eurostat population projections 2019-based (EUROPOP2019)*, Technical Note, Directorate of Social statistics, Population and migration, Luxembourg, 6 April 2020.
19. Graziani R., Keilman N. (2011), *The sensitivity of the Scaled Model of Error with respect to the choice of the correlation parameters: A simulation study*. Working Paper 37. Carlo F. Dondena Centre for Research on Social Dynamics, Università Bocconi, Milano.
20. Istat (1989), *Previsioni della popolazione residente per sesso, età e regione – Base 1.1.1988*, Note e Relazioni, n.4.
21. Istat (1997), *Previsioni della popolazione residente per sesso, età e regione – Base 1.1.1996*, Informazioni, n. 34.
22. Istat (1989), *Previsioni del numero di famiglie italiane dal 1995 al 2020*. Notiziario. Serie 4. Foglio 41 – Anno X No. 19. Dicembre 1989.
23. Istat (2001), *Previsioni della popolazione residente base 1° gennaio 2000*, Statistiche in breve, www.istat.it.
24. Istat (2003), *Previsioni della popolazione residente per sesso, età e regione, base 1.1.2001*, Informazioni n.13.
25. Istat (2006), *Previsioni demografiche nazionali 1° gennaio 2005-1° gennaio 2050*, www.istat.it, Nota informativa, 22 marzo 2006.
26. Istat (2008), *Previsioni demografiche 1° gennaio 2007-1° gennaio 2051*, www.istat.it, Nota Informativa, 19 giugno 2008.
27. Istat (2011), *Il futuro demografico del paese - Previsioni regionali della popolazione residente al 2065*, Statistiche Report, www.istat.it, 28 dicembre 2011.
28. Istat (2017), *Il futuro demografico del Paese – Previsioni regionali della popolazione residente al 2065*, Statistiche Report, www.istat.it, aprile 2017.
29. Istat (2018), *Il futuro demografico del Paese – Previsioni regionali della popolazione residente al 2065 – base 1.1.2017*, Statistiche Report, www.istat.it, maggio 2018.
30. Istat (2019), *Il futuro demografico del paese - Previsioni regionali della popolazione residente al 2065 – base 1.1.2018*, Nota metodologica, www.istat.it, ottobre 2019.
31. Istat (2021), *Ricostruzione della popolazione residente per sesso, età e comune, Anni 2002-2018*, Nota informativa, www.istat.it, marzo 2021.
32. Istat (2021), *Indicatori demografici – Anno 2020*, Statistiche Report, www.istat.it, 3 maggio 2021.
33. Keilman N. (2018) *Family Projection Methods: A Review*, March 2018, DOI: 10.1007/978-3-319-93227-9_12. In book: *Analytical Family Demography*. Publisher: Springer
34. Keilman, N. e Brunborg, H. (1995), *Household Projections for Norway, 1990-2020*, Part I: *Macrosimulations*, Statistics Norway, Oslo-Kongsvinger.
35. Keilman, N. Kuijsten A. and Vossen A. (1988), *Modelling Household Formation and dissolution*. Clarendon Press - Oxford.
36. Keilman N., Pham D.Q. e Hetle A. (2002), *Why population forecasts should be probabilistic - illustrated by the case of Norway*, Demographic Research, 6(15): 409-454.
37. Kono S. (1987), *The headship rate method for projecting households*, in Bongaarts J., Burch T., Wachter K., *Family Demography, Methods and their Applications*, Clarendon Press-Oxford.
38. Lavit, C. (1988), *Analyse conjointe de tableaux quantitatifs*, Masson, Paris.
39. Lavit, C., Escoufier, Y., Sabatier, R. and Traissac, P. (1994), *The ACT (Statis method)*. Computational Statistics and Data Analysis, Volume 18, Issue 1, 97–119.
40. Lee R.D. (1998), *Probabilistic Approaches to Population Forecasting*, Population e Development Review, 24: 156-190.
41. Lee R.D., Carter L.R. (1992), *Modeling and forecasting U.S. Mortality*, Journal of the American Statistical Association, September, vol. 87, n.419.
42. Lee R.D., Miller T. (2001), *Evaluating the performance of the Lee-Carter method for forecasting mortality*, Demography, November, vol. 39, p. 537-549.
43. Lutz, W., Sanderson W.C. e Scherbov S. (1998) *Expert-Based Probabilistic Population Projections*, Population e Development Review, 24: 139-155.
44. McDonald, P. and R. Kippen. (1998), *Household Trends and Projections: Victoria 1986–2011*. Canberra: Demography Program, The Australian National University.
45. Marsili M. (2007), *Demographic projections: the impact of net international migration on population ageing in Italy*, Atti del Convegno Intermedio della SIS 2007 “Rischio e Previsione”, Università Ca' Foscari, Venezia, 6-8 giugno.

46. Marsili M. (2020) *Scenari demografici, previsioni per l'uso*, Atti della 13° Conferenza nazionale di statistica, Dall'incertezza alla decisione consapevole: un percorso da fare insieme, Roma, 4-6 luglio 2018, pagg. 246-252, Istat, 2020.
47. ONS – Office for National Statistics (2020), *Methodology used to produce household projections for England: 2018-based. User guidance about uses, methodology, assumptions and input data for household projections for England*. <https://www.ons.gov.uk>.
48. Paciorek, A. (2014), *The Long and the Short of Household Formation*. Real Estate Economics, Forthcoming, Available at SSRN: <https://ssrn.com/abstract=2469334>
49. Rogers A. (1985), *Regional Population Projection Models*. Beverly Hills. CA: Sage.
50. Rogers A., Castro L. (1981) *Model migration schedules*, International Institute for Applied System Analysis, Laxenberg, Austria, RR-8 1-30, November 1981.
51. Rowan S., Wright E. (2010), *Developing stochastic population forecasts for the United Kingdom: Progress report e plans for future work*. Eurostat-UNECE Work session on demographic projections, Lisbon, 28-30 aprile 2010, Methodologies e Working papers, Commissione Europea.
52. Schmertmann C.P. (2003), *A system of model fertility schedules with graphically intuitive parameters*, *Demographic Research*, 9(5): 81-110.
53. Shaw C. (2008), *The National Population Projections Expert Advisory Group: results from a questionnaire about future trends in fertility, mortality e migration*. *Population trends* n.134, Winter 2008, Office for national statistics.
54. Stoto, M. A. (1983), *The accuracy of population projections*. *Journal of the American Statistical Association*. 78: 13–20.
55. Tuljapurkar S., Lee R.D. e Li Q. (2004), *Random scenario forecast versus stochastic forecasts*. *International Statistical Review*. 72: 185–199.
56. Terra Abrami V. (1998), *Le previsioni demografiche*, Il Mulino, Bologna.
57. UNECE (2018), *Recommendations on Communicating Population Projections*, United nations economic commission for Europe, United Nations, New York and Geneva, agosto 2018.
58. United Nations(1973), *Methods of projecting households and families*, Manual VII, New York.
59. United Nations (2019), *World Population Prospects 2019: Methodology of the United Nations population estimates and projections*, Department of Economic and Social Affairs/Population Division, ST/ESA/SER.A/425.
60. Wilson T. (2013), *The sequential propensity household projection mode*, DEMOGRAPHIC RESEARCH VOLUME 28, Article 24, Pages 681-712, <http://www.demographic-research.org/Volumes/Vol28/24/> DOI: 10.4054/DemRes.2013.28.24.

Note

¹⁾ I contenuti del paragrafo fanno riferimento a una *Statistica Sperimentale* in diffusione dal 29 novembre 2021. Considerato il minimo dettaglio territoriale preso a riferimento, i dati in esso commentati vanno trattati con estrema attenzione, ricordando che le previsioni demografiche divengono tanto più incerte quanto più ci si allontana dalla base di partenza, in particolar modo in piccole entità geografiche. Va parimenti sottolineato che le previsioni demografiche rappresentano un esercizio di tipo *what-if*. Sono cioè elaborazioni nelle quali i calcoli effettuati mostrano una particolare evoluzione della popolazione che è frutto delle specifiche ipotesi adottate riguardo al comportamento demografico. Per informazioni circa il processo metodologico e per l'accesso alla banca dati www.istat.it/it/statistiche-sperimentali.