JULY 24th, 2024

statistiche report

Istat | Istituto Nazionale di Statistica

POPULATION AND HOUSEHOLD PROJECTIONS | BASE 1/1/2023

Italy tomorrow: imbalance between new and old generations will grow, differences will increase

The new projections on the country's demographic future, updated to 2023, show trends whose direction seems irreversible, albeit in a context in which there is no lack of uncertainty. The **resident population is decreasing**: from around 59 million on 1 January 2023 to 58.6 mln in 2030, 54.8 mln in 2050 and 46.1 mln in 2080.

The ratio of **working-age individuals** (15-64 years) **to non-working-age** individuals (0-14 and 65 years and over) will increase from about **three to two** in 2023 to about **one to one** in **2050**.

With an average age of 51.5 years by 2050 (50.8 for Italy), there will be a faster ageing process in the South.

About one million more families in twenty years, but more fragmented. Fewer couples with children, more couples without: by 2043 less than one in four families will be composed of a couple with children, more than one in five will have none.

63%

The share of young people aged 25-29 still living with at least one parent in 20 years' time

64% in 2023

2,08

The average number of members per household in 2043 (2,25 in 2023)

4,1 million

People aged 75 and over living alone in 2043

An increase of 1.2 million compared to 2023

www.istat.it

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

CONTACT CENTRE
contact.istat.it

More than four million fewer residents by 2050

Between 2014 and 2023, under the action of recessionary demographic dynamics, Italy lost about one million 350 thousand residents (from 60.3 million to just under 59). In line with this trend, the 'median' projection scenario envisages a further decline of 439 thousand individuals by 2030 (58.6 million), with an average annual rate of change of -1.1‰. In the medium term, the population decline would be more pronounced: from 58.6 million to 54.8 million between 2030 and 2050 (average annual rate of change of -3.3‰) (Table 1).

If the demographic assumptions made under the median scenario were to occur, the population would fall to 46.1 million by 2080, decreasing by a further 8.8 million compared to 2050 (-5.8‰ per year on average), while the overall decline from the base year (2023) would amount to 12.9 million residents.

The evolution of the total population reflects the principle, typical of demographic projections, of being more uncertain the further one moves away from the base year. In 2050, its 90% confidence interval (i.e. that its assumed value falls between two extremes with a 90% probability) ranges between 52.7 and 57.0 million. Thirty years later it is between 39.3 and 53.1 million.

In the most favourable scenario, therefore, the population could experience a loss of 'only' 5.9 million between 2023 and 2080, of which 2.0 million already by 2050. In the least favourable case, on the other hand, the population decline would reach 19.7 million individuals by 2080, 6.3 million of which already by 2050. In conclusion, within the framework of reasonable and reasoned hypotheses (i.e. those that can potentially be envisaged for the country, unless *replacement level* scenarios are assumed) the population will decrease, but the extent of the reduction can present profoundly different numerical evidence from one another, which calls to mind not only demographic but also social and economic scenarios that are just as different.

Gradual depopulation affects the entire territory, but with differences between the North, the Centre and the South that make this issue reach a significant dimension especially in the latter area. According to the median scenario, in the short term a slight but significant increase in population is expected in the North (+1.5‰ per year up to 2030), while in the Centre (-0.9‰) and especially in the South (-4.8‰) a decrease in the number of residents is expected.

In the intermediate period (2030-2050), and even more in the long term (2050-2080), this evolutionary picture expands, with a generalised decline in population in all geographic areas, but particularly in the South. Looking at the long term, the North could shrink by 2.6 million inhabitants by 2080, but by only 50,000 when looking at 2050. Quite different is the evolutionary path of the population in the South, which could shrink by 7.9 million inhabitants by 2080, 3.4 million of them already by 2050.

Obviously, the above considerations must also be evaluated in light of the profound uncertainty that underlies them. In the North, a path of progressive population growth (up to 28.7 million residents by 2080) is also potentially possible, as represented by the upper bounds of the confidence interval throughout the projection horizon. Conversely, in both the Center and the South such a possibility is never contemplated, even under the most favourable scenario assumptions.

TABLE 1. RESIDENT POPULATION BY GEOGRAPHICAL AREA. MEDIAN SCENARIO AND 90% CONFIDENCE INTERVALS. Years 2023-2080, January 1st, million (*)

Geographic area	2023	2030	2040	2050	2080
North	27,4	27,7	27,7	27,4	24,8
		[27,6 / 27,8]	[27,2 / 28,2]	[26,2 / 28,5]	[21,0 / 28,7]
Centre	11,7	11,6	11,4	11,0	9,3
		[11,6 / 11,7]	[11,2 / 11,6]	[10,6 / 11,5]	[7,9 / 10,8]
South	19,9	19,2	17,9	16,4	11,9
		[19,1 / 19,3]	[17,6 / 18,2]	[15,9 / 17,0]	[10,3 / 13,7]
ITALIA	**59,0**	**58,6**	**57,0**	**54,8**	**46,1**
		[58,3 / 58,8]	[56 / 58,1]	[52,7 / 57,0]	[39,3 / 53,1]

(*) Values under the confidence intervals in square brackets.

Structural factors will influence births and deaths, uncertain the migration scenario

The median scenario of the forecasts shows that in the transition that will lead the population from today's 59 million individuals to about 46 in 2080, we glimpse, in aggregate, 21.5 million births, 44.9 million deaths, 18.2 million immigrations from abroad versus 8 million emigrations. In the most reliable scenario, therefore, the population changes radically, and not only in quantitative terms. The current older generations, in fact, bearers of their own values, customs, levels of education and skills, will give way to the new ones, who in turn will be bearers of equal but evolved characteristics.

To what extent this transformation can occur depends on the uncertainty associated with the various assumptions about future demographic behaviour, but not to the extent that, for example, the current gap between births and deaths is brought back into balance. Even in the most favourable birth and death scenarios, the number of births does not compensate for the number of deaths. For example, the upper bound of the confidence interval for births (a scenario in which the average number of children per woman increases to 1.85 in 2080) identifies a steadily lower amount of deaths expected along the lower bound (Figure 1).

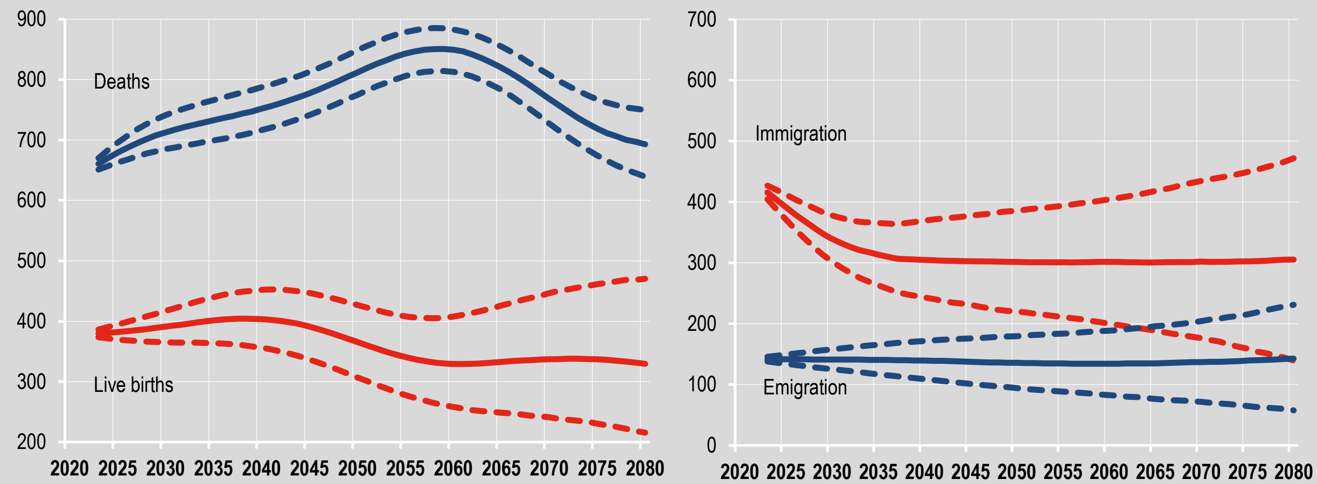
In the median scenario, where fertility is projected to grow from 1.20 children per woman in 2023 to 1.46 in 2080, births peak at 404,000 in 2038. Thereafter, the projected increase in average reproductive levels does not lead to a parallel increase in births, as it is countered by a gradual decline in women of childbearing age. It should be borne in mind that in 2023 the number of women of childbearing age amounts to 11.6 million and that, according to the median scenario, this contingent is set to shrink to 9.2 million in 2050 and 7.7 million in 2080. On the other hand, the prospect of fertility rising to 1.85 children per woman in 2080, recording an intermediate of 1.6 children per woman in 2050, i.e. what is assumed in the most favourable scenario, returns a level of births of less than 500,000 per year.

Similar structural perturbations will affect the evolution of mortality, for which a sustained number of deaths is expected annually, up to a peak of 851,000 in 2059 according to the median scenario. This will happen even in a context of good expectations on the evolution of life expectancy (86.1 and 89.7 years that expected at birth in 2080, respectively for men and women, with a gain of 4.8 years for the former and 4.4 years for the latter over 2023).

The median scenario contemplates positive net foreign migration movements. A very intense first phase up to 2040, corresponding to an average net flow of over 200,000 per year, is followed by a stabilisation phase lasting until 2080 at an annual average of 165,000.

Future migration flows do not counterbalance the negative sign of the natural dynamic. Nevertheless, they are characterised by uncertainty, due to the presence of multiple factors (migratory pressures in the countries of origin, the country's attractiveness in economic-employment terms, the instability of the international geopolitical framework characterised by war crises and the potential triggering of periods of economic recession alternating with periods of recovery). The analysis of long-term results must therefore be accompanied by great caution; from this point of view it is significant that the 90% confidence interval of the net foreign migration balance returns extremes ranging from -20 thousand to +349 thousand units in 2080.

FIGURE 1. NATURAL AND MIGRATORY POPULATION MOVEMENT, MEDIAN SCENARIO, AND 90% CONFIDENCE INTERVALS IN ITALY. Years 2023-2080, January 1st, thousand.



The imbalance between old and new generations will grow

The structure of the resident population has for years been subject to an imbalance between new and old generations, due to the typically Italian combination of increasing longevity and a persistently low fertility rate. Today, the country has the following age distribution: 12.4 per cent of individuals are up to 14 years old; 63.5 per cent between 15 and 64 years old; 24.0 per cent from 65 years old and up. The mean age, meanwhile, has risen to 46.4 years, which places Italy, immediately after Japan, among the world leaders in demographic transition, together with other Mediterranean countries (Portugal, Greece, Spain) and Germany.

Future prospects imply an amplification of this process, which is more governed by the current age distribution of the population than by assumed changes about the evolution of fertility, mortality and migration dynamics, based on a ratio of importance, roughly, of two-thirds and one-third respectively.

In 2050, persons aged 65 and over could represent 34.5% of the total according to the median scenario, while the 90% confidence interval presents a minimum of 33.1% and a maximum of 35.8%. Significant growth is also expected for the population aged 85 and over, in which a larger share of frail individuals will be concentrated, from 3.8% in 2023 to 7.2% in 2050 with confidence margins between 6.4 and 8%.

Whatever happens, therefore, the impact on social protection policies will be important, having to cope with the needs of an increasing share of the elderly. On the social security side, for example, the assumptions on life expectancy at age 65 contemplated in the median scenario predict a significant increase, under current legislation, in the retirement age. From the current 67 years, it would increase to 68 years and 2 months from 2035, to 69 years from 2045 and to 69 years and 10 months from 2055.

Young people up to 14 years of age, although in the median scenario fertility is expected to partially recover, could represent 11.2 per cent of the total by 2050, showing a moderate decrease in a relative sense but not in an absolute sense. In fact, in terms of intergenerational relations, there will be an unbalanced ratio of over three to one between the over-65s and the young.

Contributing to the absolute and relative growth of the elderly population will be above all the transit of the large generations of the baby-boom years (born in the 1960s and first half of the 1970s) between the adult and senile ages, with a concomitant reduction in the working-age population. In the next thirty years, in fact, the population aged 15-64 years would fall to 54.4% based on the median scenario, with a potential range between 53.3% and 55.5%, highlighting also here a certain evolutionary framework, with important repercussions on the labour market and on the needs to be guaranteed to the welfare system.

Among future demographic transformations, the ageing process in the South should be highlighted (Table 2). Although this geographic area still has a younger age profile, the mean age of its residents will move from 45.5 years in 2023 to 51.5 years in 2050 (median scenario), surpassing both the North, which will reach a mean age of 50.2 years in the same year, and the Centre, which will be 51.1 years at that time.

TABLE 2. MEAN AGE OF THE POPULATION BY GEOGRAPHIC AREA, MEDIAN SCENARIO AND 90% CONFIDENCE INTERVALS Years 2023-2080, January 1st, in years and tenth of year (*).

Geographic area	2023	2030	2040	2050	2080
North	46,8	48,0	49,3	50,2	50,4
		[47,8 / 48,1]	[48,7 / 49,8]	[49,2 / 51,2]	[47,9 / 53,2]
Centre	47,0	48,5	50,1	51,1	51,0
		[48,4 / 48,6]	[49,6 / 50,6]	[50,1 / 52,1]	[48,5 / 53,8]
South	45,5	47,5	49,9	51,5	52,1
		[47,3 / 47,6]	[49,4 / 50,4]	[50,6 / 52,5]	[49,5 / 54,9]
ITALY	**46,4**	**47,9**	**49,6**	**50,8**	**50,9**
		[47,8 / 48,0]	[49,1 / 50,1]	[49,8 / 51,8]	[48,4 / 53,8]

(*) Values under the confidence intervals in square brackets.

In 2043 there will be about 27 million households in Italy

Over the next 20 years, an increase of about 930,000 households is expected: from 26 million in 2023 to 26.9 million in 2043 (+3.5%). These are increasingly smaller households, characterised by greater fragmentation, whose average number of members will fall from 2.25 in 2023 to 2.08 in 2043. Households with at least one family (i.e. characterised by the presence of at least one couple or parent-child relationship) will also vary their average size from 2.94 to 2.79 members.

The increase in the number of households will mainly result from a growth in households without a family nucleus (non-family households) (+16%), which will rise from 10 million to 11.5 million, accounting for 42.9% of total households in 2043 (in 2023 they were 38.3%). On the contrary, family households (with at least one family nucleus) will decrease by more than 4%: these households, which in 2023 amounted to 16.1 million (61.7% of the total), will fall to 15.4 million in 2043, thus constituting only 57.1% of households (Table 3).

Such a decline in households with nuclei stems from the long-term consequences of Italy's ongoing socio-demographic dynamics. Indeed, the aging of the population, with an increasing life expectancy, generates more single people, the prolonged decline in the birth rate increases the number of people without children, while the rise in marital instability, as a result of the greater number of couples' dissolution, results in a growing number of single persons and single parents.

In 20 years' time four out of 10 households will consist of only one person.

Increasing life expectancy and marital instability mean that the number of people living alone, real 'micro-families', will grow overall by 15 per cent, increasing from 9.3 million in 2023 to 10.7 in 2043. This increase, both absolute and relative, explains to a greater extent the overall growth in the total number of households.

For one-person households the gender differences are substantial. Men living alone will see a 10% increase from 4.2 to 4.7 million in 2043. Single women are expected to grow even more (+20%), leading to an increase from 5.1 to 6 million.

Single-person households, due to their age composition, have an important social impact, considering that it is especially in older age groups that lone persons increase significantly.

Already in 2023, among the 9.3 million single persons, those aged 65 and over will amount to 4.4 million, constituting 47.5 per cent of the total. In the years to come, the proportion of over-65s in the total number of single-person households grows substantially. In 2043, thanks to a growth of as much as 40 per cent, the number of people over 65 living alone will reach 6.2 million, making up 57.7 per cent of the 10.7 million people expected to live alone.

TABLE 3. NUMBER OF HOUSEHOLDS BY TYPE. Years 2023*, 2033, 2043, median scenario, thousand and percentage

	Year 2023	Year 2033	Year 2043	Year 2023	Year 2033	Year 2043
	ABSOLUTE VALUES			PERCENTAGE VALUES		
TOTAL HOUSEHOLDS	26.018	26.724	26.930	100,0	100,0	100,0
with at least one family nucleus	***16.053***	***15.916***	***15.385***	***61,7***	***59,6***	***57,1***
Childless couples	**5.273**	**5.714**	**5.878**	**20,3**	**21,4**	**21,8**
Couples with children	**7.756**	**6.990**	**6.203**	**29,8**	**26,2**	**23,0**
with at least a child under 20	*5.007*	*4.240*	*3.863*	*19,2*	*15,9*	*14,3*
Single mother with children	**2.217**	**2.305**	**2.328**	**8,5**	**8,6**	**8,6**
with at least a child under 20	*953*	*966*	*1.020*	*3,7*	*3,6*	*3,8*
Single father with children	**498**	**599**	**672**	**1,9**	**2,2**	**2,5**
with at least a child under 20	*188*	*208*	*231*	*0,7*	*0,8*	*0,9*
Households with 2+ family nuclei	**309**	**307**	**303**	**1,2**	**1,1**	**1,1**
without family nuclei	***9.965***	***10.808***	***11.546***	***38,3***	***40,4***	***42,9***
Lone persons	**9.306**	**10.062**	**10.737**	**35,8**	**37,7**	**39,9**
Male lone persons	*4.249*	*4.535*	*4.689*	*16,3*	*17,0*	*17,4*
Female lone persons	*5.057*	*5.527*	*6.047*	*19,4*	*20,7*	*22,5*
Multiperson households	**659**	**746**	**809**	**2,5**	**2,8**	**3,0**

(*) The multipurpose survey on "Daily Life Aspects" data are released on a two-year average. Here, the data refer to 1st January. For 2023 this can give rise to differences. Since 2023, the survey has also introduced a new weighting methodology, in order to make its results more consistent with those of the Permanent Population Census (more details in Methodological Note).

About 4.1 million lone persons will be over 75 years old

Up to the age of 64, the condition of living alone, whether voluntary or not, now involves 4.9 million individuals, 60.5% of whom are men. Over the next 10 years this segment of the population is set to remain fairly stable (4.8 million in 2033). In the following decade, on the other hand, in line with the overall decline that will characterise the population in adulthood, the number of single persons under 64 years of age will also start to decline, leading to 4.5 million by 2043.

Also beyond the age of 65, living alone has a specific gender distribution. However, in contrast to the previous example, here women predominate numerically, in relation to the well-recognised survival advantage they can benefit from. If already in 2023 the number of lone women over 65 amounted to 3.1 million, in the next 20 years this will rise to 4.3 million, an increase of 38%. Among lone men aged 65 and over, on the other hand, a growth of 600,000 (+45%, from 1.3 million to about 1.9 million) is expected by 2043, which will contribute to keeping the gender ratio stable at about seven women and three men for every 10 lone persons aged 65 and over.

The fact of living alone, often not dictated by a voluntary lifestyle choice, can affect the level of autonomy of the very elderly. If, in fact, it is less and less frequent for individuals aged 65 and over to experience limitations in their functional capacities, the problem is quite different when the 75 age threshold is exceeded, when they are more subject to specific needs and frailties linked to ageing. The number of people over 75 who could be living alone, in particular, is set to rise by more than 1.2 million (including 860,000 women) over 20 years, reaching 4.1 million lone individuals (including 3 million women) in 2043.

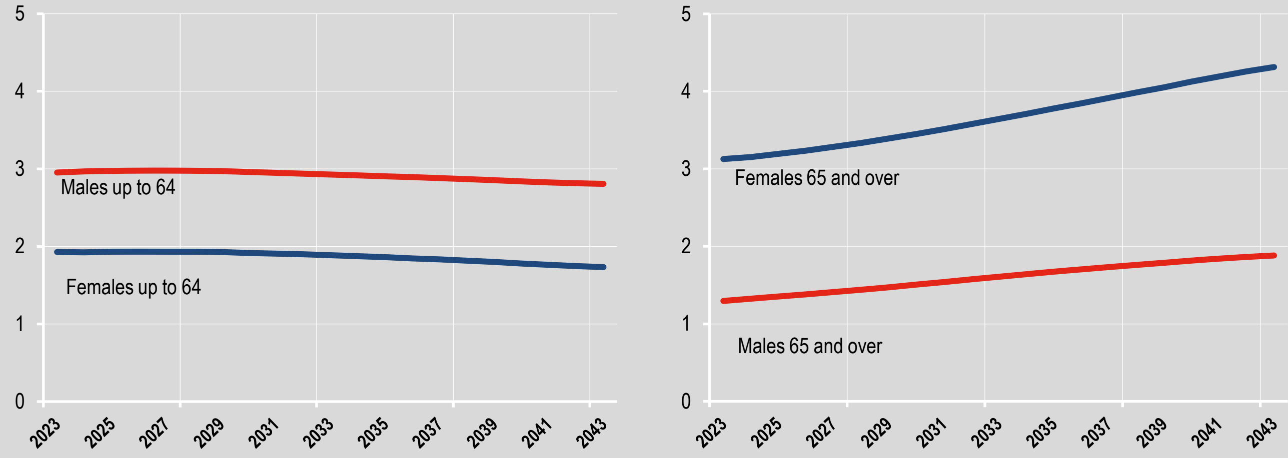
Couples with children decreasing

As a result of the prolonged low fertility rate and based on the assumptions considered in the median scenario, the decline in couples with children is expected to continue. This household type, which today accounts for almost three out of 10 households (29.8%), could fall to less than a quarter of the total in 2043 (23.0%).

Between 2023 and 2043, the number of couples with children will decrease from 7.8 to 6.2 million households (-20%). The largest decrease will occur among couples with at least one child aged between 0 and 19 years (-23%): this household type, which today gathers five million families, will decrease to 3.9 million in 2043, with a share of the total falling from 19.2% to 14.3%.

On the contrary, childless couples are expected to increase from 5.3 million in 2023 to 5.9 million after 20 years (+11%). Their share in the total number of households thus grows from 20.3 to 21.8 per cent.

FIGURE 2. **LONE PERSONS BY SEX AND AGE-CLASS. MEDIAN SCENARIO, ITALY.** Years 2023-2043.



Increase in childless couples and lone parents

The opposite path between couples with and without children, the former decreasing and the latter increasing, leads to a reduction in the numerical distance between the two types of couple from 2.5 million in 2023 to 325,000 in 2043. This structural change, which announces an overtaking of the latter over the former in the near future, could be realised in some areas of the country in the near future. In the North, for example, childless households could overtake those with children from 2040 (in the North-east as early as 2037), while in the Centre the surpassing would take place in 2043.

The increased spread of marital instability in the country will lead to a moderate increase in single-parent households, which will rise from 10.4 per cent of total households in 2023 to 11.1 per cent in 2043. Hindering the growth of this family type is both the low fertility rate and the tendency of lone individuals to regroup in other families or form new couples. In 2023, there will be 2.7 million single parents: 2.2 million mothers and 500,000 fathers, representing 8.5 per cent and 1.9 per cent respectively of the total number of households. After 20 years, the total number rises to three million. Among these, single fathers, remaining a minority compared to single mothers, increase to 670,000 (2.5% of total families) while single mothers grow to 2.3 million (8.6%).

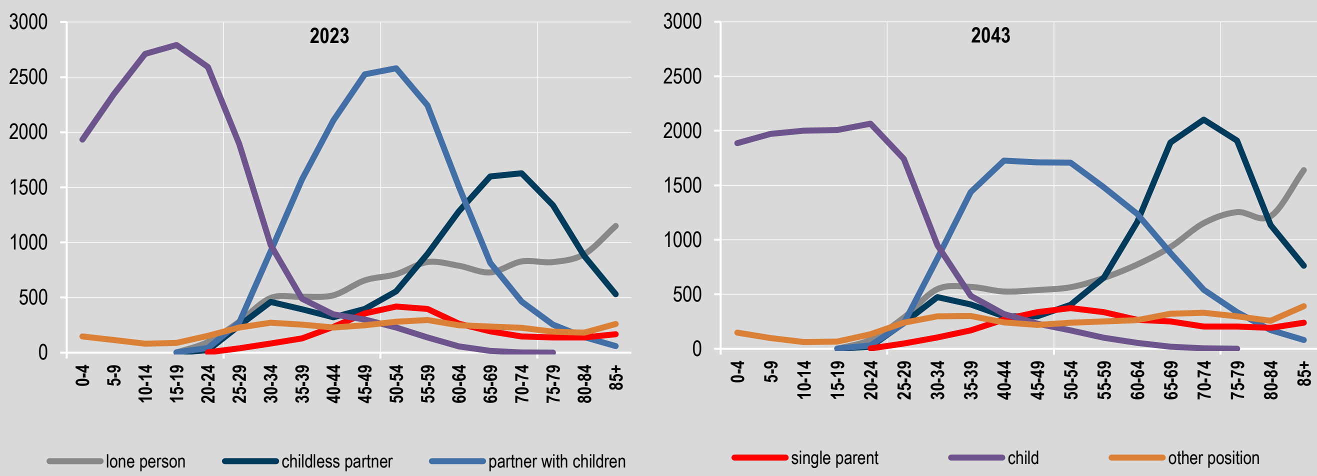
The household roles change

The effects of demographic changes also affect the roles played by people within households. Low fertility, for example, has an effect on both the presence of parents and children, just as an increase in the elderly population translates into a growth in single-parent families. The demographic scenarios for the next 20 years do not suggest any change in trends that would reverse these trajectories.

Analysing, therefore, the roles assumed by people over the course of their lives, we can observe the main changes that will characterise the next 20 years. As a result of low fertility, there will be significantly fewer children: between the ages of 15 and 19, the peak of 2.8 million children in 2023 will fall to just over two million in 2043. People in couples with children will in particular decrease between the ages of 50 and 54 (from 2.6 million to 1.7), while between the ages of 70 and 74, people in couples without children will especially increase (from 1.6 million to 2.1). Finally, the number of people living alone today is at a peak of 1.1 million among the over 85s, but in 20 years' time this will increase to 1.6 million.

The sequence of transitions in the life cycle, however, does not seem to be affected by a profound change, especially among young people. Having passed the stage of adolescence and young adulthood, where today 89% of 20-24 year olds are in the position of child (88% in 2043), in the 25-29 age group, and again in the 30-34 age group, being a child represents and will continue to represent the prevailing existential condition (respectively 64% and 31% in 2023 against 63% and 30% in 2043). Thus it is only from the 35-39 age group that the condition of being a couple with children prevails (47% in 2023), partly expected to decrease in the years to come (43% in 2043).

FIGURE 3. PEOPLE BY HOUSEHOLD POSITION AND AGE GROUP, MEDIAN SCENARIO, ITALY. Years 2023 and 2043, thousand.



Household fragmentation will affect the whole country

Households change over time as a result of social and demographic behaviour specific to different parts of the country, often highlighting territorial differences in today's and tomorrow's perspective.

As early as in 2023 in the North the share of family households (with at least one family nucleus) is lower than in the South (60.6% and 64.8% respectively). The expected reduction for this type of households is substantial, so that in 2043 they could make up 56.3% of total households in the North and 59.7% in the South. For the Centre, family households (59.5% in 2023) will fall to 55.4% of total households in 2043, approaching the share expected for the North.

The household type undergoing the greatest change over the next 20 years is the couple with children. The South will maintain the highest proportion for this type of household, although it is expected to fall by more than seven percentage points (from 33.2% in 2023 to 25.4% in 2043). Similar reductions will occur in the Centre (from 27.8% to 20.6%) and in the North (from 28.4% to 22.5%).

Most of the decline affects couples with at least one child up to the age of 19, compared to couples with only children aged 20 and over. The former fall from 18.8% to 14.9% in the North and from 18.0% to 13.3% in the Centre. In the South, couples with at least one child under 20 years of age drop by about 7 percentage points. For couples with ‘young’ children, therefore, there is a process of territorial convergence, while for couples with ‘mature’ children, there is still a significant positioning of the South, which is linked to the later leaving the family of origin in this area.

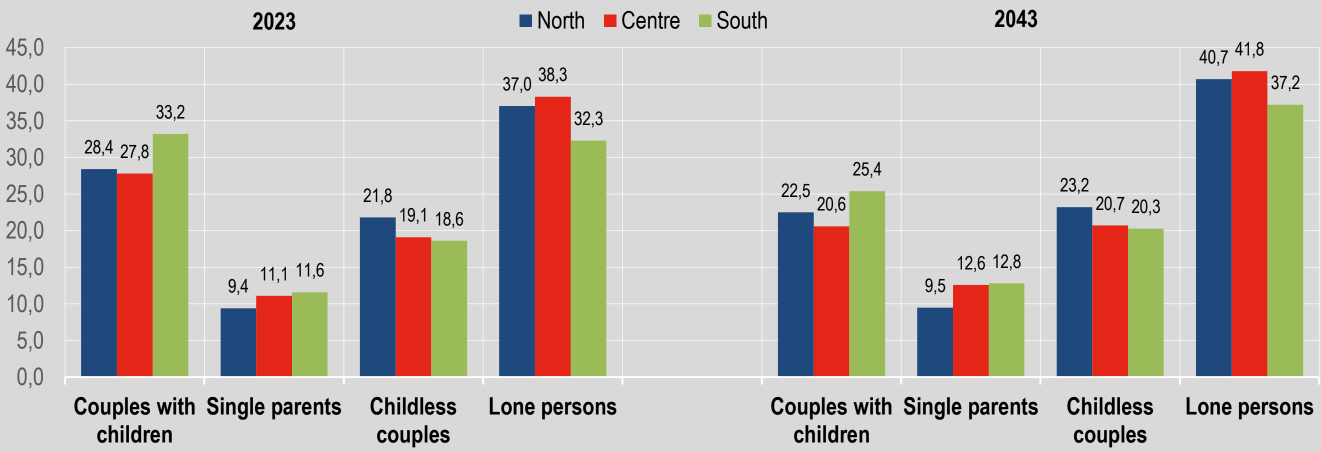
For Italy as a whole, lone persons will grow from 35.8% to 39.9%, but in the territory this average value stems from different realities: in the Centre there is a higher presence of lone persons in 2023, 38.3%, growing to 41.8%. In the North, the starting figure is lower (37.0%), but projections show that in 2043 more than four out of 10 people will live alone (40.7%). In the South, which starts from a lower initial share of lone people at 32.3%, a faster development is expected, leading this family type to represent 37.2% of households in 2043 (Figure 4).

Childless couples are growing in all areas of the country, but will continue to be more widespread in the North where, already accounting for more than a fifth of households (21.8%), they will achieve a further increase at the end of the period (23.2%). In the South, childless couples, accounting for 18.6% of households in 2023, will reach 20.3% after 20 years.

Lone parenthood is less common today in the North (9.4%) and is expected to remain fairly stable over time. In the Centre and the South, on the other hand, where levels are already above 11%, there will be a slight increase of less than two percentage points.

The combination of the expected family changes will cause the average family size to fall nationally (from 2.25 to 2.08 components) and territorially. The North and the Centre, both of which today have a value of around 2.2 members, will in 20 years' time face an average size close to 2 members. The South, today around an average size of 2.4 components, will move closer to the rest of the country with an expected average of 2.1 components.

FIGURE 4. HOUSEHOLDS BY MAIN TYPE AND GEOGRAPHIC AREA, AVERAGE SCENARIO. Years 2023 and 2043, percentage values.



Glossary

Age specific fertility (rate): the ratio of the number of live births to women between the ages of x and x + 1 and the average number of women of that age in a given year.

Average number of children per woman: the number of children a woman would have if she was subjected to the fertility calendar (in the form of age-specific fertility rates) of a given calendar year during her reproductive life span.

Birth (rate): ratio between the number of live births in the year and the average amount of the resident population, multiplied by 1,000.

Cohort component (model): the continuous calculation algorithm that in iterative mode simulates the evolution of the fundamental population equation by age group, allowing to determine the demographic flows and to obtain the surviving population at the end of each year.

Couple: two people linked by an emotional and sentimental relationship. Can be formed by opposite or same sex people. The bonds between people in couples can be formal (de jure couple: married, civilly united or de facto cohabiting pursuant to Law 76/2016) or informal (de facto couple).

Death: the cessation of any sign of life at any time after the vital birth.

Demographic projection: elaboration that shows the future development of a population when certain assumptions are made regarding the future course of mortality, fertility and migration.

Deterministic demographic projection: elaboration on the future development of a population, summarized in a single series of values obtained from a single set of demographic assumptions, which does not report any measure regarding the uncertainty usually associated with the results.

Dependency ratio: ratio between the population of inactive age (0-14 years and 65 years and over) and the population of active age (15-64 years), multiplied by 100.

Elderly dependency ratio: ratio between the population aged 65 and over and the population aged 15-64, multiplied by 100.

Emigration for abroad (rate): the ratio between the number of emigrations to abroad and the average amount of the resident population, multiplied by 1,000.

Family households (or households with at least one family nucleus): includes couples with children, couples without children, single parents, families with two or more nucleus.

Non family Households (or households without nuclei): people living alone or multi-person families; this latter do not constitute a family unit even if composed of several people.

Family nucleus: group of people linked by ties of marriage, kinship, affinity, adoption, protection, or by emotional ties, cohabitants and having habitual residence in the same Municipality. It can also be constituted by a single person.

Household: group of persons who habitually live in the same household and are bound by marriage, civil partnership, kinship, affinity, adoption, guardianship or emotional ties and who share in the household's income (by contributing to or benefiting from it) and daily expenses. Within a household there may be one or more family nucleus, but there may also be none, as in the case of households formed by an isolated member (one-person household) or several isolated members (multiperson household).

Household typology: classification based on the presence or absence of at least one nucleus and by type of nucleus.

Immigration from abroad (rate): the ratio between the number of immigrations from abroad and the average amount of the resident population, multiplied by 1,000.

Internal emigration (rate): the ratio between the number of internal emigrations and the average amount of the resident population, multiplied by 1,000.

Internal immigration (rate): the ratio between the number of internal immigrations and the average amount of the resident population, multiplied by 1,000.

Internal migration balance: difference between the number of registrations for change of residence from another Municipality and the number of de-registrations for change of residence to another Municipality.

Internal net migration (rate): the difference between the internal immigration rate and the internal emigration rate.

Life expectancy at age "x": the average number of years that a person of completed age "x" can count to survive in the hypothesis that, in the course of his subsequent life, he was subjected to the risks of mortality by age (from age "x" up) of the year of observation.

Life expectancy at birth: the average number of years that a person can count to live from birth in the hypothesis that, in the course of his existence, he was subjected to mortality risks by age of the year of observation.

Live birth: the product of conception which, once expelled or completely extracted from the maternal body, regardless of the duration of gestation, breathes or manifests other signs of life.

Mean age: mean age of the population at a certain date expressed in years and tenths of a year.

Mean age at birth: the mean age at birth of mothers expressed in years and tenths of a year, calculated considering only live births.

Migratory balance with abroad: difference between the number of registrations for change of residence from abroad and the number of de-registrations for change of residence to abroad.

Mortality (rate of): ratio between the number of deaths in the year and the average amount of the resident population, multiplied by 1,000.

Natural balance: difference between the number of births and the number of deaths.

Natural growth (rate): the difference between the birth rate and the death rate.

Net migration with abroad (rate): the difference between the immigration rate from abroad and the emigration rate with abroad.

Old age (index): ratio between the population aged 65 and over and the population aged 0-14, multiplied by 100.

Predictive (or confidence) interval: an interval associated with a random variable yet to be observed, with a specific probability that the random variable falls within it.

Probabilistic demographic projection: elaboration on the future development of a population, summarized in a set of values or in a probability distribution, in which the variables used are of a random nature that cannot be predicted with certainty and in which not all assumptions are equally probable.

Probability of death: the probability that an individual of precise age x will die before the birthday x + 1.

Projection: development expected in the future.

Projection probability of death: the probability that an individual of age x (in years completed on 1st January) will not survive within the year.

Projection probability of interregional migration: the probability that an individual of age x (in years completed on January 1st) moves residence between two regions before the end of the year.

Range: measure of the variability of a quantitative phenomenon defined by the difference between its maximum and the minimum value.

Registration and de-registration for transfer of residence: registration concerns people who have moved to a Municipality from other Municipalities or from abroad; the de-registration concerns people who have moved to another municipality or abroad.

Resident population: constituted in each Municipality (and similarly for other territorial divisions) of people with habitual residence in the Municipality itself. Persons temporarily residing in another Municipality or abroad, for the exercise of seasonal occupations or for reasons of limited duration, do not cease to belong to the resident population.

Scenario approach: the description of the context, even conceptual, in which the population is projected. In a deterministic approach it usually refers to the main or central assumption. In a stochastic it can refer to the assumption identified as mean or median.

Simulation: the quantitative implementation of a single set of demographic assumptions to be launched in the cohort-component model in order to obtain a single set of demographic projections.

Total balance: sum of the natural balance and the total migratory balance.

Total growth (rate of): the sum of the total net migration rate and the natural growth rate.

Total migratory balance: the sum of the migration balance with abroad and the internal migration balance.

Total net migration (rate): the sum of the net internal migration rate and the net migration rate with abroad.

Methodological note

1) Regional population projections by age and sex. Years 2023-2080

Istat regional demographic projections are built with the aim of representing the possible future trend of the population, both in terms of total numbers and in terms of age and sex structure. The information produced represents an important tool to support decisions in economic and social policies, such as those relating to pension, health, education and housing systems. The projections are periodically updated by reformulating the evolutionary assumptions underlying fertility, survival, international and internal migratory movements.

The new set of projections replace those based on 2022 published by Istat in September 2023. Istat is the owner and responsible for the production and dissemination of the projections, as documented in the National Statistical Program. With this new release a new three-year production cycle starts after those of the years 2016-2018 and 2019-2021[1]. This methodology was defined, between 2009 and 2015, by a working group with researchers from Istat and the Luigi Bocconi University of Milan.

The methodological approach, around which the forecasting model works, is of a semi-probabilistic nature. The fundamental characteristic of probabilistic forecasts is to consider the uncertainty associated with the predicted values, determining the confidence intervals of the demographic variables and giving the user the possibility to independently choose the degree of confidence to be assigned to the results.

Compared to the "deterministic" approach, more widely used on an international scale and also adopted by Istat in the past (up to the 2011 based projections), this represents a significant methodological advance. In fact, in the deterministic model the user does not have probability measures associated with the results. Thus, a further advantage of the probabilistic method is the fact that the user can stop to trust uncritically on the work of projection makers, who with the typical "low / high" variants define a priori the alternative boundaries to the variant retained "most likely", generally identified as "main" or “medium” or “central” scenario".

The quantification of uncertainty does not represent the only advantage of the probabilistic model. Another one is the more effective representation of the evolution of a population. In the probabilistic model, in fact, the definable scenarios are infinite on the theoretical level (although in reality, as will be seen later, a finite number is always selected), so assumptions of low survival are mixed with assumptions of high fertility or medium level of migration, or the opposite. Instead, the assumptions of the high/low scenarios of the deterministic approach are defined by pursuing an output oriented logic: the high scenario contemplates assumptions of maximum increase in survival, fertility and migrations, while, on the contrary, the low scenario contemplates only assumptions of minimum. The construction of such opposing scenarios actually captures the goal of determining a future range for the population and its structural components, but based on concomitant assumptions with low chance of occurring.

The subsequent sections contain general information and briefly illustrate the steps that made it possible to build the projections. These sections include information on the following aspects:

- base population
- projection technique
- time horizon
- panel of experts
- expert questionnaire and probabilistic model
- relationship between national and regional projections
- data
- corrective component of nowcasting for short term assumptions
- confidence intervals and median scenario
- regional fertility projections
- regional mortality projections
- regional projections on international migration
- regional projections on internal migration
- comparison with previous projections

1 The 2019-base forecasts, theoretically belonging to the second production cycle, skipped due to the unavailability in the required timeframe of the base population, for the first time in history derived from the new Permanent Population Census, as well as the need to produce a historical population reconstruction for the period 2002-2018 and to recalculate all the demographic indicators of reference for the forecasts (in particular, probability of death and specific fertility rates).

- comparison with the projections released by Eurostat and the United Nations
- data dissemination and terms of use
- contact information and personalized data requests.

Base population

The base population is the one broken down by sex, single age group and region as of 1 January 2023, as identified by the last Census of Population and Housing. The population includes all people usually residing in Italy, of any citizenship, while it does not include Italian citizens residing abroad, nor citizens illegally or irregularly present on the national territory who are not enrolled in any municipal register.

Projection technique

Projections are carried out with an iterative technique between 1 January and 31 December of each year, using the so-called "cohort-component" method. To the initial population, in correspondence of each age group, immigrations (from abroad or from other regions) are added while deaths and emigrations are subtracted (for abroad or for other regions), thus obtaining the population alive at the end of the year. Live births in the course of the year have also to be computed and, among them, those still alive as of December 31st, net of deaths and migratory movements that concern them.

For the population (stock), age is defined in completed years on 1 January (from 0 to 110 years and over). The same concept applies for flow data (births, deaths and migratory movements). This allows to identify, always and in any case, the demographic events by single year of birth of the subjects involved, ensuring the required consistency within the population equation.

It is assumed that demographic events may occur linearly at any time of the year. Between death and migration (internally or abroad) it is assumed incompatibility, that is, they cannot involve the same individual in the same year.

Deaths are determined by multiplying the resident population by age group on 1 January by the respective (projection-)probabilities of death, i.e. those involving subjects belonging to the same cohort.

Births in a given year are achieved in three steps. In the first, the average number of women for each fertile age (obtained as the average of the populations of that age at the beginning and end of the year) is multiplied by the respective fertility rate. In the second, the sum of the births by age of the mother is calculated, obtaining the total number of births in the year. In the third, births are broken down by sex using the fixed ratio of 106 male births per 100 female births.

Projections have a territorial profile and are built in the logic of the multi-regional model, a model which, with particular regard to internal migratory flows, simultaneously and coherently works the distinct territorial units of reference. The model on internal migration starts from the construction of a multi-regional matrix of migration probabilities by region of origin, region of destination, sex, and age. This matrix, applied to the population at risk of migration, identifies a coherent series of immigrants and emigrants in each forecasting year.

Time horizon

Projections cover the period between 1 January 2023 and 1 January 2080. The main purpose is to provide with information on the future development of the population in the short term (2030), and therefore to provide with information in the medium (2050) and long term (2080). With regard to this latter time reference information, data should be used with caution since the results become the more uncertain the further we go from the base year (2023). This risk is the more concrete the more attention is paid to the smaller territorial units, as in the case of some Italian regions.

Panel of experts

A panel of national experts supported Istat in formulating the demographic assumptions for Italy as a whole. The assumptions for the regions, on the other hand, were handled by Istat on the basis of a specific "bridge" methodology between the national and regional assumptions. The experts who replied to the questionnaire (with CAWI technique), providing with useful and complete information to define the assumptions, were 121. They were voluntarily recruited by Istat among the most illustrious experts in demographic-social studies. In particular, there are 69 women and 52 men, mainly employed in universities (68) or in other public research bodies (42). The mean age of the respondents is 51 years while their work experience is on average 23 years.

The survey was carried out in the first quarter of 2023, and its long-term findings are taken as a reference for the entire three-year cycle (i.e. for the forecasts in base ‘22, released last year, for those released here in base ‘23, and for those to be released in 2025 in base ‘24).

Expert questionnaire and probabilistic model

The probabilistic method adopted is based on the opinions of experts (expert-based model) to define the future evolution of the most important demographic indicators. It falls within the broader class of random scenario models. This model, used for the definition of probabilistic scenarios at a national level, is based on the elicitation of a series of parameters from which the future stochastic evolution of each demographic component is derived. Experts are asked to provide values at a given year "t" with regard to a series of summarized demographic indicators, conditional on the values assumed by the same indicators in instants of time prior to year "t" (Billari, Graziani and Melilli, 2012).

The method has the advantage of being simple and flexible. In fact, in the questionnaire, the necessary demographic components are summarized through the following indicators: the average number of children per woman; life expectancy at birth by sex; immigration and emigration from abroad. The other information necessary for the production of the projections, such as that regarding the age-breakdown of demographic events, is purposely kept out and subsequently processed in order to make the questionnaire and the forecasting model itself parsimonious.

TABLE A1. MEAN VALUES, VARIANCES AND CORRELATIONS UNDER ASSUMPTIONS OF THE EXPERTS BY DEMOGRAPHIC INDICATOR. Years 2021, 2050 and 2080

Indicator	Total fertility rate	Life expectancy at birth – Men	Life expectancy at birth – Women	Immigrations (thousand)	Emigrations (thousand)
Year 2021					
Observed value	1,25	80,3	84,8	318	158
Year 2050					
Mean assumption	1,38	84,3	87,8	302	136
High assumption	1,54	85,7	89,1	368	169
Variance	0,016	1,239	1,106	2.613	667
Year 2080					
Mean assum. conditional to 2050 mean assumption	1,45	86,2	89,6	304	142
Mean assum. conditional to 2050 high assumption	1,66	88,0	91,1	389	187
High assum. conditional to 2050 mean assumption	1,68	88,2	91,5	402	192
Varianza	0,058	4,586	3,689	10.302	2.774
Correlation 2050-2080					
Correlation coefficient	0,67	0,67	0,64	0,65	0,67

Two time points are considered for each demographic indicator: an intermediate year "t1" and a year "t2" corresponding to the last forecasting year. In the questionnaire submitted to the experts, "t0 = 2021", "t1 = 2050", "t2 = 2080", thus generating two sub-intervals, 2021-2050 and 2050-2080. Expressing the value of life expectancy at birth in the year 2080, given its expected value in 2050, is a practical example of how the mechanism works.

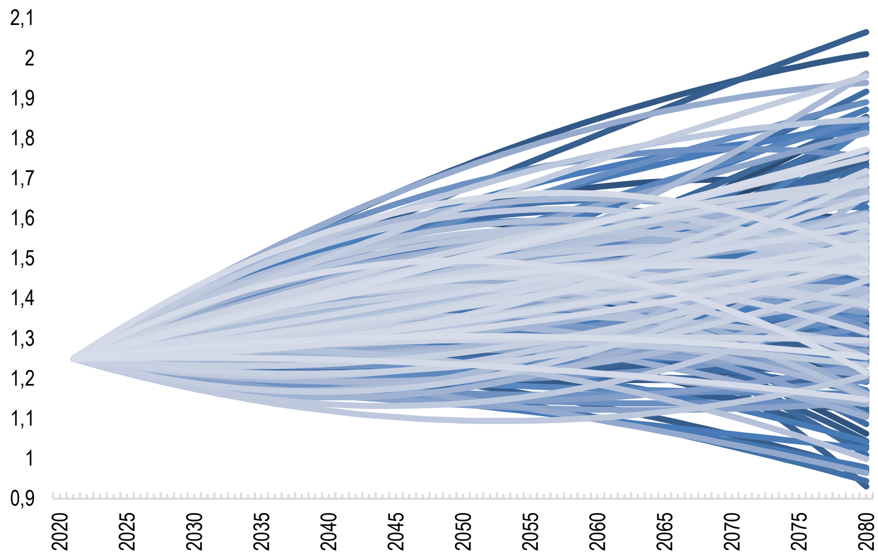
The demographic indicators are assumed, for the sake of simplicity, independent of each other (for example, the total fertility rate is not influenced by the level of migration and vice versa), although the model allows in its generalized version the possibility of interacting among them. It is also assumed that the pair of elicitations at 2050 and 2080 of a given indicator has a bivariate normal distribution.

Under these conditions, once the mean values provided by the experts have been obtained, it is possible to estimate the variance associated with each of the two future time instants as well as the correlation between the first and the second period (Table A1). On the basis of the corresponding bivariate normal distributions, 3,000 simulations were then carried out with the Markov Chain Monte Carlo method1.

The last estimation step is aimed at calculating the values of each parameter in the intermediate years of the two sub-intervals 2021-2050 and 2050-2080. This activity is carried out, for each of the 3,000 simulations, by interpolation with quadratic curves, passing through the known points corresponding to the years 2021, 2050 and 2080. Thus, the definition of 3,000 stochastic curves for each demographic indicator has been achieved at national level. As an example, Figure A1 describes the bundle of trajectories relating to the number of children per woman, obtained from the procedure described above.

The choice to consider a number of 3,000 simulations is the result of a compromise between two needs, both strategic: 1) faithfully representing the uncertainty of demographic events; 2) optimizing the machine times for processing the projections. The latter, despite today's availability of increasingly powerful and sophisticated hardware / software tools, represents a technical aspect which is anything but secondary, given the huge amount of data processed.

FIGURE A1. PROBABILISTIC EVOLUTION OF THE TOTAL FERTILITY RATE IN 3,000 SIMULATIONS OBTAINED FROM EXPERT OPIONIONS. YEARS 2021-2080



The Limesurvey electronic questionnaire

To create the electronic questionnaire for collecting data from the experts, the Limesurvey tool was used, an open-source software distributed under the GNU General Public License (GPL) and created on the LAMP platform, for the creation and management of online surveys and questionnaires.

It allows, through an intuitive user interface, the rapid development of web questionnaires and the management of all the subsequent phases of data collection, from the creation of the list of respondents and the related contact methods, to the monitoring of the survey, up to the export of the answers. Various types of questions are allowed, it supports multilingual surveys, and is fully graphically customizable via templates with responsive layouts, i.e. whose content adapts to the dimensions of the browser of the device used.

Istat began using the software more than a decade ago, hosting it and periodically updating it on its servers (exposed and internal), for different types of direct data collection processes in a web environment.

The questionnaire in question, composed of a personal data section and 6 thematic sections (of which 5 reserved for demographic forecasts and 1 for household forecasts, see paragraph 2 below), was implemented faithfully respecting the consistency and validation checks between the values of the questions within the same section and between those belonging to different sections. To this purpose, it was necessary to reprogram (in javascript) the interaction with the user to force him to insert in the various questions, from time to time, an appropriate number of decimal figures, customize the final table on family positions, as well as configure some general graphic aspects.

The list of experts was pre-loaded in the system and each of them was randomly assigned a unique participation code (token), through which the link (URL) for completing the questionnaire is composed. Each participant received an invitation email in their inbox with this link together with a brief information; he had the possibility to access the questionnaire from any browser and from any device (including smartphones) and to complete the questionnaire even in different sessions.

Finally, the system allowed the Limesurvey administrative backend to repeatedly solicit respondents who had not accessed or completed the questionnaire.

Relationship between national and regional projections

The probabilistic model releases a set of 3,000 national simulations for each summary demographic indicator. Since the objective of the Istat projections is also to give indications at a territorial level, so continuing the longstanding tradition of the multi-regional model, a "bridge" procedure has been implemented between the definition of national and regional inputs. The approach pursued is therefore top-down on the side of the assumptions building while, as will be seen later, it is bottom-up on the side of the production of final outputs.

The main action is to derive 3,000 regional stochastic scenarios from the 3,000 national ones. The first operation in this sense is to elaborate an intermediate deterministic forecast, applying the multi-regional cohort component model. From this forecast, obtained by extrapolating the regional trends considered most probable for each component (see following paragraphs), the same summary indicators of the previously described stochastic model are obtained, i.e. average number of children per woman, male and female life expectancy at birth, migratory movements with foreign countries. Such a first intermediate forecast, unique and deterministic, essentially resembles that which in a deterministic approach would be labelled with the term "central scenario".

The transition from the regional deterministic model to the regional stochastic model is achieved by multiplying, and repeating the procedure 3,000 times, the regional deterministic forecast for the relationship between the national stochastic and the deterministic forecast. In formula, indicating with "n" the generic simulation (n = 1, ..., 3,000), with "j" the regional territorial code, with DR the deterministic regional forecast, with SR the stochastic one, with DN and SN, respectively, the national deterministic and stochastic forecast, we have:

$$\mathrm{SR}_{t,n}^{j} = \mathrm{DR}_{t}^{j} \times \frac{\mathrm{SN}_{t,n}}{\mathrm{DN}_{t}}$$

thus linking, in each simulation, the vector of regional values to the national stochastic reference value. Note that with regard to the synthetic indicators of immigration and emigration from abroad, we have:

$$\mathrm{DN}_{t} = \sum_{j} \mathrm{DR}_{t}^{j}$$

$$\mathrm{SN}_{t,n} = \sum_{j} \mathrm{SR}_{t,n}^{j}$$

Once the synthetic stochastic indicators have been obtained at the regional level, we move on to the construction of the inputs necessary for the application of the cohort-component method, i.e. the (projection-)probabilities of death by age and sex, the age specific fertility rates and the distribution of immigrants / emigrants from abroad by age and sex. The procedure therefore associates each summary indicator with its own sex-age breakdown. The latter, not treated in a stochastic way, is the one that derives from the regional deterministic model and, from simulation to simulation, adapted to the specific synthetic stochastic indicator.

The coupling of the 3,000 death probability vectors (each vector develops a number of elements equal to the "number of regions X age classes X sex X forecast years") with the 3,000 fertility vectors, and the same number on immigration and emigration from abroad and, finally, with the 3,000 O / D probability matrices of internal migration, it is randomly executed.

After introducing a corrective nowcasting component (see next paragraph) relating to the very first years of forecasting, the cohort component model is then run 3,000 times, thus obtaining the required outputs: population by age and sex, demographic flows by age and sex, plus the series of demographic indicators to support the analysis, from generic rates (birth, mortality, etc.) to structural indicators (mean age, dependency rations, etc.).

The results at national level (as well as those at geographical area level) in the context of each regional simulation are obtained by sum (bottom-up approach). Therefore, the amount of the expected population, deaths, and migrations, classified by age and sex, and births by age of the mother that are determined at the national level are the sum of the forecast regional trajectories. The assumed national levels relating to the summary indicators placed into dissemination, for example regarding life expectancy or the average number of children per woman, are recalculated ex-post on the basis of these regional summaries.

It should be noted that the stochasticity introduced at the regional level, borrowed top-down from the national one and limited only to summary indicators, may result not sufficient to reproduce the randomness of the various demographic events. This is particularly true in small areas where uncertainty tends to be relatively higher. For this reason, although the number of simulations still offers ample guarantee of representativeness of the variability on a regional scale, it is more appropriate to speak of a semi-stochastic approach when referring to regional projections.

A second observation concerns the fact that in the Istat model a generalized statistical treatment of the covariance between the Regions is excluded (for example: the forecast of increase / decrease in fertility in a given region how much it conditions or how much is in turn conditioned by the forecast of increase / decrease in another). To this solution, also excluded for reasons of parsimony of the model, another one was preferred, that of territorial convergence. In fact, the initial deterministic regional model, subsequently transformed into a stochastic model through the procedure described above, is built on the assumption of very long-term convergence (2122, well beyond the last year of projections) between the regions for each fundamental demographic component. This implies that the 3,000 regional stochastic scenarios represent 3,000 different hypotheses of convergence of demographic behaviours among regions.

The main hypothesis underlying the convergence is that the socio-economic and cultural differences currently existing between the regions are destined to disappear in the long term. Therefore, their progressive cancellation would also involve a generalized rapprochement of demographic behaviours. The idea of convergence is not new in demography and there are many examples of demographic forecasts that follow it (Eurostat and the UN, in particular), including past Istat ones. In Istat projections, convergence is understood as the progressive shift of a given demographic behaviour towards a very distant point in the future which represents the instant of full regional convergence (in the sense that at that point the values would be identical for the different regions), but that in reality it is far from being reached within the time horizon considered (2023-2080). In fact, it is correct in this circumstance to speak more of a model of semi-convergence than of a model of full convergence.

Data

The assumptions defined at the regional level in the preliminary deterministic model, before the transition to the stochastic model, were obtained by extrapolating future trends from the analysis of the observed time series. In particular, these assumptions were defined using the following data series:

- for fertility, the mother's age-specific rates for the period 1977-2021;
- for mortality, the (projection-)probabilities of death by age and sex for the period 1974-2021;
- for internal and international migrations, the changes of residence by age and sex of 2016-2019 and 2021 (without 2020).

Corrective component of nowcasting for short term assumptions

Before being launched at full capacity along the time horizon with the cohort-component method, the probabilistic projections incorporate a corrective nowcasting factor (from the term nowcast = forecast of the present). With this operation we intend to ensure that the forecast relating to the very first years is as much in line with the trend that emerged in the last period or in the last historical year (jump-off effect). This type of operation is particularly suitable in years characterized by sudden, and as such unpredictable, changes in the demographic situation. This is the case, as happened in 2020 and to a lesser extent in 2021, of the effects caused by the Covid-19 pandemic on all components of the demographic change. Not only, albeit primarily, on mortality, but also on birth rates and internal and international changes of residence.

Since the base population of the projections is that recorded as of January 1, 2023, it was necessary put in place some short-term correction of the predicted inputs that affected the first projections years. The correction, in particular, takes advantage of the information from the provisional demographic balance - Year 2023, which Istat released in March 2024[2]. With this, in fact, we want to take into account not only the events that characterized the 2023, but also the subsequent years within which it is assumed that the pandemic effects may end and allow the short term inputs to be in line with medium and long term ones .

From the computational point of view, the review of the short-term assumptions is carried out by applying correction factors. For example, let E_j be the number of demographic events predicted in the first year based on the median scenario in region j. Instead, let $\hat{E}$ be the observed value of such events or, in the absence of the actually observed value, the best estimate that can be obtained (for example, using nowcasting procedures or similar statistical models). The ratio:

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

represents the correction factor to be applied to the statistical measures that give rise to type "E" events in year "b" for region j. For example, if these events were the total number of births then the quantity:

$\hat{f}_{b,x}^{n,j} = r_b^j \cdot f_{b,x}^{n,j}$ with x=14, … , 50 e n=1, … , 3000

represents the series of fertility rates by age of the mother (n-th simulation) corrected for year "b". Similar considerations apply to the determination of the correction coefficients relating to mortality and migratory movements. As regards 2023, the correction factors were constructed by comparing the data of the provisional demographic balance of each region, released in March 2024 by Istat, to the projections previously produced for that year [3].

2 Cfr.: Istat, la dinamica demografica – year 2023, https://www.istat.it/it/archivio/295586

For the years after 2023, the correction factors are applied for a limited period of the time horizon, processing weights that progressively tend to one. In particular, the number of years for which the correction factor is applied to the series of interest is obtained from:

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

with ϵ arbitrary quantity, appropriately chosen to ensure that, on regional average, the number of years to guarantee the return from short-term to medium-long term projections does not exceed 5 years. At this point, the levels of the correction factors for the years following "b", for a total of "Y" years, are given by:

$$r_t^j = \frac{r_b^j \cdot \left(b + Y^j - t\right) + (t - b)}{Y^j} \quad \text{with } t = b, b+1, \ldots, b + Y^j - 1$$

Confidence intervals and median scenario

Once the calculation procedure inherent to the 3,000 regional simulations has been launched, uncertainty is calculated for all possible information levels, from the predicted population to the flow data, also broken down by age and sex. These margins of uncertainty depend in turn on the uncertainty inherent in the future levels of mortality, fertility and migration that are also made available. The dissemination of the results contemplates the release of only the confidence intervals of 90%, 80% and 50% but it is possible to define intervals on any scale of interest. The confidence interval provides information on how likely it is that a given demographic indicator falls within predetermined limits. From this point of view it should be remembered that this probability itself represents a forecast, as it is based on hypotheses whose validity is uncertain. Furthermore, in no case should the extremes of the confidence interval be interpreted as extreme limits, upper or lower, of future demographic behaviour.

The construction of a confidence interval is here based on the determination of the percentiles in the distribution of the 3,000 simulations. For example, the 90% confidence interval for a given indicator is determined by considering the distribution values that fall between the 5th and 95th percentiles. It is also recalled that the uncertainty always refers to the domain of the specific estimated parameter. The limits of the confidence interval for a given hierarchical level are estimated on their own, and not constructed by summation of limits obtained at a hierarchically lower level of disaggregation. The criterion is also applied in non-territorial hierarchical contexts; for example in the composition by age of the population or in that by sex.

The "median scenario" was built with the aim of defining a "punctual" forecast that can be adopted as the most likely reference of future demographic evolution. This scenario corresponds to a 3001-th simulation, obtained by construction, but which in fact was not detected in the observation field of the 3,000 simulations. Its set of assumptions is identified by taking as a reference the median value between all the simulations at the level of the individual demographic components (fertility, mortality, migration) within the possible combinations of the covariates age, region and year of forecast. For example, the median scenario specific fertility rate at the age of 32 in the Tuscany region, in the year 2065, is identified as the median value with these characteristics identified among all the simulations. The same specific rate but at the following age, or in the following year, is identified with the same procedure but it probably arises from a different simulation. For the identification of the median scenario on mortality and migration, the procedure is identical but with the additional covariate of sex. Furthermore, for internal migrations, the selection also includes the region of origin and destination.

The scenario is therefore “median” from the side of the fundamental inputs. From the point of view of the outcome (population and expected flows) that this scenario generates once the procedure for cohort-components has been launched, for the typical properties of the median it returns values very close to the median ones.

Regional fertility projections

For regional fertility the projections concerned the classic parameters of intensity and age-breakdown, i.e. the average number of children per woman and the distribution of specific fertility rates by age of the mother.

The average number of children per woman was represented using ARIMA models (n, p, k), searching, separately for each region, the one most suitable for predicting the future intensity of reproductive behaviour. On the basis of the 1977-2021 historical series the predominantly model was an ARIMA (2,0,0) with intercept.

The fertility age profile was modelled using a *quadratic splines* function system (Schmertmann, 2003). This model functionally describes the curve of specific fertility rates standardized as a function of three parameters: the age of onset of the fertile age α; the age P in which fertility reaches its maximum level; age H, subsequent to P, in which fertility is halved compared to the maximum level. By specific standardized fertility rate we mean the specific fertility rate normalized to the unit, where the value one corresponds to the maximum value observed within its age distribution.

The *quadratic splines* model fits five second-degree polynomials to the fertility curves. The final function is continuous with the first derivative also continuous. Moreover, thanks to suitable mathematical restrictions it is uniquely determined by the three parameters [α, P, H] mentioned above.

In practice, the prediction of the specific fertility rate is transformed into the prediction of the three parameters (through ARIMA models) which express it functionally, once the series has been estimated in the period 1977-2021. To do this, a hypothesis of convergence between the Italian regions was adopted, assuming that the territorial differences in terms of reproductive behaviour tend to decrease in the long term. From an operational point of view, full convergence was set in 2122. In particular, the convergence constraint provides that, from 2023 to 2122, the parameters of the regional vector [α, P, H] converge linearly to the values of a hypothetical national vector, specially designed for the operation.

Regional mortality projections

Regional mortality projections were produced using the Lee-Carter model (1992) in the variant proposed by Lee-Miller (2001), a model in which the adjustment procedure leads the fitted probabilities of death to reproduce precisely the observed level of life expectancy at birth, rather than the total deaths observed as in the original version. Furthermore, here the model is applied to death probabilities rather than to mortality rates of the original formulation.

The model approximates the logarithmic form of the probability of death using three synthetic parameters, one of which is related to the trend [k(t)] and two related to the age distribution [(a(x), b(x)].

As for fertility, also for mortality the construction of the model originates from the definition of a provisional reference scenario at national level. The forecast is determined by projecting into the future the only national trend parameter k(t), whose series is identified over the period 1974-2021, while the parameters a(x) and b(x) remain invariant over time in this phase. In particular, due to its substantial linearity, the k(t) parameter was projected to 2080 with a *random walk with drift*.

The assumptions at the regional level are derived from the provisional national reference scenario, by first estimating the regional values of the three parameters in 1974-2021 with the same methodology and, subsequently, by making each regional parameter converge to the corresponding national parameter at 2122. Therefore, as a consequence of the convergence process and unlike the classical approach of the Lee-Carter model, here the regional parameters a(x) and b(x) are also varied over time.

Regional projections on international migration

In order to capture the most recent trends, the regional projections of migratory flows with abroad focus the analysis only on the last five years, namely 2016-2019 and 2021. The year 2020 was deliberately censored to avoid incorporating into the forecasts the effects of the lockdown. This need, considering the complexity of predicting international migratory flows by resorting to analysis of long historical series, leads to use a very simple model. Without forgetting that at this level of operations it is a question of structuring an intermediate deterministic model, whose values are subsequently calibrated on the intensities produced by the expert-based stochastic model.

In the first year of the projection, the total values of immigration and emigration from abroad are set equal to the mean value observed over the last five years. In accordance with the general convergence framework of the deterministic model, it is therefore assumed that in each region inflows and outflows converge linearly in the long term (2122) at the same level, which is to the initial half sum of the two values.

Once the totals of inflows and outflows up to 2080 have been determined, the associated age and sex breakdown are derived by applying the Castro-Rogers model (Rogers and Castro, 1981) to the 2016-2019 and 2021 series. With this model it is shown that the characteristic age profile of migrants can be described, regardless of the intensity of the phenomenon, by a mathematical function composed of 4 additive components and up to 11 predictive parameters. These parameters, whose estimate in the observed period is produced thanks to a generalized procedure for non-linear models (category in which the Castro-Rogers function fully falls), are kept constant in the forecast period. The conclusive result is therefore that the global intensity of migratory flows with abroad may vary over time but on the basis of a constant composition by age.

Regional projections on internal migration

Interregional migrations are developed according to a multidimensional approach, which allows to simultaneously consider the areas of origin and destination of migratory flows, to define the entrances in a given area as the sum of the exits with that destination from all the other areas of the system. The system is by construction consistent for all the forecast years since the marginal row and column of the O/D matrix, corresponding respectively to the inflows and outflows in/from each region, give the same sum, corresponding to the amount total of movements within the national territory.

The probability of migration specific for age (110), sex (2), region of origin (21) and destination (21) represents the basic component of the O/D matrix composed of 110x2x21x21 = 97020 cells for each calendar year. The probabilities are estimated on the basis of the levels observed in the individual years of the 2015-2016 and 2021 period, censoring the year 2020 as in the case of international migration. The probability vectors thus obtained, at the level of each annuity, are subsequently modelled using the Castro-Rogers function.

Therefore, indicating with

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

a generic (projection-)probability of migration for an individual of age "x" and sex "s" between the region "i" and the region "j" relating to the year "t" (t = 2016, ..., 2019, 2021), is assumed that this represents a normal random variable with an average equal to the mean value of the five-year period and variance equal to the variance detected in the five-year period:

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

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

From the above mentioned random variables, 3,000 values are randomly extracted for each of the 97,020 elements of the O/D matrix, thus giving rise to the random creation of 3,000 different matrices. The O/D matrix relating to the median stochastic scenario is identified by taking as a reference the median value between all the simulations within the possible combinations of the covariates sex, age, region of origin and region of destination. This median matrix is also used with the preliminary purpose of producing the deterministic forecast of the population, prior to the transition to the actual stochastic model (see previous paragraph on the relationship between national and regional projections).

Note that in the context of each simulation (including the median scenario) the O/D matrix is assumed to be invariant over time. The hypothesis underlying the model is based, in fact, on maintaining a propensity for mobility that remains constant throughout the time horizon. This implies that internal migratory flows evolve over time only because of the variations affecting level and age structure of the population exposed to the risk of migration.

Comparison with previous projections

An assessment of the change that occurred between the last two rounds can be made by comparing the median scenarios of the projections based on ‘22 and ’23.

First of all, a rather limited difference should be noted between the total base population ‘23 (58 million 997 thousand) and that which had been estimated in the median scenario on the same date by the projections based on ‘22 (58 million 889 thousand).

On the side of projected flows in the common projection section (2023-2080), an assessment of 21.1 million births is captured in both exercises while the forecasts in base '23 show a higher number of deaths by 300 thousand (44.1 million vs. 44.4). Foreign migration entries, on the other hand, are more favorable for the exercise in base '23. The latter, in fact, presents in the common projective tract an overall net migration balance of 10.2 million individuals versus 9.8 in the base '22 forecast.

The difference between the final populations of the two separate forecast exercises is small (about 300,000 more as of January 1, 2080 for the median base 2023 scenario), confirming the substantial robustness of the base 2022 forecasts, despite the base population change and short-term adjustments on the budgetary components. In this respect, Table A2 highlights how the process of revisiting assumptions for all demographic components affected only the early forecast years.

TABLE A2. 2021 AND 2022 MEDIAN SCENARIO ASSUMPTIONS FOR THE MAIN DEMOGRAPHIC INDICATORS. Years 2023, 2030, 2050 and 2080.

Median scenario	Total fertility rate	Life expectancy at birth – Men	Life expectancy at birth – Women	Immigration (thousand)	Emigration (tohusand)
Anno 2023					
Base 2022	1,25	80,6	84,7	398	151
Base 2023	1,20	81,3	85,3	416	142
Anno 2030					
Base 2022	1,30	81,8	85,8	312	147
Base 2023	1,28	82,0	85,9	339	141
Anno 2050					
Base 2022	1,38	84,3	87,8	302	136
Base 2023	1,38	84,3	87,8	301	136
Anno 2080					
Base 2022	1,46	86,1	89,7	305	143
Base 2023	1,46	86,1	89,7	305	143

Comparison with the projections released by Eurostat and the United Nations

In order to compare the projections produced by Istat with those of other bodies, it makes sense to take as a reference the projections released by Eurostat and the United Nations Population Division (UNPD). For years, the statistical institute of the European Union has been carrying out the task of producing demographic forecasts on a regular basis for all member countries. The latest release is based on ’22, whose main reference scenario is the so-called baseline scenario. The UNPD, in turn, also produces demographic projections on a regular basis through the World Population Prospects, which include all the countries of the globe. In this latter case, the latest available exercise is based on ’22 and the main reference scenario is the so-called medium variant.

It should be noted in the introduction that, despite the comparability on the level of projective technique, the exercises produced by the two international organizations present some methodological differences compared to the Italian one. Among these, in the first place, the fact that the two international models examined here are uninational, i.e. they project the resident population in Italy as a whole without taking into account the demographic development of the regions.

Table A3 shows the main scenario assumptions compared. As regards migratory flows, the comparison is limited to net migration as both Eurostat and UNPD build the assumptions directly on this indicator (without distinction between immigrants and emigrants).

The UNPD projections present very limited assumptions in terms of net migration, not only in the initial projection period but over the entire projection horizon. In the medium and long term the assumptions continue to be rather differentiated between the various producer bodies. In particular, with regard to migratory movements, where compared to a UNPD that is rather cautious about Italy, Eurostat is opposed with a much more optimistic vision. This evidence is partly due to the Eurostat methodology, which, in addition to predicting the underlying evolution of net migration, incorporates an additive *replacement-migration* component into the model[3].

The assumptions on fertility are quite similar, although in the medium-long term Eurostat and UNPD produce more favorable forecasts than Istat. The assumptions on survival are also not particularly distant, however Eurostat and especially UNPD highlight very favorable expectations about the lengthening of life expectancy, which are only partially glimpsed in the Istat model.

The development of the different demographic assumptions therefore gives rise to differences in terms of expected results which, as regards the evolution of the total population, can be appreciated in Figure A2. The UNPD projections give a much more pessimistic evolution of the population which approximates the lower limit of

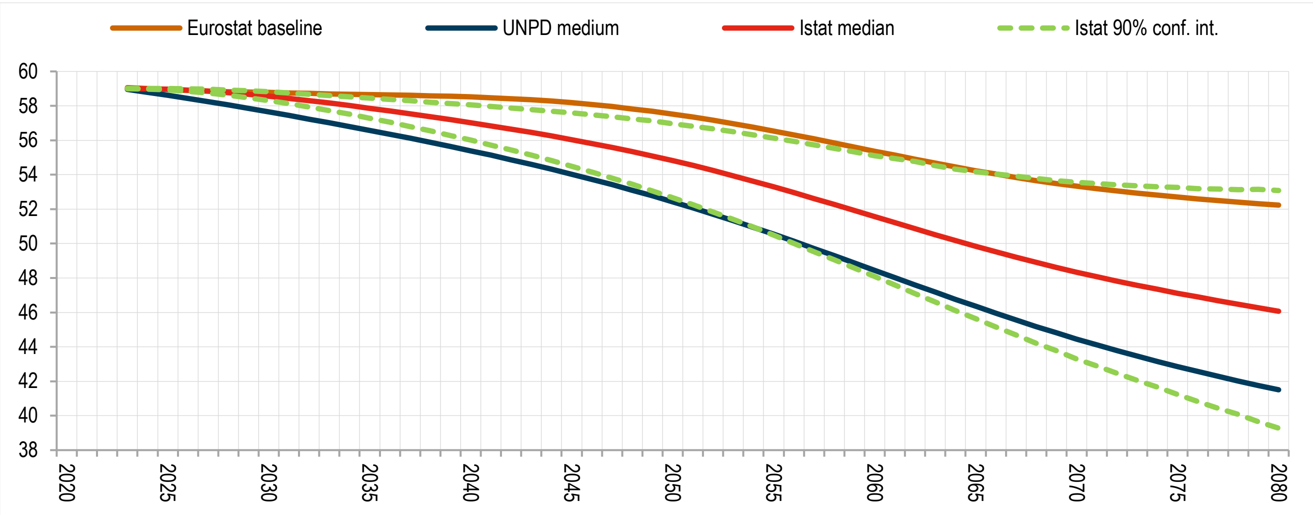
3 This component assigns in each forecast year an additional quota of net migrants in the measure equal to 10% of the reduction found in the population of working age (15-64 years).

Istat confidence interval. The Eurostat scenario, given the significant impact of a more sustained net migration, is particularly optimistic. Up to the point of maintaining a population even wider than the upper limit of the 90% confidence interval of Istat projections for most part of the time horizon. Nonetheless, the evolutionary trajectory of the population is consistent between the three scenarios. In fact, all of them foresee a progressive decline of the population, which tends to worsen in the medium-long term.

TABLE A3. COMPARISON BETWEEN LATEST MAIN ASSUMPTIONS ON ITALY MADE BY ISTAT (MEDIAN SCENARIO), EUROSTAT (BASELINE) AND UNPD (MEDIUM). Years 2023, 2030, 2050 and 2080.

Scenario	Total fertility rate	Life expectancy at birth – Men	Life expectancy at birth – Women	Net migration (thousand)
		Anno 2023		
Istat Median	1,20	81,3	85,3	274
Eurostat Baseline	1,25	81,4	85,8	241
UNPD Medium	1,30	82,1	86,1	58
		Anno 2030		
Istat Median	1,28	82,0	85,9	198
Eurostat Baseline	1,28	82,4	86,8	270
UNPD Medium	1,35	83,2	87,1	58
		Anno 2050		
Istat Median	1,38	84,3	87,8	166
Eurostat Baseline	1,37	85,0	89,0	240
UNPD Medium	1,44	85,8	89,6	58
		Anno 2080		
Istat Median	1,46	86,1	89,7	163
Eurostat Baseline	1,48	88,1	91,9	228
UNPD Medium	1,50	89,4	93,1	58

FIGURE A2. TOTAL POPULATION ACCORDING TO ISTAT, EUROSTAT AND UNPD SCENARIOS. Years 2023-2080, million.



2) Households projections, by region. Years 2023-2043

Household projections show the future trend of the number and type of households that will characterize the population in Italy from 2023 to 2043. These projections derive from the application of a static method, based on propensity rates, applied to the projected population. The purpose is to provide with an integrated system of information that can be useful to several users, both public and private, who deal with goods and services intended for families rather than for individuals. Given the importance of the role of the family, both at the protective level and for individual choices and paths, the demand for information on households arises from planning needs in various areas. First of all, we can consider the decisions to be taken in economic and social policies, such as those relating to housing, social and welfare systems for the young and the elderly. Last, improve the planning of productive strategies of durables goods for households and energy consumption is another potential task.

Territorial level and time horizon

Household projections are disseminated at regional and national level. The base population is the one observed on 1.1.2023 while the elaborations cover the period from 2023 to 2043.

Data

Several set of data have been implemented. Among them, the official probabilistic projections - base 1.1.2023 of the median scenario to be used as reference for the future evolution of the resident population by sex, age and region; the Base Population Register at 1.1.2022 and 1.1.2023 to estimate the share of population living in institutional cohabitation by sex and single year of age; the Multipurpose Survey "Aspects of daily life", which provide with a long time series (from 2002 to 2023), to derive the family structures by typology and position. The concept of household here in use is the "de facto family" that until 2022 identified a household as "*the set of people linked by ties of marriage, kinship, affinity, adoption, protection, or from emotional ties, cohabitants and having habitual residence in the same Municipality*". In 2023, the definition of household changes slightly with the inclusion of the concept of sharing economic resources: "*it is a group of people who habitually live in the same dwelling and are bound by marriage, civil partnership, kinship, affinity, adoption, guardianship or emotional ties and who share the household's income (by contributing to and/or benefiting from it) and/or daily expenses*"[4].

In addition, since the 2023 edition, the AVQ survey has introduced a new procedure for weighting data to accurately represent the target population. In fact, in order to make the survey results consistent with the evidence annually derived from the Permanent Census of Population and Housing, the calibration constraints for the construction of weights are defined within the system of "Anticipatory Estimates of Demographic and Social Indicators." Specifically, through this system, which provides the basis for weighting all Istat sample surveys, an estimate of the distribution of households by number of members is constructed in an anticipatory way, since the Permanent Census releases it in final form only the year following the year to which the survey data refer[5].

Method

The model is based on an adaptation to the Italian context of the method known as "Propensity model". It is a static method that goes beyond the classic "Headship rate model", overcoming the concept of 'head of household' and providing a much more detailed set of information. Predictions of the number of future households, their average size and composition, and the population by family role can be easily obtained. The method relies on Propensity rates, defined as the proportion of people of age x in household position i at time t:

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

For example, the propensity for a 30-year-old person to live in a couple with a partner will be given by the ratio between the number of 30-year-old people living in a couple and the total population of 30-year-olds.

The advantages of the method are many: it ties easily to population projections; there is no need to analyse transitions between potential family positions, typical of a dynamic model; it is simple to apply and provides with high detailed results. However, some drawback is also present, which arise mainly from the static nature of the method, do not allowing to reproduce the process of household formation and dissolution. Thus the application of propensity rates to the resident population may in some cases determine inconsistencies in term of global results, for example between sexes or for household positions within age-classes, a problem that it is therefore necessary to solve with ex-post adjustments.

[4] The change of AVQ's definition of household in 2023 is a consequence of the amendment of the Eurostat Regulation on Social Statistics, in which this definition is adopted.

[5] This major process innovation is based on a methodology that takes advantage of the current historical series about the distribution of households by number of members from census sources and, through an appropriate extrapolative nowcasting model, is able to estimate number of households and its breakdown for the year for which the information is needed.

The method consists of 5 steps:

- Step 1. Estimate the base-year and projected population living in households
- Step 2. Calculate household propensity rates
- Step 3. Modelling future trends of household propensity rates
- Step 4. Derive the projected population in the different household positions
- Step 5. Calculate the number, type and size of projected households.

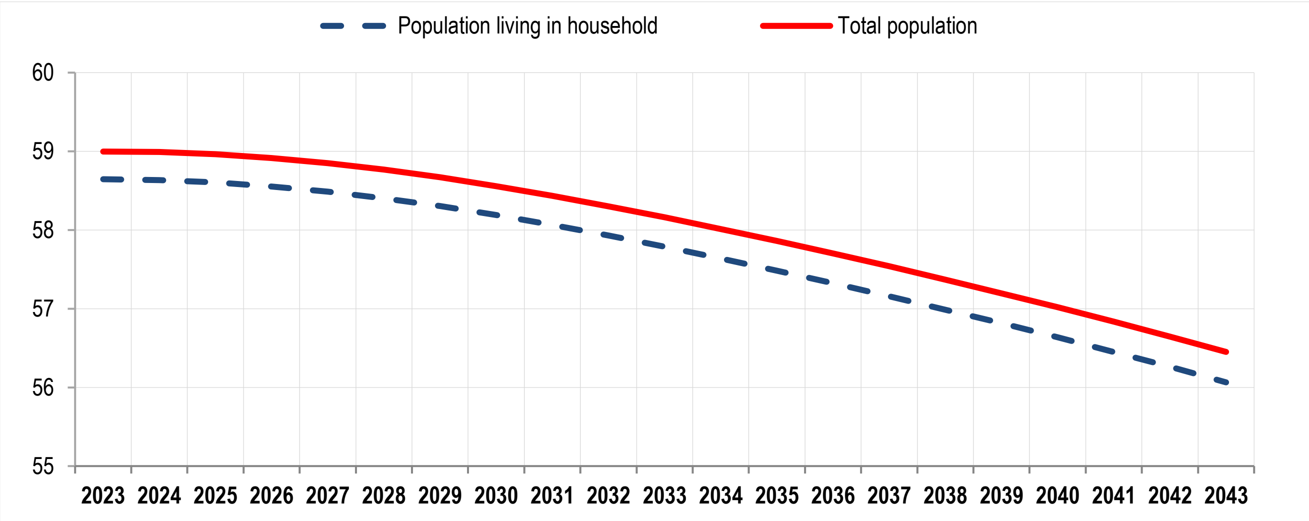
The various steps are explained in detail below.

Step 1. Estimate the base-year and projected population living in households

The base-year population is represented by the resident population by gender, age and region at January 1st, 2023, as collected from the last register-based Census. Then, making a preliminary estimate of the population living in households, excluding individuals residing in institutional households (hospitals, barracks, prisons, nursing homes, religious buildings, etc.) is necessary. This operation should then be repeated for every projected year, deducting from the regional projections (whose reference is also in this case the overall resident population) the share of the institutionalized population.

From the Base Population Register as of 1.1.2022 and 1.1.2023, the (average) percentage incidences of the population living in institutions by sex, five-year age group, and region were calculated. Given the substantial stability of the share of this population over time, these percentage values are assumed to be constant throughout the projection time horizon. Applying the 100 complement of these incidences to the total population from 2023 to 2043, we obtained the population living in households by region, sex, and age group (Figure A1).

FIGURE A3. TOTAL POPULATION AND POPULATION LIVING IN HOUSEHOLD. Years 2023-2043, median scenario, million.



Step 2. Calculate household propensity rates

The second step consists on calculating the propensity rates to live in a given household position by gender and 5-year age groups for the following 10 household positions:

1. lone person;
2. person in a childless couple;
3. person in a couple with at least one child under 20 years of age;
4. person in a couple in couple with all children aged 20 and older;
5. single parent with at least one child under 20 years of age;
6. single parent with all children aged 20 and older;
7. child (living with one parent in a couple or with a single parent);
8. other person living in a family household[6];

[6] Family refers to people who form a couple or a parent-child relationship. It means a married couple, civilly united or cohabiting, without children or with single children, or even a single parent together with one or more children who have never been married. Within a household there may be one or more families (family households), but there may also be none, as in the case of households formed by a single member (one person household) or several isolated members (multi-person household).

9. person in multi-person household (e.g., 2 siblings living together or a divorced individual who has returned home to a parent);
10. person in a household with 2 or more families.

Positions from 2 to 8 refer to individuals in one-family households. People living in households with 2 or more family nucleus have been considered in a separate category, since this typology constitutes a small share of the total number of households (approximately 1.5%).

As mentioned above, propensity rates are constructed as the proportion of persons of age x in category i. In this context, the age variable was considered in five-year classes and the rates were also disaggregated by sex, as the latest variable is very discriminating in household behaviour. Hereinafter, these rates are referred to as *Living Arrangement Propensities* (LAP):

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

where x= five-year age group 0-4, 5-9, ..., 80-84, 85+, i= family position, s=sex, t=time.

LAPs are calculated using data from the *Aspects of Daily Life* (AVQ) survey, along the entire 2002-2023 time series. Since regional estimates by sex and age groups leads to a paucity of data in small regions, it was decided to group regions into "macro-regions”.

A multivariate statistical analysis, including various sociodemographic context factors[7], has generated the following 5 groups of regions:

- Group 1 - North-west (Piemonte, Valle d'Aosta, Lombardia, Liguria);
- Group 2 - Eastern Adriatic (Veneto, Emilia-Romagna, Trentino-Alto Adige, Friuli-Venezia Giulia, Marche);
- Group 3 - Tyrrhenian (Toscana, Lazio);
- Group 4 - South (Campania, Puglia, Calabria, Sicilia);
- Group 5 - Central (Umbria, Sardegna, Abruzzo, Molise, Basilicata).

Step 3. Assumptions on future trends of household propensity rates

It is now necessary to make assumptions about the evolution of household propensities from 2023 to 2043. To this end, some modifications to the Propensity rates method have been introduced. The new approach is based on the introduction of a new synthetic indicator, constructed as the sum by age of the LAP, weighted by the years lived at the various ages (L_x). This new indicator is named *Total Propensity Rate per household position* (TPT):

$$TPR_{i,s,t} = \sum_{x=0-4}^{85+} LAP_{x,i,s,t} * L_{x,s,t} = \sum_{x=0-4}^{85+} \frac{P_{x,i,s,t}}{P_{x,s,t}} * 100 * L_{x,s,t}$$

where i=household position, s=sex, x=five-year age class, t=time.

$L_{x,s,t}$, representing the number of years lived in the age class x by sex s in year t, are derived from the projected life tables of the median scenario.

The TPR for a given household position represents how many years on average a generation of individuals expects to live in that position, assuming over the life course the family behaviours and mortality conditions as observed in a given calendar year. It is, therefore, a life expectancy in that family status, shifted from the cross-sectional to the longitudinal observational dimension. In other words, it takes on the same meaning that better-known cross-sectional indicators have, such as the average number of children per woman, the life expectancy at birth or the total marriage rate.

If in 2003 a man counted on living as a single person an average of 5.5 years (out of a total life expectancy of 77.2), in 2023 the expected time in this state rises to 11.2 years (out of a total of 81.1). In contrast, as a result of declining birth rates, in 2003 women expected to live 14.2 years (out of a total of 82.8) as a person in a couple with at least a child under 20, but in 2023 this expected time has fallen to 13 years (out of a total life expectancy that has since risen to 85.2 years). As a final example, the time in "child" status has increased from 30.5 to 31 years for males and from 27.7 to 28.6 for females, due to the prolonged stay of young people within the family of origin (Table A4).

[7] In order to identify homogeneous groups of regions, united by common family structures and similar evolution over time, a dynamic principal component analysis was carried out using the STATIS methodology. The analysis examined the main socio-demographic variables at the regional level in the years 2002-2019, including: fertility rates, mean age at birth, average size of families, separation and divorce rates, female employment rates, internal and foreign migration rates, quotes of some family types (single people, couples with and without children, single parents, etc.). The procedure was optimized by eliminating the variables with low latent variability explained by the axis.

TABLE A4. TOTAL PROPENSITY RATES BY HOUSEHOLD POSITION AND SEX. Years 2003, 2013 and 2023.

HOUSEHOLD POSITION	MALES			FEMALES		
	2003	2013	2023	2003	2013	2023
Lone person	5,5	7,9	11,2	10,8	12,3	13,4
Partner in a childless couple	13,4	14,2	14,2	12,4	13,1	13,3
Partner with at least a child <20	14,3	13,0	12,1	14,2	13,3	13,0
Partner with all children >=20 years	8,5	7,9	6,5	8,2	7,3	6,1
Lone parent with at least a child <20	0,2	0,2	0,4	1,4	1,7	2,6
Lone parent with all children >=20 years	0,6	0,8	0,8	2,7	3,0	3,1
Child	30,5	30,9	31,0	27,7	27,9	28,6
Person in a multi-person household	1,0	1,0	0,9	1,8	1,5	1,2
Person in a household with 2 or more families	1,00	1,2	1,8	1,3	1,6	1,8
Other position	2,1	2,8	2,1	2,3	3,1	2,1
Total	**77,2**	**80,0**	**81,1**	**82,8**	**84,8**	**85,2**

In order to hypothesize future trends in propensities, we proceeded to project the Total Propensity Rates by single family position, and then to estimate its distribution broken down by age group ($\mathrm{LAP}_{x,i,s,t}$) in each projected year. Predicting total intensity in a first step made it possible, on the one hand, to more easily translate the assumptions about family behaviour and, on the other, to keep together the trends in the various household positions. These latter, if projected separately by single age group, would be more difficult to control with the risk of obtaining unreliable results (e.g., a higher rate for the “child family position” at intermediate ages than at younger ones).

Before making the TPR extrapolations, however, it was necessary to address the changes introduced in 2023 in the weighting system of the AVQ survey. As mentioned in the section on data, the AVQ survey in 2023 uses for the first time weighting coefficients that are constrained to the total of households by region provided as part of the “Anticipatory Estimates of Demographic and Social Indicators” system and constructed on the basis of RBI-Census data. This innovation allows sample results to be more consistent with registry-census data. However, it entails differences from the results of the projections previous editions, mainly generating further increase of lone persons and decrease of couples with children.

Therefore, to prevent the values for 2023 from excessively affecting the extrapolations of TPRs, a correction of the last 5 years of the observed series of TPRs was made so that the effect of the new weights would be more gradual.

In summary, the procedure was based on the calculation of a correction factor as the ratio of the 2023 TPR obtained with the new weight bound to the anticipatory estimates of total households and the 2023 TPR calculated with the traditional weight. A linear growth of the correction factor was assumed from 2019 to 2023. For each year in this time frame, the traditionally calculated TPR was multiplied by the respective correction factor. Finally, the TPRs were adjusted, so that for each year their sum by household positions was equal to the life expectancy in that year.

The final goal of Step 3, which is to define the projected LAPs from 2023 to 2043 by region, was achieved by first performing the projection in the 5 established territorial groups (Step 3.1), and then moving from these to a regional detail (Step 3.2).

Step 3.1 Projecting LAPs in the 5 territorial groups

The total intensity of each family position and sex ($\mathrm{TPR}_{i,s,t}$) was predicted by trend extrapolation over the period 2002-2023, using time series analysis models. ARIMA, Random walk with drift or Linear Trend type models were applied for each family position and sex (Table A5).

TABLE A5. PREDICTIVE MODELS OF TOTAL PROPENSITY RATES BY HOUSEHOLD POSITION AND SEX
(Prevailing model among the 5 territorial groups) *

HOUSEHOLD POSITION	Males	Females
Lone person	RWD || ARIMA(1,0,0)	RWD
Person in a multi-person household	RWD	RWD
Partner in a childless couple	ARIMA(1,0,0)	ARIMA(0,2,1)
Partner with at least a child <20	RWD	ARIMA(1,0,0)
Partner with all children >=20 anni	RWD || ARIMA(1,1,0)	RWD || ARIMA(1,1,0)
Lone parent with at least a child <20	ARIMA(0,2,2)	Linear_Trend
Lone parent with all children >=20 anni	RWD	ARIMA(1,2,1)
Child	ARIMA(1,1,2)	ARIMA(0,2,1)
Other person	RWD	ARIMA(1,0,0)
Person in a household with 2 or more families	ARIMA(1,0,0)	Linear_Trend

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

Regarding the first activity, an example of the application of the above models is shown in Figure A4, where the behavior of the Northwest spatial group for the main family positions is depicted. More generally, with regard to each area of the country, variations in time spent in various household position are assumed, resulting in:

- an increase of “single people”;
- a decrease of “partners with children”;
- a slight increase of “partners without children”;
- an slight increase of people in “child” position;
- a slight increase of “lone parents”, especially fathers;
- a slight decrease in “other people” living in households with one family nucleus and a slight increase for people in households consisting of two or more family nuclei
-

Estimation of the predicted distribution by age, i.e. $\mathrm{LAP}_{x,i,s,t}$ from 2023 to 2043, was obtained using predicted TPR, predicted years lived L_x, and observed distributions from AVQ survey data over the three-year period 2021-23. For this latter purpose, the mean 2021-23 distributions by single position in the household were weighed by two coefficients: one to account for the predicted TPR in year t relative to that in the 2021-23:

$$\mathrm{WP}_{s,i,t} = \frac{\mathrm{TPR}_{s,i,t}}{\mathrm{TPR}_{s,i,2021-23}} \qquad t = 2023, \dots, 2043$$

and a coefficient expressing changes in mortality over time:

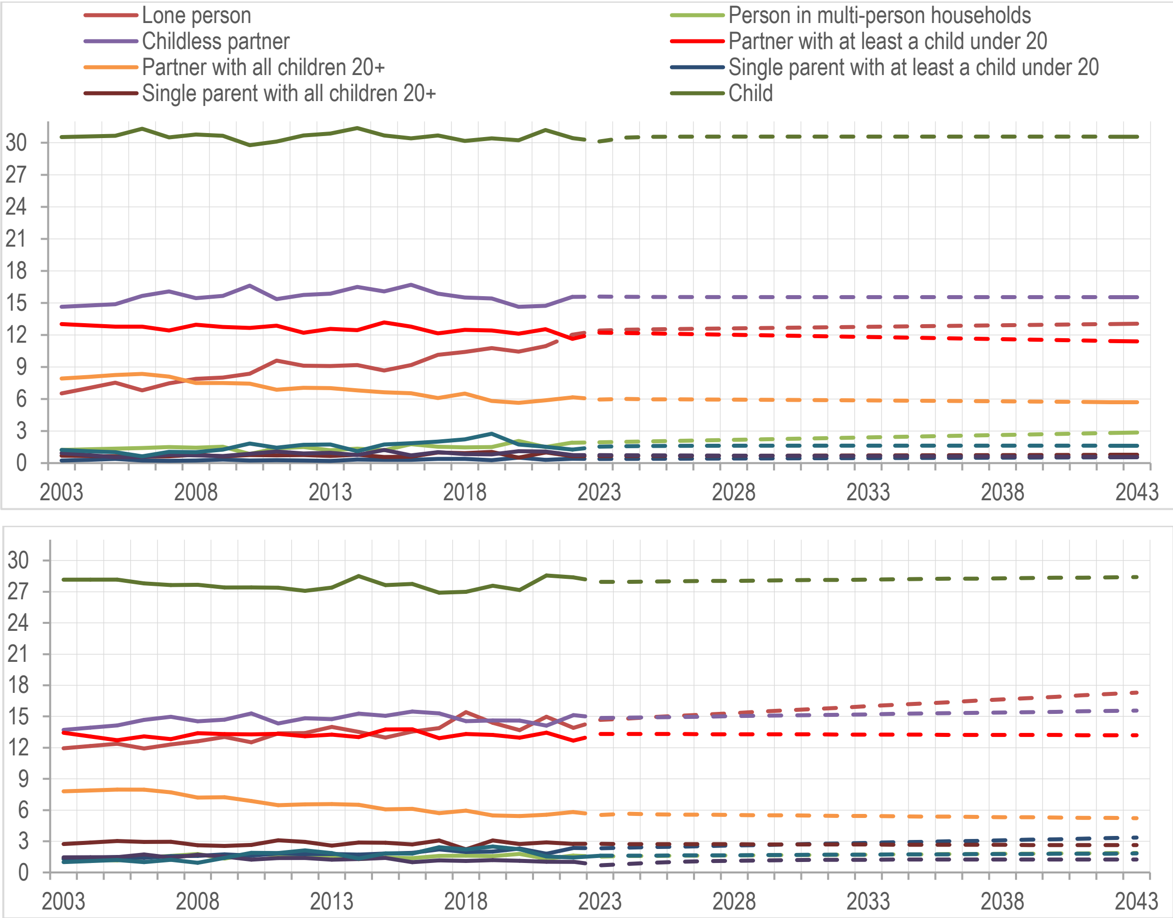
$$\mathrm{WL}_{x,s,t} = \frac{\mathrm{L}_{x,s,2021-23}}{\mathrm{L}_{x,s,t}} \qquad t = 2023, \dots, 2043$$

Therefore, household propensities throughout the projection horizon were calculated using the formula:

$$\mathrm{LAP}_{x,s,i,t} = \mathrm{LAP}_{x,s,i,2021-23} * \mathrm{WP}_{s,i,t} * \mathrm{WL}_{x,s,t} \qquad t = 2023, \dots, 2043$$

where: x=age groups 0-4, ... ,85+, s=sex, i=household position.

FIGURE A4. TOTAL PROPENSITY RATES BY HOUSEHOLD POSITION AND SEX. North-west. Years 2003-2043.



Since the method of estimation did not assume annual variation about the age distribution of LAPs, it was implicitly assumed that behaviours in terms of family choices would maintain in the future an age-group distribution proportional to that found in the AVQ survey in the three-year period 2021-2023. Last, the sum of LAPs by household position in each age group approximates but does not always equal the value of 100, so ex-post adjustments were necessary.

Step 3.2. Projecting LAPs in the regions

In order to project households at the regional level, it is necessary to ensure that each region has its own socio-demographic specificity within the projecting group to which it belongs. Considering that, a regional correction factor has been defined to be applied to the LAP projections of the various territorial groups in order to obtain those specific to each region:

$$\mathrm{FC}_{r,i} = \frac{\mathrm{TPR}_{2021-23,i,r}}{\mathrm{TPR}_{2021-23,i,G}}$$

where i=household position, r=region, G=group to which region r belongs.

The projected LAPs for the spatial groups are then multiplied by the regional correction factor calculated in this way, determining the series of regional LAPs from 2023 to 2043. For example, for the single male person, the detected TPR in Piemonte is 11.9 while in group 1 it is 11.1. The correction factor is therefore equivalent to 1.07 in this case. This means that since Piemonte has a higher TPR than the group to which it belongs, an adjustment must be made by multiplying all LAPs at different ages and projected years by 1.07, increasing the level slightly.

Step 4. Obtain the projected population by single household position.

In this step, the regional propensities are applied to the projected population living in households, as it was obtained in Step 1. The projected population in the different household states by sex, age group, and region from 2023 to 2043 is then derived.

Step 5. Calculate the number, type, and size of projected households

The projected number of households is obtained directly from the population separated by family position, gender, and age, as:

- each "single person" represents 1 household (coefficient=1);
- persons in a couple constitute 0.5 of a family (coefficient=0.5);
- each "single parent" represents 1 family (coefficient=1);
- “multi-person households" are obtained by dividing the number of persons living in multi-person households by the average size of this type of household, substantially stable over time and equal to about 2.1 members (coefficient=2.1);
- households "with 2 or more families” are obtained by dividing the number of persons living in households with 2 or more families by the average size of this type of household, which assumes time series coefficient values between 5.1 and 5.4 depending on the territorial group of reference.

Applying these coefficients to the population of sex s and age x yields as a final product the number of households by the family types of interest.

The average number of members is then calculated by dividing the population living in the household to the number of households. It can be disaggregated for total households and those with at least one family (excluding single persons and multi-person households).

For dissemination purposes, multi-person households and those with 2 or more families are considered together in the "other type of household" item.

3) Data dissemination and terms of use

The detailed picture of the assumptions underlying the projections and the main results can be consulted on dati.istat.it (topic: Population and families> Population projections) and demo.istat.it.

The dissemination of the **population projections** is implemented into the three following sections: population structure by sex and single age group; components of the population change; main demographic indicators. Each table shows the values of the median scenario and the lower and upper limits of the 90%, 80% and 50% confidence intervals.

The components of the population balance include:

- population at start and end of the year, total variation
- births, deaths, natural change
- in-migration and out-migration, net migratory balance
- interregional changes of residence, net internal migration balance.

The data described above and those relating to the age distribution of the population are rounded to the nearest unit.

Regarding the demographic indicators, the tables include:

- birth, mortality and natural growth rates
- immigration rate from abroad, emigration for abroad and net migration rate with abroad
- internal immigration rate, internal emigration rate and net internal migration
- total net migration rate and total growth rate
- mean age of the population
- % of population 0-14 years, 15-64 years, 65 years and over, 85 years and over
- structural dependency ratio, elderly dependency ratio and aging index
- total fertility rate
- life expectancy at birth and at 65 years of age by sex.

The dissemination of data relating to **household projections** is articulated into three sections including tables that can be processed on the structure by sex, five-year age group and household position of the population, on the distribution of household by type and on the average number of household members. All results refer to the median scenario.

The reproduction of the information contained in this note and in the databases dati.istat.it and demo.istat.it is left free, provided that the Istat source is quotes.

Istat periodically produces population projection under the activity line “System of Population Estimates and Projections,” in accordance with the National Statistical Program, project “Population Projections” (PSN code IST-01448).

References

1. 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.
2. 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.
3. 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.
4. 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.
5. Istat (1989), *Previsioni della popolazione residente per sesso, età e regione – Base 1.1.1988*, Note e Relazioni, n.4.
6. Istat (1989), *Previsioni del numero di famiglie italiane dal 1995 al 2020*. Notiziario. Serie 4. Foglio 41 – Anno X No. 19. Dicembre 1989.
7. Istat (1997), *Previsioni della popolazione residente per sesso, età e regione – Base 1.1.1996*, Informazioni, n. 34.
8. Istat (2001), *Previsioni della popolazione residente base 1° gennaio 2000*, Statistiche in breve, www.istat.it.
9. Istat (2003), *Previsioni della popolazione residente per sesso, età e regione, base 1.1.2001*, Informazioni n.13.
10. Istat (2006), *Previsioni demografiche nazionali 1° gennaio 2005-1° gennaio 2050*, www.istat.it, Nota informativa, 22 marzo 2006.
11. Istat (2008), *Previsioni demografiche 1° gennaio 2007-1° gennaio 2051*, www.istat.it, Nota Informativa, 19 giugno 2008.
12. Istat (2011), *Il futuro demografico del paese - Previsioni regionali della popolazione residente al 2065*, Statistiche Report, www.istat.it, 28 dicembre 2011.
13. Istat (2017), *Il futuro demografico del Paese – Previsioni regionali della popolazione residente al 2065*, Statistiche Report, www.istat.it, aprile 2017.
14. 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.
15. 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.
16. Istat (2021), *Previsioni della popolazione residente e delle famiglie – base 1.1.2020*, Statistiche Report, www.istat.it, novembre 2021.
17. Istat (2022), *Previsioni della popolazione residente e delle famiglie – base 1.1.2021*, Statistiche Report, www.istat.it, settembre 2022.
18. Istat (2023), *Previsioni della popolazione residente e delle famiglie – base 1.1.2022*, Statistiche Report, www.istat.it, settembre 2023.
19. Istat (2024), *Indicatori demografici – Anno 2023*, Statistiche Report, www.istat.it, marzo 2024.
20. 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.
21. 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.
22. Terra Abrami V. (1998), *Le previsioni demografiche*, Il Mulino, Bologna.
23. UNECE (2018), *Recommendations on Communicating Population Projections*, United nations economic commission for Europe, United Nations, New York and Geneva, agosto 2018.