SEPTEMBER 22th 2022

statistiche report

Istat Istituto Nazionale di Statistica

HOUSEHOLDS AND POPULATION PROJECTIONS | BASE 1/1/2021

Fewer residents, more elderly, smaller families

Latest **population projections**, updated to 2021, confirm the presence of a potential crisis scenario. A **decreasing population**: from 59.2 million as of January 1, 2021 to 57.9 million in 2030, to 54.2 million in 2050 and to 47.7 million in 2070.

The ratio between **individuals of working age** (15-64 years) and **not** (0-14 and 65 years and over) will go from about **3 to 2** in 2021 to about **1 to 1** in **2050**.

The demographic crisis of the territories: a population decline is expected in 4 out of 5 municipalities within 10 years, in 9 out of 10 in rural municipalities.

The **number of families** is expected to **grow** but with an ever smaller mean number of members. **Fewer couples with children**, **more couples without**: by 2041 only 1 in 4 families made up of a couple with children; more than 1 in 5 will be childless.

34.9%

Individuals aged 65 and over in 2050

From 23.5% in 2021.

2049

The year in which deaths could double live births (788 thousand against 390 thousand)

10.2 million

People expected to live alone in 2041

From 8,5 million in 2021.

www.istat.it

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

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

A continuous decline of the population

The resident population is expected to decrease over the next decade according to the "median" scenario (Table 1): from 59.2 million as of January 1st, 2021 (starting point of the projections) to 57.9 million in 2030 for an annual rate of change of -2.5 per thousand. In a medium-term perspective, however, the decrease in the population would be more pronounced: from 57.9 million to 54.2 million between 2030 and 2050, with an average annual rate of change of -3.3 per thousand.

In the long term, the consequences of the expected demographic dynamics on the population become more important. Between 2050 and 2070 the population would decrease by a further 6.4 million (-6.3 per thousand on annual average). Under this assumption, the total population would amount to 47.7 million in 2070, resulting in an overall loss of 11.5 million residents compared to today.

Demographic projections are uncertain by definition. Such a characteristic is the more relevant the more time is further away from the base year. The future evolution of the total population fully reflects this principle after only a few years of projection. In 2050, the 90% confidence interval of the total population ranges between 51.1 and 57.5 million. Twenty years later it ranges from 41.2 to 55.1 million residents.

Thus, according to the more favourable assumption the population could record a loss of "only" 4.2 million between 2021 and 2070, on the other hand, a decline of 18 million could be achieved. It is almost certain that the population will decrease. In fact, although the possibility that the demographic dynamics could lead to a population in 2070 larger than today is not entirely excluded, the empirical probability of this event is minimal, being equal to 1% (percentage of cases in favour of the event on the total of the simulations conducted).

The issue affects the whole Country, albeit with differences between the Centre-North and the South. According to the median scenario, in the short term it is expected that in the North (-0.9 per thousand per year until 2030) and in the Centre (-1.6) there may be a smaller reduction in the population than in the South (-5.3). In the medium term (2030-2050) and long term (2050-2070) this trend strengthens, with a population decline shared by all geographical areas, but more strongly in the southern one. For example, in the less disadvantaged North there is an average annual reduction of 1.4 per thousand in 2030-2050 and one of 4.2 per thousand in 2050-2070. On the contrary, in the South the reduction is 6.8 and 10.1 per thousand, respectively, in the two periods.

The evolution of the population by geographical area is also marked by uncertainty. For the North, this uncertainty makes difficult to understand the direction of demographic change. Oriented towards growth, as shown by the upper limit of the confidence interval in 2070 (28.3 million), or to decrease by looking at the lower one (20.8). Its final population, therefore, is between two values, respectively, below and above that in the base year, although the median scenario (24.4 million) shows that its decrease is more likely. The Centre and the South show a range of assumptions for the year 2070 whose maximum values (11.3 and 15.4 million, respectively) are lower than their base populations.

TABLE 1. RESIDENT POPULATION BY GEOGRAPHICAL AREA. MEDIAN SCENARIO AND 90% CONFIDENCE INTERVALS. Years 2021-2070, January 1st, million (*)

Geographic area	2021	2030	2040	2050	2070
North	27.5	27.3	27.0	26.5	24.4
		[27.0 / 27.5]	[26.2 / 27.9]	[24.9 / 28.3]	[20.8 / 28.3]
Centre	11.8	11.6	11.4	11.0	9.8
		[11.5 / 11.7]	[11.1 / 11.7]	[10.4 / 11.7]	[8.4 / 11.3]
South	20.0	19.0	18.0	16.6	13.6
		[18.9 / 19.2]	[17.5 / 18.4]	[15.8 / 17.5]	[11.9 / 15.4]
ITALY	**59.2**	**57.9**	**56.4**	**54.2**	**47.7**
		[57.4 / 58.4]	[54.8 / 58.0]	[51.1 / 57.5]	[41.2 / 55.1]

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

Demographic dynamics addressed to a growing imbalance

For fifteen consecutive years Italy has been facing the reality of a negative natural change, a factor that underlies the process of population reduction, despite the partial counterpart of positive migratory dynamics with foreign countries.

Assumptions on births and deaths amplify this process, measuring an evident tendency to annually record negative balances for the natural movement of the population. Not even in the most favourable opposing scenarios would the projected number of births come to compensate for that of deaths. For example, the upper bound of the 90% confidence interval for births, corresponding to scenarios in which the total fertility rate grows to 1.88 in 2070, would still determine a lower number of births than the number of deaths under its lower confidence limit (Figure 1).

In the median scenario, where a growth in fertility is contemplated from 1.25 children per woman in the base year to 1.55 in 2070, the maximum number of births achieved would be equal to 424 thousand units in 2038. After this year, the further increase in average reproductive levels does not lead to a parallel increase in births, as women of childbearing age will tend to decrease as well as age on average, reducing the country's reproductive potential.

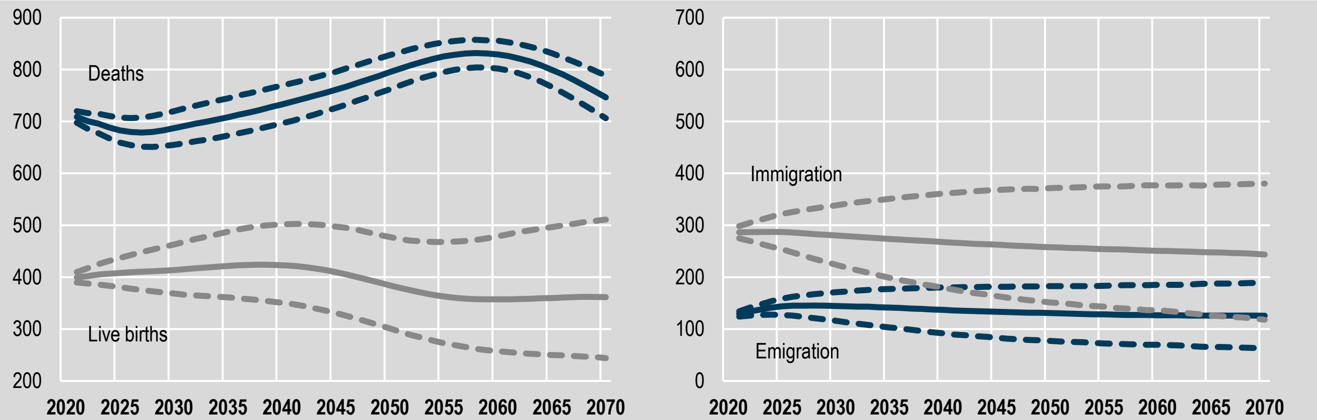
The evolution of mortality may in turn depend on similar perturbations of structural origin, which will continue to express a sustained and growing number of death events year after year, up to a maximum of 832 thousand in 2058 according to the median scenario. This even in a context of good expectations on the evolution of life expectancy (86.5 and 89.5 years that expected at birth in 2070, respectively for men and women) and, therefore, in line with what will be the level of aging of the population.

Migratory flows will not counterbalance the negative natural dynamics. Nonetheless, migration is characterized by deep uncertainty, as there are various factors that can give rise to diversified scenarios. To stick to what has happened in recent years, just think of the drastic reduction in migration dictated by the pandemic in 2020, the subsequent economic recovery started in 2021 which acted as a lever for attracting immigrants, the climate of confidence generated by the prospects on the PNRR and, finally, to the current uncertainties determined by the war and energy crisis on an international level.

The median scenario contemplates largely positive net migratory movements with foreign countries, albeit with a slightly decreasing trend, from over 150 thousand units per year to about 120 thousand between 2021 and 2070. Cumulating over the entire forecasting period, the median scenario therefore prefigures a permanent settlement of 13.2 million immigrants, on the other hand the amount of emigrants abroad would amount to 6.7 million.

The analysis of such long-term results must necessarily be accompanied by great caution, to the point that the 90% confidence interval of the net migration balance with foreign countries returns in 2070 values ranging from -20 thousand to + 268 thousand, revealing, in fact, two separate pictures. On the one hand that of an attractive country, on the other that of a country that could change its current nature of hospitality to return to being a place from which to emigrate.

FIGURE 1. NATURAL AND MIGRATORY CHANGE OF THE POPULATION, MEDIAN SCENARIO AND 90% CONFIDENCE INTERVALS. Italy, Years 2021-2070, thousand.



Growth of the elderly population and accentuation of structural imbalances

Today the population aged 65 and over represents 23.5% of the total, those up to 14 years of age 12.9%, those in the 15-64 age group 63.6%, while the mean age of the population has approached 46 years. The country's population is already well into an accentuated and prolonged phase of aging. From future prospects, an amplification of this process arises, mostly governed by the current breakdown by age of the population, and only to a lesser extent by the imagined changes regarding the evolution of fertility, mortality and migration. Based on a ratio of importance, approximately, of two thirds and one third respectively.

By 2050, people aged 65 and over could represent 34.9 of the total according to the median scenario, while the 90% confidence interval ranges from 33% to 36.7%. Whatever happens, therefore, further steps will be needed to adapt social protection policies to such a growing share of the elderly population.

Young people up to 14 years of age, although in the median scenario a recovering fertility is expected, could represent 11.7% of the total by 2050, thus recording a slight decline. On the level of intergenerational relationships, however, the issue of an unbalanced relationship between the over 65s and children would arise, to an extent of about 3 to 1.

In the meantime, the transit of the large generations born in the baby boom years (born in the 60s and first half of the 70s) between the adult and senile ages will contribute to the absolute and relative growth of the elderly population, with concomitant and sudden reduction of the working age population. In the next thirty years, in fact, the population aged 15-64 would drop from 63.6% (37.7 million) to 53.4% (28.9 million) based on the median scenario, with a potential range between 52% and 54.8%. As for the elderly population, therefore, here too a certain evolutionary picture is expected, with potential effects on the labor market, on economic planning, on the maintenance of the level of welfare necessary for the country.

A partial rebalancing in the population structure could occur in the long term as the baby boom generations will tend to become extinct. According to the median scenario, the 15-64 year olds could therefore return to 54.3% by 2070 while the over 65 year olds could decrease to 34.1%. On the other hand, the youth population remained stable with a level of 11.6%.

Among the demographic transformations that will affect the country, the marked aging process that will affect the South is to be highlighted (Table 2). Although this geographical area still presents a younger age profile today, the mean age of its residents passes from 45 years in 2021 to 49.9 years in 2040 (median scenario), at that point surpassing the North which reaches a mean age of 49.2 years, starting in the base year from a higher level (46.4). Looking, instead, at the long-term prospects, the South would slow down but would not stop its path, reaching a mean age of the population close to 52 years. At that point, on the other hand, both the North (49.7 years) and the Center (51.1) would have already started the opposite path, towards an age structure in small part rejuvenated.

TABLE 2. MEAN AGE OF THE POPULATION BY GEOGRAPHIC AREA, MEDIAN SCENARIO AND 90% CONFIDENCE INTERVALS .Years 2021-2070, January 1st, in years and tenth of year (*).

Geographic area	2021	2030	2040	2050	2070
North	46.4	48.0	49.2	49.9	49.7
		[47.7 / 48.2]	[48.3 / 49.9]	[48.4 / 51.3]	[47.1 / 52.5]
Centre	46.6	48.5	50.2	51.2	51.1
		[48.3 / 48.7]	[49.4 / 50.9]	[49.7 / 52.6]	[48.5 / 53.8]
South	45.0	47.5	49.9	51.5	51.9
		[47.2 / 47.7]	[49.1 / 50.6]	[50.1 / 52.9]	[49.3 / 54.7]
ITALY	**45.9**	**47.9**	**49.6**	**50.6**	**50.6**
		[47.7 / 48.1]	[48.8 / 50.4]	[49.2 / 52.1]	[48.0 / 53.4]

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

4 out of 5 municipalities in demographic decline within ten years

An increasing number of municipalities will face a demographic decline within 10 years. According to the median scenario, 80% should be in such conditions by 2031. This is due to the low fertility, which uniformly affects the age structure of the populations at the base, but also to unfavorable migratory levels for some territorial realities, where both abroad and internal emigration are stronger.

It is foreseen that between 2021 and 2031 the municipalities of rural areas may overall record a population reduction of 5.5%, going from 10.1 to 9.5 million residents (Table 3). In these areas, the municipalities with a negative population balance are 86% of the total. The issue affects above all the areas of the South, where the municipalities of rural areas with a negative balance are 94% of the total and where there is a population reduction of 8.8%.

For the 1,060 municipalities that fall into *Internal Areas*, particular areas of the national territory that are characterized by a physical distance from the supply of essential services, the demographic condition is even more unfavorable. Here, in fact, the share of municipalities with a negative population balance in the decade rises to 94%, with a total population reduction of 9.1% (but 10.4% in the South).

The municipalities with intermediate density (small towns and suburbs) are in a relatively better situation. Here, the expected demographic decline is 1.9% (the population passes from 28.3 to 27.7 million in the decade) and the share of municipalities affected by the demographic decline is equal to 70% of the total, which however rises to 84% in the South. Finally, although to a lesser extent, densely populated cities will also be affected by depopulation. The attractiveness of areas with strong urbanization will mean that in the decade their overall population decline is only 1.8%, with 65% of the municipalities destined to register a negative balance among their residents.

In 2041 one million more families but smaller

Within twenty years the number of families is expected to increase by about one million units: from 25.3 million in 2021 to 26.3 million in 2041 (+3.8%). These are increasingly smaller families, characterized by greater fragmentation, whose average number of members is expected to drop from 2.3 people in 2021 to 2.1 in 2041. Even families with at least one nucleus (characterized by the presence of at least one couple or parent-child relationship) will change their average size from 3.0 to 2.8 members.
The increase in the total number of families largely depends on the evolution of families without a nucleus, which increasing by 20.5% in the period 2021-2041(from 9 to about 11 million) would represent 41.4% of total families. On the contrary, families with at least one nucleus would follow an opposite trend, presenting a decrease of 5.4% in the 20 years considered. These families, today equal to 16.3 million or 64.3% of the total, in 2041 would drop to 15.4 million, thus representing 58.6%.

TABLE 3. POPULATION BY URBANIZATION / INTERNAL AREA AND GEOGRAPHIC AREA – MEDIAN SCENARIO.
Years 2021 and 2031, January 1st, thousand *

Geographic area	Year	City or Densely populated area	Small town, suburb, intermediate density	Rural or scarcely populated area	Internal area	Non internal area
North	2021	9,331	13,333	4,823	494	26,992
	2031	9,339	13,236	4,657	459	26,773
Centre	2021	4,485	5,393	1,910	375	11,412
	2031	4,441	5,342	1,814	344	11,254
South	2021	7,067	9,545	3,351	1,069	18,893
	2031	6,732	9,150	3,057	959	17,981
ITALY	**2021**	**20,882**	**28,270**	**10,084**	**1,939**	**57,297**
	2031	20,512	27,728	9,528	1,762	56,007

(*) The degree of urbanization responds to the classification Degurba (Population Statistics by regular grid, www.istat.it/it/archivio/155162). Internal areas, defined by the Social Cohesion Agency, are areas of the Italian territory characterized by a significant distance from the supply of essential services. They include 1,060 Municipalities (www.agenziacoesione.gov.it).

In twenty years over 10 million lonely people

The decline in families with a nucleus derives from the long-term consequences of the socio-demographic dynamics: the aging of the population, with the increase in life expectancy, generates a greater number of lonely people; the prolonged decline in the birth rate increases the number of people without children, while the increase in marital instability determines an increasing number of lone individuals and single parents.

The absolute growth of the total number of families is mainly due to lonely people, however associated with the concept of family, even if micro. Men living alone would have an increase of 18.4%, reaching over 4 million in 2041. Lone women are expected to increase even more, from 4.9 to almost 6 million, with an increase of 22.4%. Then, single-component families, because of their composition by age, have an important social impact: it is, in fact, mainly in advanced age that the number of lonely people increases.

While in 2021 the share of single people aged 65 and over represents half of those living alone, in 2041 it would reach 60%. In absolute terms, lonely people would reach 10.2 million (+20%), of which 6.1 million will be over 65 years old and over (+44%). In 2021, among men living alone, about one in three is over 65 (32.3%), while among women the ratio is over three in five (63.1%). Over the years, projections show a scenario in which the incidence of men and women aged 65 and over in single-member families as a whole increases substantially, so that in 2041 men would make up 42.5% and women 72.2% of them.

The increase in survival among the elderly, many of whom living alone, could lead to a future increase in care needs. However, a greater number of lonely elderly people can also generate positive implications. A longer survival, combined with a better quality of life, could allow these people to play an active role in society: for example, as already happens today and probably in the future, by supporting the families of their children in care of grandchildren and guaranteeing them economic support, participating in the economic cycle as consumers of welfare services but also as capital investors.

Couples with children will decrease

On the basis of the fertility levels found in recent years and the assumptions made under the median scenario, a substantial decrease in couples with children is expected. This type of family, which today represents about a third of total families (32.5%), in 2041 could represent less than a quarter (24.1%). Between 2021 and 2041 the consistency of families with children would decrease by 23%, going from 8.2 to 6.3 million. In particular, taking into account the age of children, the most significant decrease would be recorded among couples with at least one child between the ages of 0 and 19 (-26%). This latter type of family would decrease from 5.3 million in 2021 to 3.9 million in 2041, the share dropping from 21% to 15% of total families (Table 4).

TABLE 4. NUMBER OF HOUSEHOLDS BY TYPE. Years 2021*, 2031, 2041, median scenario, thousand.

TYPE	2021	2031	2041	TYPE	2021	2031	2041
Total number of families	25,323	25,895	26,289	Lone mother	2,197	2,289	2,318
of which with nucleus	*16,285*	*15,996*	*15,400*	*of which at least one child <20 years*	*902*	*909*	*951*
Couple with children	8,232	7,253	6,332	Couple without children	5,003	5,463	5,657
of which at least one child <20 years	*5,301*	*4,413*	*3,931*	Living alone - Men	3,584	3,883	4,242
Lone father	532	667	770	Living alone - Women	4,874	5,380	5,967
of which at least one child <20 years	*162*	*186*	*206*	Other **	902	959	1,003

(*) The multipurpose survey on "Daily Life Aspects" are released on a two-year average. Here, however, the data refer to 1st January. For 2021 this can give rise to differences.

(**) Multi-person households (several people living together who do not form a nucleus) and households with more than one nucleus.

Couples without children and lone parents on the rise

Childless couples would increase from 5 to 5.7 million, for an increase of 13%, and with a share of the total that would rise from 19.8 to 21.5%. If these trends would continue with the same intensity even beyond 2041, childless couples could outnumber those with children by 2045.

Marital instability, increasingly widespread in the country, will contribute to the increase of families made up of a single parent, male or female, with one or more children. In 2021 there were a total of 2.7 million single parents, more mothers (2.2 million) than fathers (just over 500 thousand) which respectively represent 8.7% and 2.1% of the total families.

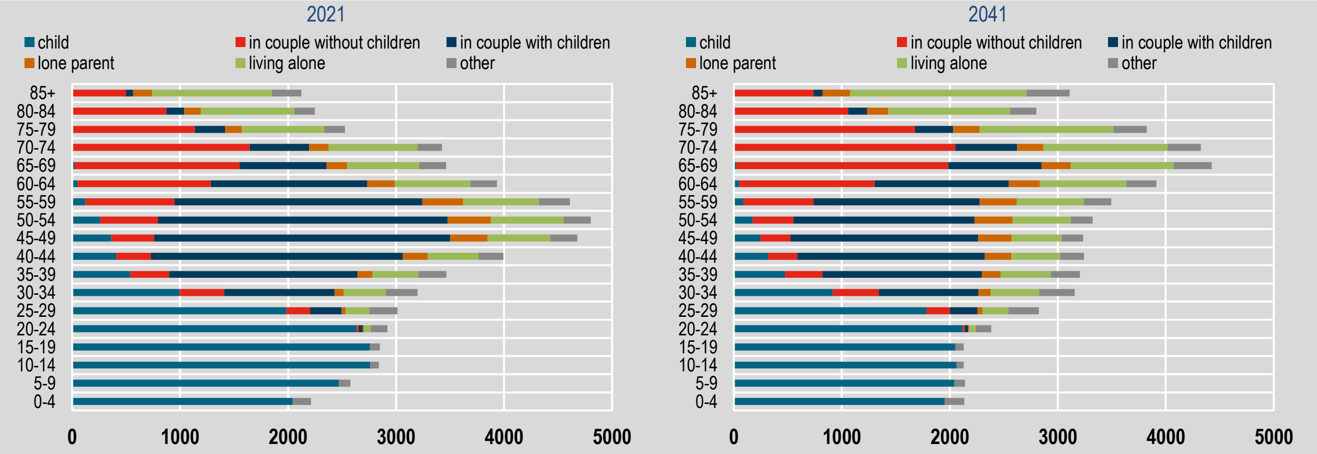
In the past, following a breakup of the couple, children were generally entrusted to mothers. Since the promulgation of the law on joint custody in 2006 this prevalence has been decreasing. This has led to an ever greater spread of fathers as foster parents in separation or divorce sentences. By 2041, single fathers, while remaining a minority compared to single mothers, could amount to about 800 thousand (2.9% of the total number of families). In that year, single mothers would arrive to 2.3 million (8.8% of the total), so that the total of single parents would be 3.1 million.

The increase in single parents, from 10.8% to 11.7% of total families, remains modest, as it is contrasted by the continuous decline in births, as well as by the tendency to re-aggregate to other families or to form new ones. Analysing the composition of single-parent families by age of children, it is noted that the increase, while maintaining a limited consistency, will involve mainly single fathers with children aged 20 and over, who in 2041 would exceed 500 thousand units. In parallel, single fathers with at least one child aged less than 20 years would be just over 200 thousand.

Family roles will be affected by demographic and social transformations

Past and future demographic dynamics determine a reduction of the new generations, in both absolute and relative terms. The age structure of the population already shows a high imbalance in favour of the older generations and at the moment there are no factors that could suggest a change of direction. Looking at demographic projections, a turnaround in the number of births in the years to come appears unlikely, even in the face of favourable fertility assumptions. This is due to a decreasing number of women of childbearing age, on the one hand, and the prolonged tendency to postpone parenthood on the other.
The comparison between the population in 2021 and that projected in 2041, broken down by family role, shows the demographic and social changes that are expected in the forthcoming twenty years. In particular, there is an increase in single parents, single people and couples without children, the latter especially if elderly (Figure 2).

FIGURE 2. POPULATION BY POSITION IN THE HOUSEHOLD AND FIVE-YEAR AGE GROUP. Years 2021 and 2041, median scenario, thousand.



Growing lifespan for older couples without children

The increase in couples without children will have greater intensity among people aged 65 and over, for whom the prolonged survival of the partner and the exit of the children from the family means that the period in which they remain as a couple is prolonged.

Among lonely people, the increase would be substantial for women aged 65 and over and men aged 75 onwards. On the contrary, the decrease that would occur in the middle ages for people in couples are important, especially for those couples between 35 and 59 years of age with at least one child under 20 years. Among those couples who have only children aged 20 and over the decrease will mostly occur between 50 and 64 years. Furthermore, due to the low birth rate, also the position of child in the young age records a decrease, although it remains prevalent up to 29 years of age due to the permanence in the family of origin.

Family transformations in every territorial area

The family types respond to demographic dynamics and social behaviour specific to the different areas of the country, with more marked differences between the North and the South.
In the North, in 2021, the share of households with at least one nucleus is lower, precisely 64% against 67% in the South (Table 5). The expected change for this type of families is substantial: in 2041 they could make up 58% of total families in the North and 61% in the South, registering a reduction of 6 percentage points in both cases. In the Center, households with a nucleus would have a similar reduction, equal to about 5 percentage points, making up 57% of total households, with an approach to the North.
The family type couple with children is the one that is expected to undergo the most evident change in the next twenty years. In the South, the expected decline is around 9 percentage points (from 37% in 2021 to 28% in 2041), while in the North (from 31% to 23%) and in the Center (from 30% to 22%) it is around 8%, so that the South would still maintain a higher proportion of couples with children.
Most of the reduction that will affect the "couple with children" typology actually concerns couples with at least one child under 20 years of age. In the North, this type of couple falls from 21% in 2021 to 16% in 2041, achieving a reduction of 5 percentage points compared to the 8 overall lost by couples with children, regardless of the age of the children. In the Center, it goes from 19% to 13%, with 6 points lost out of 8 overall. In the South, projections point to a wider demographic crisis. Here, couples with at least one child under the age of 20 would decrease by 7 percentage points out of the 9 overall of couples with children. The final result, therefore, is that for couples with "young" children there is a process of territorial convergence. The same, on the contrary, cannot be said for couples with "mature" children, where a difference remains in favour of the South, partly due to the fact that in this area of the country the times for leaving the family of origin are longer.

TABLE 5. HOUSEHOLDS BY TYPE AND GEOGRAPHIC AREA. Years 2021, 2031, 2041, median scenario, % values.

TYPE	North 2021	North 2031	North 2041	Centre 2021	Centre 2031	Centre 2041	South 2021	South 2031	South 2041
HOUSEHOLD WITH NUCLEUS	63.6	60.8	57.8	61.6	59.6	56.7	67.0	64.7	61.1
HOUSEHOLD WITHOUT NUCLEUS	36.4	39.2	42.2	38.4	40.4	43.3	33.0	35.3	38.9
Couple with children	30.7	26.1	22.5	29.5	25.2	21.5	37.2	32.8	28.3
of which at least one child <20 years	*20.6*	*17.0*	*15.5*	*19.3*	*15.5*	*13.4*	*22.5*	*18.1*	*15.1*
Lone father	2.1	2.6	3.0	2.4	3.0	3.3	1.9	2.3	2.6
of which at least one child <20 years	*0.6*	*0.6*	*0.7*	*0.8*	*0.9*	*1.0*	*0.7*	*0.8*	*0.9*
Lone mother	7.9	7.9	7.8	9.6	10.2	10.4	9.2	9.4	9.4
of which at least one child <20 years	*3.3*	*3.2*	*3.4*	*3.9*	*3.8*	*3.8*	*3.7*	*3.7*	*3.7*
Couple without children	21.9	23.1	23.4	18.4	19.5	19.8	17.4	19.0	19.7
Living alone - Men	14.5	15.9	17.3	15.6	16.1	17.0	12.7	13.0	13.7
Living alone - Women	19.8	21.1	22.6	20.3	21.5	23.0	17.8	19.9	22.7
Other	3.2	3.3	3.5	4.2	4.6	4.9	3.7	3.7	3.6

Average family size homogeneously decreasing

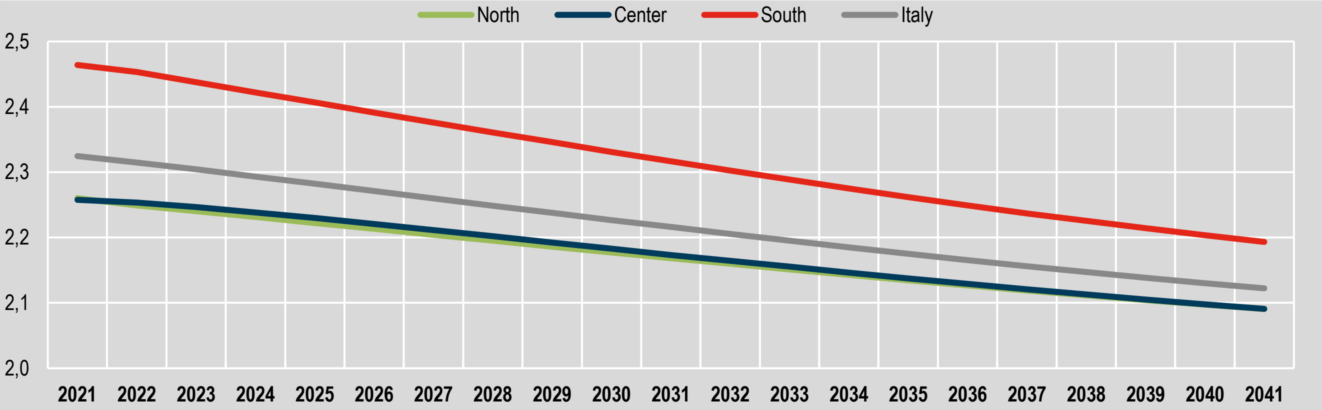
Gender differences in survival determine an annual growth in the number of women living alone. For Italy as a whole, it is expected that this type of family could make up about 23% of total families by 2041, from a current value of more than 19%, thus generating a variation of about 4 percentage points. In the Center and in the North (from 20% to 23%) the same variation is about 3 percentage points, while in the South is 5 (from 18% to 23%), giving rise to a process of convergence in the years to come.

On the contrary, among men who live alone, the process of convergence affects only the North with regard to the Center. Both geographical areas should show a 17% share of single-male households in 2041, compared to a national mean of 16%. The South maintains its territorial specificity, presenting a growing but more limited evolution, thanks to which the proportion of single men would be below 14% of total families by 2041. This is mainly due to two factors: on the one hand the presence of territorial differences as regards life expectancy, on the other hand the existence of a greater predisposition on the part of men, in particular in the South, to enter a second union in the case of widowhood or following the dissolution of a previous union.

Childless couples will continue to be more common in the North, despite achieving a modest increase (from 22% to over 23%). A more important change is expected in the South, where, in the face of a less widespread initial situation, couples without children would increase from 17% to about 20% of all families, reaching the values of the Center (from over 18% to about 20%). In the Center, however, single parents are more present, increasing their share from 12% in 2021 to about 14% in 2041, while in the North and the South the levels reached in 2041 would be approximately 11% and 12% of families.

The combination of the planned family transformations would ensure that the mean family size continues to decline, not only nationally (from 2.3 to 2.1 members), but also among the various geographical entities. The North and the Center, with very similar current values and future trajectories, will reach an average value of components close to the national mean. The South, thanks to past higher fertility rates, has always been characterized by the presence of families on average larger than in the rest of the country. Today, with more contained reproductive levels also in the South, this record (2.5 components) tends to become less clear and in 2041 the expectation is for a further decrease down to 2.2 components (Figure 3).

FIGURA 3. AVERAGE HOUSEHOLD SIZE BY GEOGRAPHIC AREA. Years 2021-2041, median scenario.



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.

Families with nucleus: includes couples with children, couples without children, single parents, families with two or more nucleus.

Families without nucleus: people living alone or multi-person families; this latter do not constitute a family unit even if composed of several people.

Family: 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.

Family nucleus: set of people who form a couple or a parent-child relationship. It means a married couple, civilly united or cohabiting, with or without children, or even a single parent together with one or more children who have never been married. Within a family there may be one or more family nucleus, but there may also be none, as in the case of families formed by an isolated member (single-component families) or several isolated members (other resident persons).

Family 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 projetions by age and sex. Years 2021-2070

Istat's 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 2020 published by Istat in November 2021. Istat is the owner and responsible for the production and dissemination of the projections, as documented in the National Statistical Program. The methodological framework underlying the current exercise is the same as that implemented in the previous three-year cycle, which resulted in the sequential release of the forecasts based on January 1st 2016, 2017 and 2018. 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
- 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 2021, 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 2021 and 1 January 2070. 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 (2070). 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 (2021). 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 86. They were voluntarily recruited among the participants in the 13th edition of the Population Study Days organized by the Italian Association for Population Studies (AISP), which took place in Milan between 24 and 26 January 2019 at Bocconi University. In particular, there are 50 women and 36 men, mainly employed in universities (21 in Northern Italy, 11 respectively in Central and Southern Italy and 10 in foreign universities) or in other public research bodies (24). The mean age of the respondents is 44 years while their work experience is on average 16 years.

In all the phases that involved the building of the methodological framework underlying the projections, Istat made use of the concrete cooperation of Francesco Billari and Rebecca Graziani of the Bocconi University in Milan.

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

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 2019, 2050 and 2080

Indicator	Total fertility rate	Life expectancy at birth – Men	Life expectancy at birth – Women	Immigrations (thousand)	Emigrations (thousand)
			Year 2019		
Observed value	1.27	81.1	85.4	333	180
			Year 2050		
Mean assumption	1.51	84.7	88.1	256	131
High assumption	1.75	86.3	89.5	343	172
Variance	0.034	1,441	1,309	4,593	1,042
			Year 2080		
Mean assum. conditional to 2050 mean assumption	1.55	87.1	90.0	240	127
Mean assum. conditional to 2050 high assumption	1.74	88.4	91.0	305	158
High assum. conditional to 2050 mean assumption	1.75	88.5	91.5	348	186
Variance	0.044	2,180	2,002	7,523	2,675
			Correlation 2050-2080		
Correlation coefficient	0.68	0.66	0.54	0.51	0.46

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 = 2019", "t1 = 2050", "t2 = 2080", thus generating two sub-intervals, 2019-2050 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.

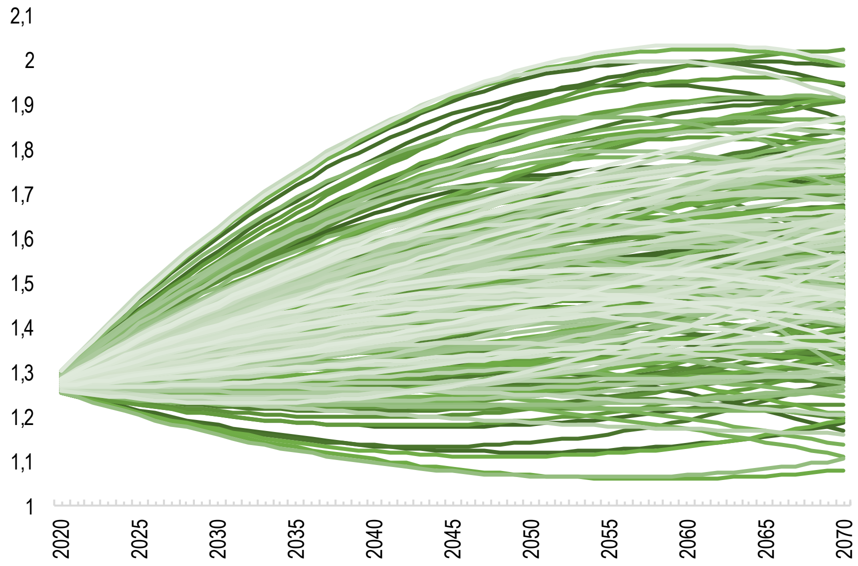
Under these conditions, once the mean values provided by the experts have been obtained, it is possible to estimate the variance associated with each of the two future time instants as well as the correlation between the first and the second period (Table A1). On the basis of the corresponding bivariate normal distributions, 3,000 simulations were then carried out with the Markov Chain Monte Carlo method[1].

The last estimation step is aimed at calculating the values of each parameter in the intermediate years of the two sub-intervals 2019-2050 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 2019, 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.

[1] The choice fell on the years 2050 and 2080 in order to identify two time intervals of similar length. Although the availability of information collected from the experts allows us to extend the time horizon up to 2080, it was decided to stop the iterative simulation exercise up to 2070 in order to derive a forecasting period of precise 50 years.

FIGURE A1. PROBABILISTIC EVOLUTION OF THE TOTAL FERTILITY RATE IN 3,000 SIMULATIONS OBTAINED FROM EXPERT OPIONIONS. YEARS 2020-2070



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 (2120, 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 (2020-2070). 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 1977-2019;
- for mortality, the (projection-)probabilities of death by age and sex for the period 1974-2019;
- for internal and international migrations, the changes of residence by age and sex of 2015-2019.

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. No forecasting model applied to the historical series mentioned in the previous paragraph could have been able to accurately predict the shock caused by the pandemic. Certainly not the over 740 thousand deaths found for all causes in 2020 (about 100 thousand more than expected), but not even the further decline in births (405 thousand) in the context of an overall picture already compromised by the well-known contraction of reproductive behaviours. Nor, finally, could the strong contraction recorded in migration be foreseeable, as a result of the measures undertaken at national level to contain the spread of the virus (lockdown).

Since the base population of the projections is that recorded as of January 1, 2021, it was necessary put in place some short-term correction of the predicted inputs that affected the first projections years. With this, in fact, we want to take into account not only the exceptional events that characterized the 2021, 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[2].

From the computational point of view, the review of the short-term assumptions is carried out by applying correction factors. For example, let E_b^j be the number of demographic events predicted in the first year based on the median scenario in region j. Instead, let $\widehat{\mathrm{E}}_b^j$ 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:

$$\mathrm{r}_\mathrm{b}^\mathrm{j} = \widehat{\mathrm{E}}_\mathrm{b}^\mathrm{j} / \mathrm{E}_\mathrm{b}^\mathrm{j}$$

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:

$$\hat{\mathrm{f}}_{\mathrm{b,x}}^{\mathrm{n,j}} = \mathrm{r}_\mathrm{b}^\mathrm{j} \cdot \mathrm{f}_{\mathrm{b,x}}^{\mathrm{n,j}} \text{ with x=14, … , 50 and n=1, … , 3000}$$

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 2021, the correction factors were constructed by comparing the data of the provisional demographic balance of each region, released in March 2022 by Istat, to the projections previously produced for that year[3].

For the years after 2021, 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:

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

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:

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

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

[2] Furthermore, bearing in mind the iterative calculation mechanism offered by the cohort component method, i.e. a mechanism of continuous stock-flow interaction over time, the correction imposed in the first years also affects the outcome of all subsequent years, up to 2070.

[3] Cfr.: Istat, la dinamica demografica – anno 2021, https://www.istat.it/it/archivio/267834.

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 2040, 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 1977-2019 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 1977-2019. 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 2120. In particular, the convergence constraint provides that, from 2020 to 2120, 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 1974-2019, 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 2070 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 1974-2019 with the same methodology and, subsequently, by making each regional parameter converge to the corresponding national parameter at 2120. 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 2015-2019. 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 (2120) at the same level, which is to the initial half sum of the two values.

Once the totals of inflows and outflows up to 2070 have been determined, the associated age and sex breakdown are derived by applying the Castro-Rogers model (Rogers and Castro, 1981) to the 2015-2019 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 2015-2019 period. 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 = 2015, ..., 2019), 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 2020 and 2021.

First of all, a rather limited difference should be noted between the total base population 2021 (59 million 236 thousand) and that which had been estimated in the median scenario on the same date by the projections based on 2020 (59 million 249 thousand).

On the side of the expected flows in the period of common projection (2020-2065), a slightly better assessment can be seen in the projections based on 2021, where for example 19.6 million births were expected against the 19.5 million of the previous exercise. Even the other components of the population change, although not such as to overturn the results emerged from the projections based on 2020, are more favourable for the exercise. The latter, in fact, has 69 thousand fewer deaths in the common projection period, 65 thousand more immigrants from abroad and 19 thousand fewer emigrants to abroad.

The difference between the final populations of the two distinct projections is also small (as of January 1, 2070, just 136 thousand units more for the median scenario 2021 based), confirming the substantial stability of the projections based on 2020, despite the change in the base population and short-term adjustments to the balance components. From this point of view, table A2 highlights how the process of reviewing the assumptions for all demographic components only affected the first years of forecasting.

TABLE A2. 2020 AND 2021 MEDIAN SCENARIO ASSUMPTIONS FOR THE MAIN DEMOGRAPHIC INDICATORS. Years 2021, 2030, 2050 and 2065.

Median scenario	Total fertility rate	Life expectancy at birth – Men	Life expectancy at birth – Women	Immigration (thousand)	Emigration (tohusand)
			Year 2021		
Base 2020	1.21	80.1	84.9	271	130
Base 2021	1.25	80.0	84.6	286	129
			Year 2030		
Base 2020	1.37	82.2	86.2	279	146
Base 2021	1.37	82.2	86.2	281	145
			Year 2050		
Base 2020	1.50	84.7	88.1	258	131
Base 2021	1.50	84.7	88.1	258	131
			Year 2070		
Base 2020	1.55	86.5	89.5	244	126
Base 2021	1.55	86.5	89.5	244	126

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 2019, 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 2021 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 Eurostat projections are based on 1 January 2019, that is, they project a population that is not in line with the results of the 2020 census and, in particular, with respect to the latter, a significantly higher population. Secondly, it should be mentioned 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).

For all the demographic components, the assumptions are initially very different between Istat and Eurostat/UNPD. This is due to the fact that, unlike the Istat scenario, Eurostat projections do not take into account the 2020 demographic shock produced by the Covid-19 pandemic. The UNPD projections, in turn, 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[4].

The assumptions on fertility are quite similar, although in the medium-long term Eurostat and UNPD produce less favourable forecasts than Istat. The assumptions on survival are also not particularly distant, however Eurostat

[4] 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).

and especially UNPD highlight very favourable expectations about the lengthening of life expectancy, which are only partially glimpsed in the Istat model.

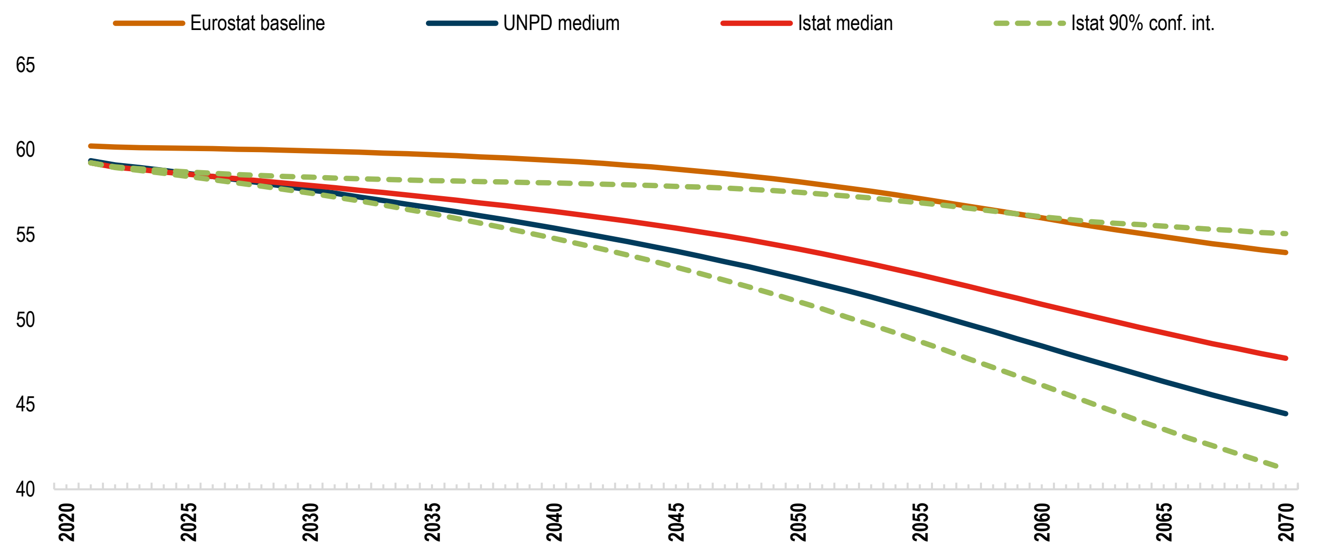
The development of the different demographic assumptions therefore gives rise to differences in terms of expected results which, as regards the evolution of the total population, can be appreciated in Figure A2. Unfortunately, the initial difference due to the different population bases adopted in the Eurostat projections make the comparison with the Istat scenario rather spurious. In turn, the UNPD projections, although aligned from the start with those of Istat, give a much more pessimistic evolution of the population.

TABLE A3. COMPARISON BETWEEN LATEST MAIN ASSUMPTIONS ON ITALY MADE BY ISTAT (MEDIAN SCENARIO), EUROSTAT (BASELINE) AND UNPD (MEDIUM). Years 2021, 2030, 2050 and 2070.

Scenario	Total fertility rate	Life expectancy at birth – Men	Life expectancy at birth – Women	Net migration (thousand)
		Year 2021		
Istat Median	1.25	80.0	84.6	157
Eurostat Baseline	1.33	81.4	85.8	178
UNPD Medium	1.28	80.5	85.1	28
		Year 2030		
Istat Median	1.37	82.2	86.2	136
Eurostat Baseline	1.37	82.6	86.9	224
UNPD Medium	1.35	83.2	87.1	58
		Year 2050		
Istat Median	1.50	84.7	88.1	127
Eurostat Baseline	1.45	84.9	89.0	214
UNPD Medium	1.44	85.8	89.6	58
		Year 2070		
Istat Median	1.55	86.5	89.5	118
Eurostat Baseline	1.52	87.0	90.9	207
UNPD Medium	1.48	88.2	91.9	58

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. 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. The UNPD scenario, on the other hand, tends to be about halfway between the Istat median scenario and its lower confidence interval.

FIGURE A2. TOTAL POPULATION ACCORDING TO ISTAT, EUROSTAT AND UNPD SCENARIOS. Years 2021-2070, million.



2) Regional households projections. Years 2021-2041

Household projections show the future trend of the number and type of households that will characterize the population in Italy from 2021 to 2041. These projections derive from the application of a static method, based on propensity rates, applied to the projected population. The purpose is to provide 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 durable 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 1.1.2021 while the elaborations cover the period from 2021 to 2041.

Data

Several set of data have been implemented. Among them, the official probabilistic projections - base 1.1.2021 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 1.1.2020 and 1.1.2021 to estimate the share of population living in institutional cohabitations by sex and single year of age; the Multipurpose Survey "Aspects of daily life", which provided with a long time series (from 2002 to 2021), to derive the family structures by typology and household position. The concept of household here in use is the "de facto family". Adopted in all Istat Multipurpose surveys, it identifies 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*".

Model

The model is based on an adaptation to the Italian context of the method known as "Propensity model". Such a model has been used for recent years by the Australian Bureau of Statistics (ABS) to project households in Australia and New Zealand (ABS, 2019). 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 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, there are also some drawbacks, which stem mainly from the static nature of the method, which does not allow the process of household formation and dissolution to be reproduced. Thus, the application of propensity rates to the resident population may in some cases lead to inconsistencies in terms of overall results, such as between sexes or for household positions within age groups, a problem that therefore needs to be resolved by ex-post adjustments.

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 by household position
- Step 5. Calculate the number, type and size of projected households.

The various steps are explained in detail below.

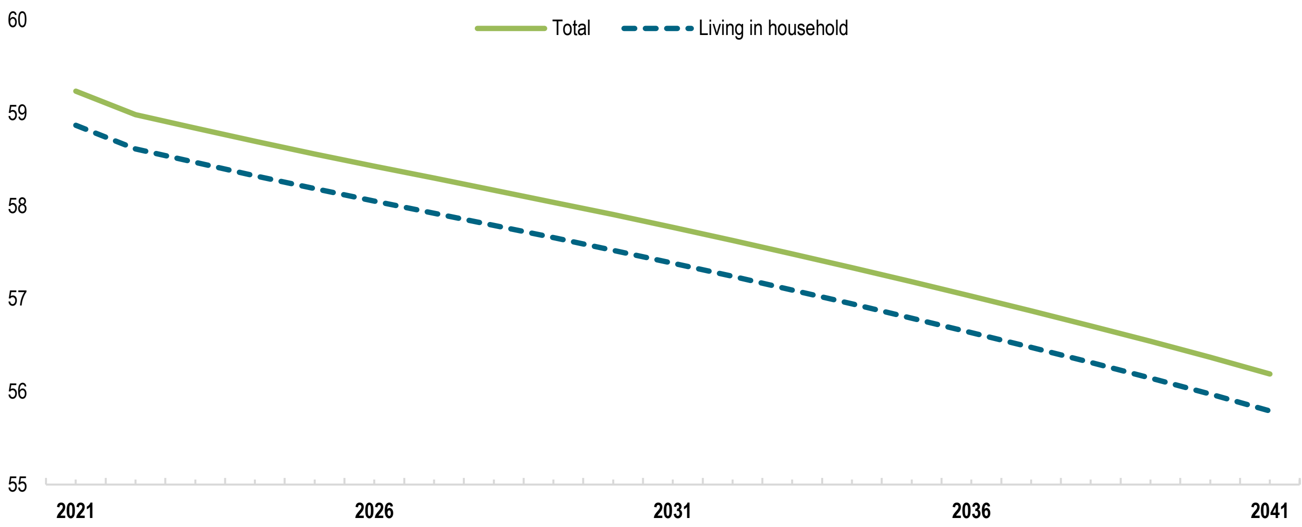
Step 1. Estimate the base-year and projected population living in households

The base year population is the total resident population by sex, age, and region as of January 1, 2021, as collected by the latest register-based Census of Population. Therefore, it is necessary to make a preliminary estimate of the resident population in households, excluding individuals residing in institutional cohabitation (hospitals, barracks, prisons, nursing homes, religious buildings, etc.). This was then repeated for each projection year, subtracting from the regional projections (whose reference is again the total resident population) the share of the institutionalized population.

Using data from the Basic Population Registry as of 1.1.2020 and 1.1.2021, the (average) percentage incidence 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, a constant trend was assumed throughout the forecast period.

By applying the 100's complement of these incidences to the total population from 2021 to 2041, we obtained the population living in the household by region, sex and age class (Figure A3).

FIGURE A3. TOTAL POPULATION AND POPULATION LIVING IN HOUSEHOLD. Years 2021-2041, 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. Compared with the previous edition, 11 family positions are considered, and no longer 8, as the distinction is introduced for parents according to the age of the children (whether "up to 19 year-old" or "20 years and older").

1. lone person;
2. person in multi-person household (e.g., 2 siblings living together);
3. person in a childless couple;
4. person in a couple with at least one child under 20;
5. person in a couple with only children 20 and over;
6. lone parent with at least one child under 20;
7. lone parent with only children 20 and over
8. child under 20 (living with at least one parent);
9. child 20 and over (living with at least one parent);
10. other person living in a one-family household[5];
11. person in a household with 2 or more families.

1.

Positions 3 through 10 refer to individuals in one-family households. In general, position "child" follows the definition adopted in all Multiscope surveys, which considers children only if never married. People living in households with 2 or more families 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):

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

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

[5] 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).

LAPs are calculated using data from the *Aspects of Daily Life* (AVQ) survey, along the entire 2002-2021 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[6], 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

The prediction of the future evolution of the household propensities was based on the introduction of a synthetic indicator summarizing the population's family behaviors over time. This indicator, which we will indicate by the name Total Intensity Rate by household Position (TPT), is given by the sum by age of LAP weighted by years lived at various ages.

$$\mathrm{TPT}_{i,s,t} = \sum_{x=0-4}^{85+} \mathrm{LAP}_{x,i,s,t} * \mathrm{L}_{x,s,t} = \sum_{x=0-4}^{85+} \frac{\mathrm{P}_{x,i,s,t}}{\mathrm{P}_{x,s,t}} * 100 * \mathrm{L}_{x,s,t}$$

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

$\mathrm{L}_{x,s,t}$, 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.

Under the hypothesis of independence between mortality and household position, the TPT 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.

TABLE A4. TOTAL INTENSITY RATES BY FAMILY POSITION (TPT) AND SEX. Years 2002-2021

Household position	MEN					WOMEN				
	2002	2007	2012	2017	2021	2002	2007	2012	2017	2021
Lone person	5.8	6.4	8.1	8.9	9.1	10.4	11.1	12.1	12.4	12.8
Person in a childless couple	13.1	14.2	14.1	14.2	13.1	12.1	13.0	13.0	13.2	12.4
Person in couple with at least 1 child aged <20	14.5	13.8	12.8	12.7	12.5	14.3	13.9	13.1	13.1	13.1
Person in couple with only children aged >=20	8.8	8.5	7.9	7.3	6.9	8.5	8.1	7.2	6.7	6.5
Lone parent with at least 1 child aged <20	0.2	0.2	0.3	0.3	0.4	1.2	1.4	1.9	2.0	2.1
Lone parent with only children aged >=20	0.5	0.6	0.8	0.8	0.9	2.9	3.0	3.2	3.1	3.1
Child aged <20	19.2	19.3	19.1	19.0	19.0	19.1	19.3	19.0	19.0	18.9
Child aged >=20	11.2	11.5	11.9	12.2	13.0	8.5	8.7	8.9	8.9	10.1
Person in multipersonal household	0.9	1.2	1.3	1.4	1.5	1.7	1.6	1.5	1.7	1.7
Person in a household with 2 or more families	2.1	2.0	2.5	2.7	2.7	2.3	2.2	2.7	3.2	2.9
Other person	1.0	0.9	1.0	1.0	1.0	2.0	1.7	1.7	1.4	1.3
Totale	**77.2**	**78.6**	**79.7**	**80.5**	**80.1**	**83.0**	**83.9**	**84.5**	**84.8**	**84.8**

As can be seen in Table A4, while in 2002 a man counted on living as a single person for an average of 5.8 years (out of a total life expectancy of 77.2), in 2021 the expected time in this state rises to 9.1 years (out of a total of 80.1). As a result of the declining birth rate, on the other hand, in 2002 women expected to live as a person in a couple with at least one child until age 19 for 14.3 years (out of a total of 83), but in 2021 this expected time has dropped to 13.1 years (albeit out of a total life expectancy that has since risen to 84.8 years). As a final example,

[6] 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.

the time in the "child" state increased from 30.4 to 31.2 years for males and from 27.6 to 29.0 for females, due to the prolonged stay of young people within the family of origin.

In order to hypothesize future trends in propensities, we proceeded to project the total intensity rates by household position, and then to estimate its distribution broken down by age group ($LAP_{x,i,s,t}$) 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).

The final goal of Step 3, which is to define the projected LAPs from 2021 to 2041 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 household position and sex ($TPT_{i,s,t}$) was predicted through trend extrapolation over the period 2002-2021, using time series analysis models. For each household position, sex and geographical group the models described in Table A5 were applied.

The assumptions regarding the future evolution of the TPTs emerge from Figure A4, where, as an example, the North-West territorial group is shown. On brief, the variations in time spent in different household positions express the following changes:

- an increase of “lone persons”;
- a decrease of “partners with children”;
- a slight increase of “partners without children”;
- an increase of people in “child” position;
- a slight increase of “lone parents”, especially fathers;
- a substantial stability of "other people" living with a family household and of “persons in households with 2+ families”.

TABLE A5. PREDICTIVE MODELS* FOR TOTAL INTENSITY RATES BY HOUSEHOLD POSITION AND SEX.**

Household position	Men	Women
Lone person	RWD || ARIMA(1,0,0)	RWD
Person in a childless couple	ARIMA(2,0,0)	RWD
Person in couple with at least 1 child aged <20	RWD	RWD
Person in couple with only children aged >=20	RWD || ARIMA(2,1,0)	RWD || ARIMA(2,1,0)
Lone parent with at least 1 child aged <20	Linear Trend	Linear Trend
Lone parent with only children aged >=20	RWD	RWD
Child aged <20	RWD	RWD
Child aged >=20	AR2	ARIMA(2,1,0)
Person in multipersonal household	RWD	RWD
Person in a household with 2 or more families	ARIMA(1,1,0)	ARIMA(1,1,0)
Other person	RWD	ARIMA(1,0,0)

*Prevailing model for the 5 territorial groups. **RWD=*Random Walk with Drift model*; ARIMA=*AutoRegressive Integrated Moving Average model*.

Estimation of the predicted distribution by age, i.e. $LAP_{x,i,s,t}$ from 2021 to 2041, was obtained using predicted TPT, predicted years lived and observed distributions from AVQ survey data over the three-year period 2019-21. For this latter purpose, the mean 2019-21 distributions by position in the household were weighed by two coefficients: one to account for the predicted TPT in year t relative to that in the 2019-21:

$$WP_{s,i,t} = \frac{TPT_{s,i,t}}{TPT_{s,i,2019-21}} \qquad t = 2021, \dots, 2041$$

and a coefficient expressing changes in mortality over time:

$$WL_{x,s,t} = \frac{L_{x,s,2019-21}}{L_{x,s,t}} \qquad t = 2021, \dots, 2041$$

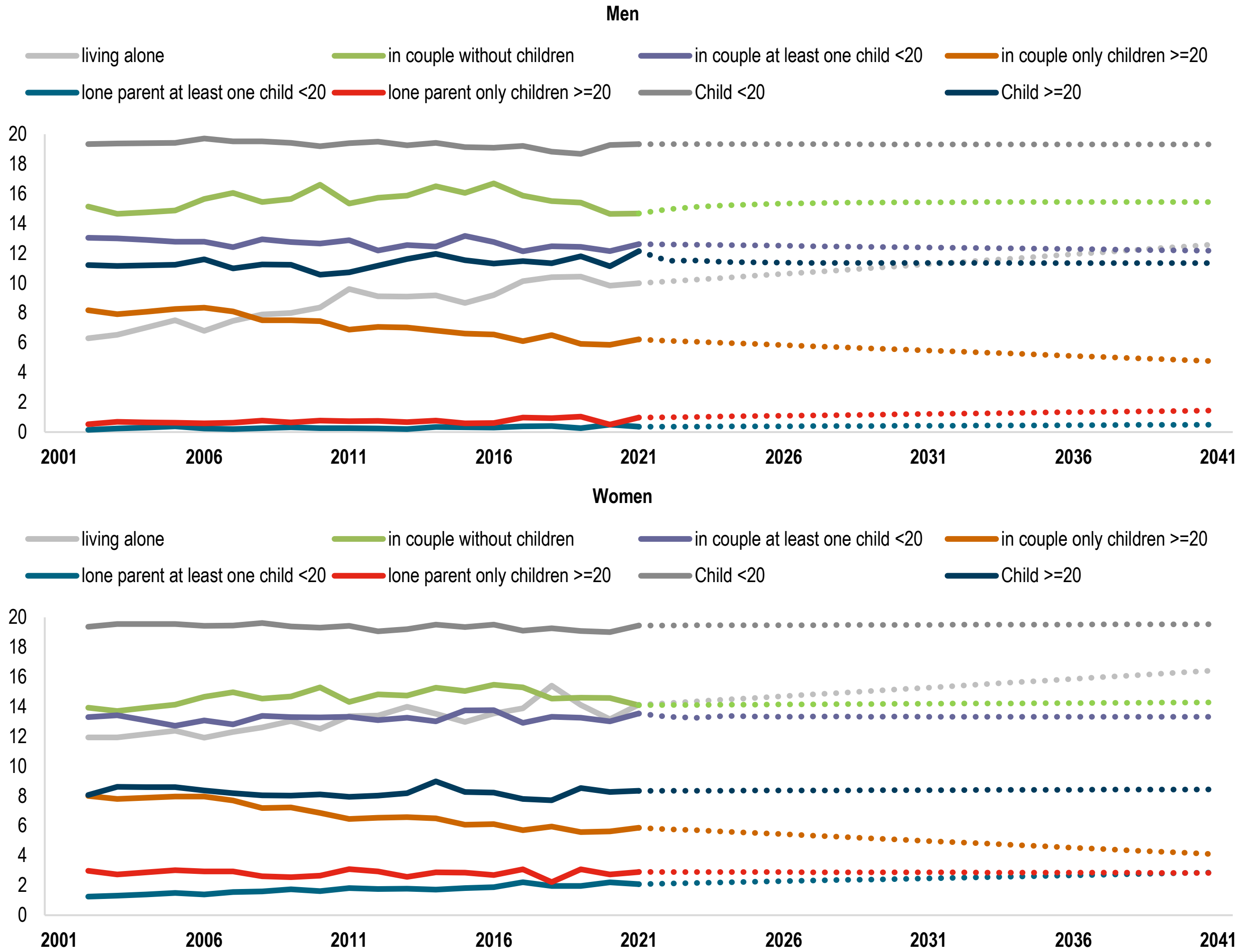
Therefore, household propensities throughout the projection horizon were calculated using the formula:

$$LAP_{x,s,i,t} = LAP_{x,s,i,2019-21} * WP_{s,i,t} * WL_{x,s,t} \qquad t = 2021, \dots, 2041$$

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

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 2019-2021. 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.

FIGURA A4. TOTAL INTENSITY RATES BY FAMILY POSITION. North-west, Years 2002-2041



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:

$$FC_{r,i} = \frac{TPT_{2019-21,i,r}}{TPT_{2019-21,i,G}}$$

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

Next, the projected LAPs for the spatial groups are multiplied by the regional adjustment factor calculated in this way, determining the series of regional LAPs from 2021 to 2041. For example, for the male single-person household position, the TPT found in Piemonte is 11.06 while in group 1 it is 10.08. The correction factor is therefore equivalent in this case to 1.10. This means that, since Piemonte has a TPT higher than that of the group to which it belongs, it is necessary to make an adjustment by multiplying all the LAPs at the various ages and the various projection years by 1.10, 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. We thus derive the projected population by household status, sex, age group and region from 2021 to 2041.

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 5,2 and 5.5 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 family position of the population, on the distribution of families by type and on the mean number of family 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 demographic projections within the framework of the line of activity "Population estimates and projections", in accordance with the provisions of the National Statistical Program (code PSN IST-01448).

References

1. ABS - Australian Bureau of Statistics (2019), *Household and Family Projections*, *Australia. Household and family projections (based on different assumptions of living arrangements) for Australia, states and territories and capital cities. Reference period: 2016 – 2041*. https://www.abs.gov.au/.
2. ABS - Australian Bureau of Statistics (2019), *Household and Family Projections, Australia methodology. Reference period: 2016 – 2041*. https://www.abs.gov.au/methodologies/household-and-family-projections-australia-methodology/2016-2041
3. Alho J. e Keilman N. (2010), *On future household structure*. Journal of the Royal Statistical Society Series A, 2010, vol. 173, Issue 1, 117-143.
4. Alho J.M. e Nikeer T. (2004), *Uncertain population of Europe - summary results from a stochastic forecast*. http://www.stat.fi/tup/euupe/rp_reports_e_pub.html.
5. Alho J.M., Spencer B.D. (2005), *Statistical demography and forecasting*. New York: Springer.
6. Bernard A., Bell M. (2012), *A Comparison of Internal Migration Age Profile Smoothing Methods*. Working Paper 2012/01. Queensle Centre for Population Research, The University of Queensle.
7. Billari, F.C., Corsetti G., Graziani R., Marsili M. e Melilli E. (2014), *Towards stochastic forecasts of the Italian population: an experiment with conditional expert elicitations*. Proceedings of the 6th Eurostat/UNECE Work Session on Demographic Projections, pagg. 326-338, Istat, 2014.
8. Billari, F.C., Corsetti G., Graziani R., Marsili M. e Melilli E. (2014), *A stochastic multi-regional model for Italian population projections*. Budapest, 25-28 giugno 2014, European Population Conference. http://epc2014.princeton.edu/papers/140361.
9. Billari, F.C., Graziani R. e Melilli E. (2012), *Stochastic population forecasts based on conditional expert opinions*. Journal of the Royal Statistical Society. Series A. 175(2): 491-511.
10. Blangiardo G., Barbiano di Belgiojoso E., Bonomi P. (2012), *Le previsioni demografiche delle famiglie*. In: Donati P. (a cura di), *La famiglia in Italia. Sfide sociali e innovazioni nei Servizi*. Osservatorio Nazionale sulla Famiglia. Rapporto biennale 2011-2012. Volume I Aspetti demografici, sociali e legislativi. Pagg 91-123.
11. Box G. E. P., Jenkins G. M., Reinsel G. C. and Ljung G. M. (2015), *Time Series Analysis: Forecasting and Control*, 5th Edition. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1.
12. Booth H. (2006), *Demographic forecasting: 1980 to 2005 in review*, International Journal of Forecasting, 22: 547–581.
13. Cooper J., Bell M. e Les M. (1995), *Household and Family forecasting Models: a review.* Review of IPC Long-Terrm Projections Model. Paper No. 3. Australian Housing and Urban Research institute in conjunction with Demographics Australia.
14. CBS (2011), *Key figures of the population forecasts 2010-2060*. Statline, Centraal Bureau voor der statistiek, http://statline.cbs.nl/statweb/.
15. Corsetti G., Marsili M. (2012), *A stochastic population projection from the perspective of a national statistical office*. European Population Conference. Stoccolma, 13-16 giugno 2012, EAPS. http://epc2012.princeton.edu/papers/120635.
16. Corsetti G., Marsili M. (2013), *Previsioni stocastiche della popolazione nell'ottica di un Istituto nazionale di statistica*. Rivista di statistica ufficiale, n. 2-3, p. 5-29, Istat.
17. Eurostat (2015), *People in the EU: who are we and how do we live?* - 2015 edition, Luxembourg: Publications Office of the European Union.
18. Eurostat (2020), *Methodology of the Eurostat population projections 2019-based (EUROPOP2019)*, Technical Note, Directorate of Social statistics, Population and migration, Luxembourg, 6 April 2020.
19. Graziani R., Keilman N. (2011), *The sensitivity of the Scaled Model of Error with respect to the choice of the correlation parameters: A simulation study*. Working Paper 37. Carlo F. Dondena Centre for Research on Social Dynamics, Università Bocconi, Milano.
20. Istat (1989), *Previsioni della popolazione residente per sesso, età e regione – Base 1.1.1988*, Note e Relazioni, n.4.
21. Istat (1997), *Previsioni della popolazione residente per sesso, età e regione – Base 1.1.1996*, Informazioni, n. 34.
22. Istat (1989), *Previsioni del numero di famiglie italiane dal 1995 al 2020*. Notiziario. Serie 4. Foglio 41 – Anno X No. 19. Dicembre 1989.
23. Istat (2001), *Previsioni della popolazione residente base 1° gennaio 2000*, Statistiche in breve, www.istat.it.
24. Istat (2003), *Previsioni della popolazione residente per sesso, età e regione, base 1.1.2001*, Informazioni n.13.
25. Istat (2006), *Previsioni demografiche nazionali 1° gennaio 2005-1° gennaio 2050*, www.istat.it, Nota informativa, 22 marzo 2006.

26. Istat (2008), *Previsioni demografiche 1° gennaio 2007-1° gennaio 2051*, www.istat.it, Nota Informativa, 19 giugno 2008.
27. Istat (2011), *Il futuro demografico del paese - Previsioni regionali della popolazione residente al 2065*, Statistiche Report, www.istat.it, 28 dicembre 2011.
28. Istat (2017), *Il futuro demografico del Paese – Previsioni regionali della popolazione residente al 2065*, Statistiche Report, www.istat.it, aprile 2017.
29. Istat (2018), *Il futuro demografico del Paese – Previsioni regionali della popolazione residente al 2065 – base 1.1.2017*", Statistiche Report, www.istat.it, maggio 2018.
30. Istat (2019), *Il futuro demografico del paese - Previsioni regionali della popolazione residente al 2065 – base 1.1.2018*, Nota metodologica, www.istat.it, ottobre 2019.
31. Istat (2021), *Ricostruzione della popolazione residente per sesso, età e comune, Anni 2002-2018*, Nota informativa, www.istat.it, marzo 2021.
32. Istat (2021), *Previsioni della popolazione residente e delle famiglie – base 1.1.2020*, Statistiche Report, www.istat.it, novembre 2021.
33. Istat (2022), *Indicatori demografici – Anno 2021*, Statistiche Report, www.istat.it, 8 aprile 2022.
34. Keilman N. (2018) *Family Projection Methods: A Review*, March 2018, DOI: 10.1007/978-3-319-93227-9_12. In book: Analytical Family Demography. Publisher: Springer
35. Keilman, N. e Brunborg, H. (1995), *Household Projections for Norway, 1990-2020*, Part I: Macrosimulations, Statistics Norway, Oslo-Kongsvinger.
36. Keilman, N. Kuijsten A. and Vossen A. (1988), *Modelling Household Formation and dissolution*. Clarendon Press - Oxford.
37. Keilman N., Pham D.Q. e Hetle A. (2002), *Why population forecasts should be probabilistic - illustrated by the case of Norway*, Demographic Research, 6(15): 409-454.
38. Kono S. (1987), *The headshhip rate method for projecting households*, in Bongaarts J., Burch T., Wachter K., *Family Demography, Methodhs and their Applications*, Clarendon Press-Oxford.
39. Lavit, C. (1988), *Analyse conjointe de tableaux quantitatifs*, Masson, Paris.
40. Lavit, C., Escoufier, Y., Sabatier, R. and Traissac, P. (1994), *The ACT (Statis method)*. Computational Statistics and Data Analysis, Volume 18, Issue 1, 97–119.
41. Lee R.D. (1998), *Probabilistic Approaches to Population Forecasting,* Population e Development Review, 24: 156-190.
42. Lee R.D., Carter L.R. (1992), *Modeling and forecasting U.S. Mortality*, Journal of the American Statistical Association, September, vol. 87, n.419.
43. Lee R.D., Miller T. (2001), *Evaluating the performance of the Lee-Carter method for forecasting mortality*, Demography, November, vol. 39, p. 537-549.
44. Lutz, W., Sanderson W.C. e Scherbov S. (1998) *Expert-Based Probabilistic Population Projections*, Population e Development Review, 24: 139-155.
45. McDonald, P. and R. Kippen. (1998), *Household Trends and Projections: Victoria 1986–2011*. Canberra: Demography Program, The Australian National University.
46. 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.
47. 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.
48. ONS – Office for National Statistics (2020), *Methodology used to produce household projections for England: 2018-based. User guidance about uses, methodology, assumptions and input data for household projections for England.* https://www.ons.gov.uk.
49. Paciorek, A. (2014), *The Long and the Short of Household Formation*. Real Estate Economics, Forthcoming, Available at SSRN: https://ssrn.com/abstract=2469334
50. Rogers A. (1985), *Regional Population Projection Models*. Beverly Hills. CA: Sage.
51. Rogers A., Castro L. (1981) *Model migration schedules*, International Institute for Applied System Analysis, Laxenberg, Austria, RR-8 1-30, November 1981.
52. Rowan S., Wright E. (2010), *Developing stochastic population forecasts for the United Kingdom: Progress report e plans for future work*. Eurostat-UNECE Work session on demographic projections, Lisbon, 28-30 aprile 2010, Methodologies e Working papers, Commissione Europea.
53. Schmertmann C.P. (2003), *A system of model fertility schedules with graphically intuitive parameters*, Demographic Research, 9(5): 81-110.
54. Shaw C. (2008), *The National Population Projections Expert Advisory Group: results from a questionnaire about future trends in fertility, mortality e migration*. Population trends n.134, Winter 2008, Office for national statistics.
55. Stoto, M. A. (1983), *The accuracy of population projections*. Journal of the American Statistical Association. 78: 13–20.
56. Tuljapurkar S., Lee R.D. e Li Q. (2004), *Random scenario forecast versus stochastic forecasts*. International Statistical Review. 72: 185–199.
57. Terra Abrami V. (1998), *Le previsioni demografiche*, Il Mulino, Bologna.

58. UNECE (2018), *Recommendations on Communicating Population Projections*, United nations economic commission for Europe, United Nations, New York and Geneva, agosto 2018.
59. United Nations(1973), *Methods of projecting households and families*, Manual VII, New York.
60. United Nations (2019), *World Population Prospects 2019: Methodology of the United Nations population estimates and projections*, Department of Economic and Social Affairs/Population Division, ST/ESA/SER.A/425.
61. Wilson T. (2013), *The sequential propensity household projection mode*, DEMOGRAPHIC RESEARCH VOLUME 28, Article 24, Pages 681-712, http://www.demographic-research.org/Volumes/Vol28/24/ DOI: 10.4054/DemRes.2013.28.24.