UDK: 314.48:314.74](497.2)
314.74:314.48](497.2)
311.3:314.74](497.2)
Izvorni znanstveni rad
Primljeno: 14. 05. 2003.
Prihvaćeno: 10. 06. 2003.

NIKOLA TCHOLAKOV
University of National and World Economy, Sofia

Migration Statistics, Population Accounts and Life Tables – the Methodology Used in Bulgaria

SUMMARY

Vital statistics in Bulgaria are considered to be comparatively detailed, complete and precise. Professionals at the National Statistical Institute prefer to use these data together with more direct methods to produce life tables with single year age groups separately for males and females. However, external migration might present a conceptual problem for such methods, while more sophisticated methodologies might not be applicable because migration data are usually scarce and very unreliable. Migration might affect the derivation of life tables in two main ways. First, migration is a competitive risk to mortality and should be encompassed carefully, particularly for years with a high number of emigrants as is the typical case for Bulgaria. Secondly, loose migration registration or any other deficiencies or distortions in migration data might appear to pose a serous problem in the use of population accounts in life table calculations, and for statistical analysis in general. Toward the end of the 1960s there was a significant outflow of migrants to Turkey which was never mirrored by statistical data in the same detail as mortality. Underground and illicit emigration from Bulgaria after World War II and the massive departure of Turks during the 1980s might also have been a factor in misjudging the size of population groups between censuses and, hence, under- or overestimating mortality. Given the reluctance of the political authorities at that time to reveal these facts to the public, it was very difficult for statisticians to obtain relevant data, perform life table calculations in a consistent manner or produce reliable demographic analysis. This paper examines the methodology of the Institute and addresses some related issues made upon observations of the practice of the official statistics in Bulgaria in this area.

KEY WORDS: vital statistics, migration, natural increase, life table, population accounts, population balance, population projection

The Bulgarian statistical authority – the National Statistical Institute[1] – published life tables periodically during the second half of the last century. These could be found in several issues of the *Statistical Yearbook of Bulgaria*, in the more specialized yearbook *Population* and in the dedicated brochure *Mortality and Life Expectancy Tables for the Population of Bulgaria*, which was published by the Institute once or twice in a decade. Several independent Bulgarian authors (Сугарев, 1975; Чолаков, 1975; Малешков, 1982) have dealt with this matter, making a lot of comparisons and suggesting different me-

[1] The National Statistical Institute (www.nsi.bg) is the Bulgarian statistical authority. In the past, it used to be an independent body governed by Parliament, a division of the Ministry of Transportation and Communication, a separate administrative unit, with different names such as Directorate of Statistics, Central Statistical Office, State Administration of Information, Committee for Unified System of Social Information and others. The abbreviation NSI will be used to denote this Institute.

thods for the construction of life tables, incl. life tables by causes of death. Data for the population of Bulgaria have also been used abroad (Keyfitz, 1968; Preston, Keyfitz and Schoen, 1972) to obtain corresponding life tables.

The first life table for the Bulgarian population emerged in 1919, corresponding to mortality during 1900–1905 separately for males and females. Several other life tables were published later, mainly around census years 1921–26, 1927–34 and 1935–39, produced by the Bulgarian Academy of Sciences. Old-age accumulation used to be a serious burden for deriving life table indicators at that time and therefore the age/sex structures were subject to careful smoothing (National Statistical Institute, 1983).

The life table for 1956–57 is derived by NSI from vital statistics for that period categorized by age, sex and generation, and corresponding age/sex distribution of the population from the 1956 census. Almost the same approach was applied later. Since 1960, several life tables for males and females have been derived from vital statistics data along with the projected age/sex structure of the population based on information from the last census. For the 1960–62 life table for example, the 1956 census data were supposedly used along with vital statistics and migration data for the period 1956–1962. Old-age-accumulation proved to be negligible (National Statistical Institute, 1983) and population accounts were assumed to be accurate. Together with 1960–62, life tables created on these lines are 1969–71, 1978–80, 1988–90, 1995–97 and 1998–2000. The 1965–67, 1974–76, 1984–86 and 1999–2001 life tables are based on 1965, 1975, 1985 and 2001 census data respectively, thus implementing the 1956–57 model and avoiding the need for population projection.

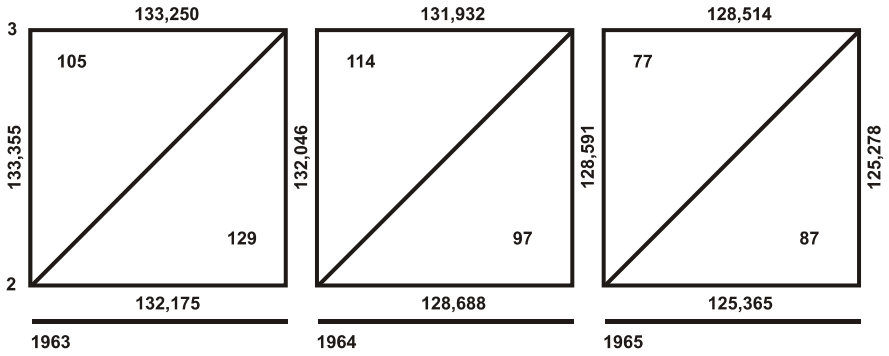
Vital statistics in Bulgaria are considered to be comparatively detailed, complete and precise. Professionals at the National Statistical Institute prefer to use these data together with more direct methods and to produce life tables with single year age groups separately for males and females. However, external migration might present a conceptual problem for such methods, while more sophisticated methodologies might not be applicable because migration data are usually scarce and very unreliable. Migration might affect the derivation of life tables in two main ways. First, migration is a competitive risk to mortality and should be encompassed carefully, particularly for years with a high number of emigrants as is the typical case for Bulgaria. Secondly, loose migration registration or any other deficiencies or distortions in migration data might appear to pose a serous problem in the use of population accounts in life table calculations.

Toward the end of the 1960s there was a significant outflow of migrants to Turkey which was never mirrored by statistical data in the same detail as mortality. Underground and illicit emigration from Bulgaria after World War II and the massive departure of Turks during the 1980s might have also been a factor in misjudging the size of population groups between censuses and, hence, under- or overestimating mortality. Given the reluctance of the political authorities at that time to reveal these facts to the public, it was very difficult for the statisticians to obtain relevant data, perform life table calculations in a consistent manner or produce reliable demographic analysis. This paper examines the methodology of the Institute and addresses some related issues made upon observations of the practice of the official statistics in Bulgaria in this area.

Life table methodology of National Statistical Institute of Bulgaria

This technique has best been described in specialized publications like *Mortality and Life Expectancy Tables for the Population of Bulgaria* (National Statistical Institute, 1983, 1988) and its logic is clearly presented by Keyfitz (1968: 9–11) with the help of the Lexis diagram, although Bulgarian practice differs in some minor details (Русев, 1973). The method is known among Bulgarian demographers as the Becker-Zeiner method and has been in usage since 1960.

Figure 1: Mapping of three successive generations between age 2 and 3



Consider Fig. 1 which is a small part of the Lexis diagram for the population of Bulgaria for the years 1963–1965.[2] It shows a single age slice which represents the life of a couple of generations between the exact ages of 2 and 3. The horizontal figures show the number of children who have reached their second and third birthday. One can see that from the generation born in 1962 (initial number 134,148, not shown on the chart) 128,688 reached the exact age of 2 during 1964. Then, 97 of them died before the end of the year; hence 128,591 survived the beginning of 1965. During the following year, 77 died before their third birthday; therefore, only 128,514 reached that anniversary. Thus, the vertical numbers denote survivors to the end of each year with the figures within triangles denoting deaths for a given year, additionally categorized by age and year of birth.

Such changes could be tracked the opposite way. Starting from the 1965 Population Census data with 125,278 at age 2 but not 3, add the number of deaths 87 to obtain 125,365, which is actually the number of the 2nd birthdays during 1965.

Once the data on mortality are available in the form of triangles categorized by age, generation and year of death (possibly by sex too) it is very easy to reconstruct the whole generation down to the beginning. There should be a clear correspondence between the data on life births and consecutive deaths, while the generation balance in a forward pro-

[2] Data from the specialized demographic yearbook for the period under consideration (this changed title several times during the second half of the 20th century: *Демографска статистика*, *Население*, *Население и демографски процеси*) and the specialized publication *Население по пол и възраст на НР България – общо и по окръзи през периода 1965 –1967 г.*, see references.

jection should match up exactly with the corresponding figures of future population censuses; there is, however, one very important exception – migration. Lack of adequate data on migration might destroy the logic of this simple and elegant chain calculation method suggested by Wilhelm Lexis long ago. Bulgarian vital statistics provide all the necessary data to keep this population account consistent only in the absence of migration.

The Lexis diagram shows how easy it is to calculate two key age-specific indicators. One is the probability a person who has reached a given exact age to reach the end of the current year. Fig. 1 suggests three possibilities to derive this probability for age 2. The data for 1963 yield:

$$p_2' = \frac{132,046}{132,175} = 0.999024.$$

For 1964:

$$p_2' = \frac{128,591}{128,688} = 0.999246.$$

And for 1965:

$$p_2' = \frac{125,278}{125,365} = 0.999306.$$

It is apparent that later generations seem to have a better chance of survival but the observed difference might not be statistically significant enough to support such a hypothesis. In any case, an average probability can be derived from the above data to reflect the whole 3-year period:

$$p_2' = \frac{132,046+128,591+125,278}{132,175+128,688+125,365} = 0.999190.$$

The other indicator is the probability of a person alive at the end of a given year to survive their next birthday. Fig. 1 again provides three possibilities to derive these probabilities. For 1963:

$$p_2'' = \frac{133,250}{133,355} = 0.999213.$$

For 1964:

$$p_2'' = \frac{131,932}{132,046} = 0.999137.$$

And for 1965:

$$p_2'' = \frac{128,514}{128,591} = 0.999401.$$

The average probability for the whole three year period is:

$$p_2'' = \frac{133,250+131,932+128,514}{133,355+132,046+128,591} = 0.999249.$$

This and the previous averaging procedures are implemented by NSI in order to avoid random fluctuations in the data of separate years as suggested by Sugarev (Сугарев, 1975: 163–164).

Using these two types of probability it is easy to derive all life table indicators in a simple and straightforward way. The backbone for all calculations are l_x and L_x (x = 0,1,2,…). The starting point is l_0=100,000 which is used to calculate the corresponding age group of the stationary population: $L_0 = l_0 p'_0$. Next comes $L_0 = l_0 p''_0$. Using the next layer of data in the Lexis diagram one may obtain $L_1 = l_1 p'_1$ and $l_2 = l_1 p''_1$, etc. The recursive formulae are:

$$L_x = p' l_x \tag{1}$$

$$l_{x+1} = p_x' p''_x \text{ or } l_{x+1} = p''_x L_x \tag{2}$$

for x = 0,1,2,…, with initial value l_0 = 100,000.

Obviously, the probability alive at the exact age x to survive until the next exact age $x + 1$ is $p_x' p''_x$ and the probability alive at age x to $x + 1$ to survive for the next 365 days is $p''_x p'_{x+1}$.

Quite surprisingly, however, one discovers that (1) is used for the first five ages only x = 0,1,2,3,4 and for the remaining $x \geq 5$, NSI adopts another approximation (Keyfitz, 1968; Сугарев, 1975; Чолаков, 1975) with a different formula being used instead:

$$L_x = \frac{l_x + l_{x+1}}{2} + \frac{d_{x+1} - d_{x-1}}{24}$$

which is based on the precise relationship (Keyfitz, 1968: 6):

$$L_x = \int_0^1 l_{x+t} dt \tag{3}$$

This rather unusual shift in methodology is difficult to explain. If mortality data (the triangles in Fig. 1) or population accounts (and/or census data) are unreliable or distorted in any way, then not only (1) but (2) might produce invalid results and a quite separate approach to life table calculation would be called for.

The rest of the life table indicators could easily be calculated from l_x and L_x. Life expectancies, which are the corner stone of life tables, could be obtained by first totaling L_x:

$$T_x = L_x + L_{x+1} + L_{x+2} + \ldots$$

and then taking *per capita* life potential[3]:

$$\mathring{e}_x = \frac{T_x}{l_x}$$

Life expectancy in Bulgaria has shown a steady increase over the last couple of decades: from about 66 years in the 1950's to 72 at the turn of the century. The increase comes mainly from a substantial decrease in infant mortality and is more significant for

[3] Formula (3) is the key to life expectancies, not L_x as a stationary population.

females than males. Nowadays[4] life expectancy at birth for females is over 74 years, while for males it is roughly 68. This difference was about 4 years half a century ago.

Early age mortality and related issues

There are some other observations connected to infant mortality and the way NSI calculates the corresponding life table indicators. Less known in Bulgaria is the following indicator (Keyfitz, 1968: 12) which shows the average number of years lived in the interval x to $x + 1$ by those who die in it:

$$a_x = \frac{L_x - l_{x+1}}{l_x - l_{x+1}} \qquad (4)$$

Fig. 1 provides two means of calculating this indicator. For the generation born 1961:

$$a_2 = \frac{114}{129 + 114} = 0.469$$

and for the next generation:

$$a_2 = \frac{77}{77 + 97} = 0.443$$

As expected, the value of this indicator should be close to 0.5 as long as the mortality risk is uniformly distributed within the age interval x to $x + 1$. This is not quite the case for the early ages, where the risk rapidly decreases from relatively high infant mortality at the beginning of life. For the age 0, the value of a_0 for the population of Bulgaria is little below 0.20. This figure may be less or more depending on two factors: the level of infant mortality and the speed of its decrease during the first months of life.

For older ages, the value of a_x should be above 0.5 because those who die in the age interval x to $x + 1$ are more likely to be at an age closer to $x + 1$ rather than x because of the rising risk through age.

The indicator ax has never being systematically used by NSI in calculating life tables or at least for verification purposes, although it may be found in the method implemented for the first five ages. This gives rise to another examination of the practice of the Institute, which reveals that the methodology might not have been followed in a consistent manner all the time. For example, the life table for 1995–1997 (National Statistical Institute, 1997) shows for females: $l_0 = 100{,}000$, $l_1 = 98{,}592$ and $L_0 = 98{,}602$. The value of L_0 appears to be so close to l_1 that:

$$a_0 = \frac{98{,}602 - 98{,}592}{100{,}000 - 98{,}592} = 0.007$$

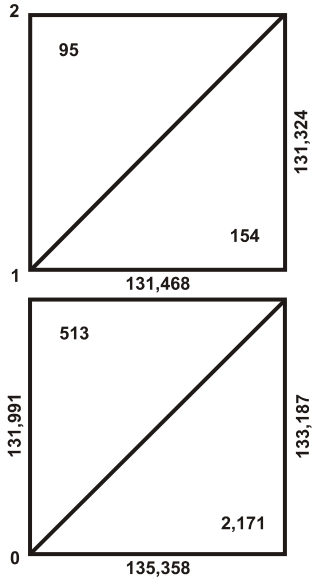
which is a very doubtful result. For males the situation is even worse: $l_0 = 100{,}000$, $l_1 = 98{,}229$ and $L_0 = 98{,}214$, which in combination is simply nonsense because L_0 as an average of l_x over the first twelve months of life (see 3) should be between l_0 and l_1.

[4] The latest figures are 68,53 for males and 75,23 for females from the 1999–2001 life table (National Statistical Institute, 2003).

The figures from the life table for 1999–2001 produce $a_0 = 0.076$ for males and $a_0 = 0.140$ for females. Both appear to be very low, although the latter might be acceptable. On the other hand, from the Lexis diagram of the deaths for this period (National Statistical Institute, 2003) one may directly obtain $a_0 = 0.143$ for males and $a_0 = 0.186$ for females as averages for the generations born in 1999 and 2000.

There is also a problem with the way the Institute performs the Lexis algorithm to project age groups through the years. Starting in 1979, one can see a mix up and incorrect accounting of survivors at end of the year. Fig. 2 illustrates the proper way this should be performed, without allowing for migration or simply assuming net migration to be zero. From the 1979 generation consisting of 135,358 life births, one should subtract the lower triangle 2,171 deaths thus obtaining 133,187 survivors at end of 1979. Instead, NSI wrongly subtracts both triangles for age 0 thus producing 132,674, giving the impression that 513 deaths belong to this generation and not to those actually born the previous year 1978.

Figure 2: Lexis chart, both sexes, 1979



A similar mix up could be found for the ages which followed. For example, from 131,991 alive from the 1978 generation at age 0 at the beginning of 1979, one should first subtract 513 deaths below the age of 1 (thus obtaining 131,468) and then 154 at age 1 to 2. The result should be 131,324. The NSI miscalculation is 131,991 minus 154 and minus 95 again ignoring the fact that these deaths belong to different generations (154 deaths occur among the 1978 generation, while 95 are from the generation born in1977).

Such miscalculations are distrubing because data for the first several ages could be easily verified with the help of vital statistics, their confirmation not being solely dependant on census information and these age groups are known to be less affected by migration.

Another example could be found among adult ages. At the beginning of 1979, there are 124,217 at age 50. During that year, 352 of them died before and 413 after reaching age 51. Therefore, there should be 123,452 at age 51 at the end of 1979, not counting external migration. According to NSI figures, however, this age group is 123,449, which could be derived only if the total number of deaths at age 51 to 52 during 1979 (which is 768) is subtracted from the initial 124,217 alive at age 50 at the beginning of 1979, again not including external migration.

Perhaps the differences come from NSI taking migration into account? Possible, but highly unlikely! Otherwise, the coincidence with the number of deaths is difficult to explain. Moreover, for the total population (all ages) there is no difference for 1979 – net migration is zero for males, zero for females, zero for both sexes! In addition, external migration data published by NSI for 1979 show no migrants at all.

Old age statistics and terminal age of the life table

Closely related to life expectancies, although in a technical sense only, is the terminal age of the life tables. For actuarial and insurance purposes this should be 100 or even higher. For demographic analysis, however, 80 or 85 is usually enough. The National Statistical Institute's typical choice of 100 is accurate enough to serve almost all objectives but this is strictly dependent on precise population registration and a high quality of statistical data for aged people.

Figure 3: Old age mortality, Bulgaria, 1961 –1991

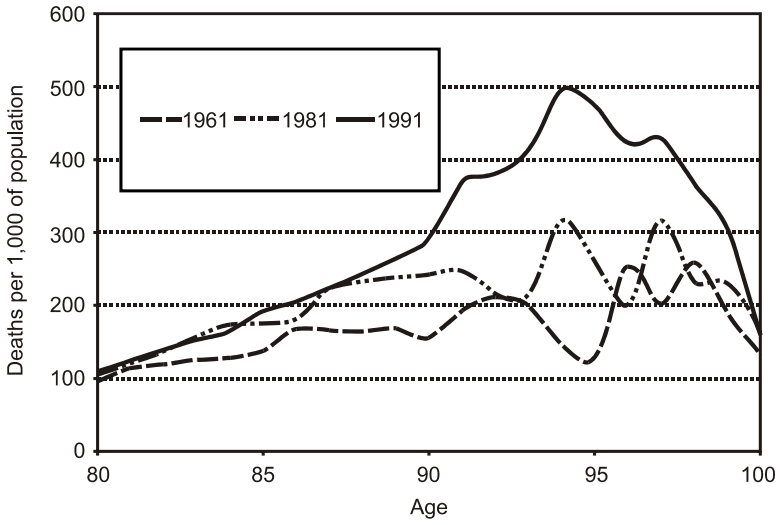


Fig. 3 shows the age pattern of mortality for old people in Bulgaria for 1961, 1981 and 1991. Up to age 80–85, the profile is relatively clear but after that the data show a huge variation. The 1991 curve reaches a maximum at about 95 and then shows a decrease in mortality. This shape is typical for the distribution of mortality for the very old in Bulgaria during the last 30 years and it is difficult to find any explanation for this fact. One possible source of confusion might be some lapses in registration during population censuses and completion of deaths certificates or simply defective enumeration. There may be some social and economic factors for the remarkable good health of very aged people in Bulgaria, or some interrelation between mortality at different ages which results in such a profile. This is a controversial matter and the interested reader might consult Coale and Demeny (1966), Preston (1970), Големанов, Христов (2001).

On the other hand, however, data about aged people contain mainly small numbers subject to large variation which makes direct statistical estimation a difficult task. In such a situation, one may not expect the terminal age of the life tables and its indicators to be derived with much precision.

The terminal age does not mean the last accessible age and that above it none could survive; it merely denotes the end of the table and suggests that events after that age are insignificant from a demographic point of view while all events below the terminal age should be described explicitly. The Bulgarian life table for 1984–1986 (National Statistical Institute, 1988) shows 129 males and 602 females survive beyond the age of 100. The figures show that each of these 129 males has an expected 0.80 years of future life. The corresponding expectancy for the 602 females is 0.84. The terminal age in the 1995–1997 life table decreased to 90 and the figures show 2,588 males and 7,009 females survive the age of 90. Males at that age have a remaining life expectancy of 3.01 years and females 3.83.

Determining the accuracy of these data is somewhat tricky. As mentioned earlier, common statistical data for very aged people might not be overly reliable for a precise estimation of the terminal age indicators. In addition, the corresponding data on mortality and resident population are based on open-ended age intervals and the triangle form like Fig. 1 is usually unavailable. This makes it very difficult to apply direct methods similar to the one already discussed around formulae (1) and (2).

One way to have an idea of how long those who survive the age of 100 live is to suppose that the age structure of the adult population is close to stationary. If so, the reciprocal of the age specific mortality rate should be close to e°_{100}. Table 1 shows the corresponding values for the population of Bulgaria for the period 1981–1990.

The relation between the age-specific mortality rate for the last open-ended terminal group and the expected length of life can be shown this way (Coale, Demeny, 1966: 13). The last age-specific mortality rate of the life table population is:

$$\frac{d_{100}}{L_{100}}$$

which actually equals:

$$\frac{l_{100}}{T_{100}} = \frac{1}{\mathring{e}_{100}}.$$

If the actual last age mortality rate is close to that of the life table population one may use the above formula to obtain a reasonable approximation for e°_{100}.

Table 1: Mortality rates and life expectancy approximation for age 100, Bulgaria

	Mortality rates			Reciprocal of mortality rates		
Year	Both sexes	Males	Females	Both sexes	Males	Females
1981	0.167724	0.192308	0.157527	5.962	5.200	6.348
1982	0.121193	0.153368	0.107207	8.251	6.520	9.328
1983	0.121719	0.137405	0.114924	8.216	7.278	8.701
1984	0.119919	0.118660	0.120462	8.339	8.427	8.301
1985	0.193466	0.218579	0.182154	5.169	4.575	5.490
1986	0.365875	0.342246	0.377457	2.733	2.922	2.649
1987	0.227106	0.206128	0.237381	4.403	4.851	4.213
1988	0.205825	0.222222	0.197674	4.858	4.500	5.059
1989	0.193246	0.150000	0.215297	5.175	6.667	4.645
1990	0.163907	0.149038	0.171717	6.101	6.710	5.824

Note: Calculations made using vital statistics from the demographic yearbook *Население* for the period 1981–1990.

It should be noted, however, that the correct evaluation of the last age indicators might be more of academic interest rather than practical importance because the terminal age group contributes very little to the life expectancy for other ages. Nevertheless, one may feel that the corresponding numbers of the Bulgarian life table for 1984–1986 look too low if compared with the values in the three right-most columns of Table 1. Perhaps the decrease of the terminal age to 90 for the life table 1995–1997 is, on reflection, a successful step in the right direction. For the 1999–2001 life table, however, the NSI reverted to 100 as a terminal age with $e^{\circ}_{100} = 0.48$ for both males and females.

How migration might affect the methodology

In real life, however, migration does exist and the above scheme needs modification to encompass possible drop outs and new entrants. If migration data are available in the same triangular form categorized by age, generation and year of exit/entrance (possibly by sex also) then this modification should be straightforward. However, migration calculation and statistics in Bulgaria are far from this perfect model.

But whatever the quality and relevance of migration data, population figures are expected to be precise. The triangle form of mortality data published by NSI in the 1960's together with the yearly information on life births help to easily calculate the balance at end of each year.[5] For example, there are 137,861 life births in Bulgaria in 1961. Of those

[5] Data from the demographic yearbook *Демографска статистика* for the period and the specialized publication *Население по пол и възраст на НР България – общо и по окръзи през периода 1965–1967 г.*

3,972 died before the end of the year and therefore 133,889 are expected to reach the beginning of 1962 being still at age 0. This number could be further projected into 1962 by first subtracting 1,172 deaths still at age 0 and then 344 at age 1, all deaths being from the same 1961 generation. This yields 132,373 which should be the size of the age group 1 to 2 at end of 1962, without counting migration.

Table 2: Projected and accounted population at the end of 1965

Generation	Age	Projected population	Accounted population	Difference
1965	0	122,878	130,200	-7,322
1964	1	126,526	122,800	3,726
1963	2	127,008	125,278	1,730
1962	3	128,449	127,814	635
1961	4	131,742	130,172	1,570
Total		636,603	636,264	339

Note: Calculations made using data from the demographic yearbook *Демографска статистика* for the period 1961–1965. Population accounts taken from the specialized publication *Население по пол и възраст на НР България – общо и по окръзи през периода 1965–1967 г.*

Table 3: External migration, all ages, Bulgaria 1961–1965

	Immigrants			Emigrants			Net migration		
Year	Both sexes	Males	Females	Both sexes	Males	Females	Both sexes	Males	Females
1961	17	15	2	84	38	46	-67	-23	-44
1962	16	14	2	100	47	53	-84	-33	-51
1963	7	7	0	119	58	61	-112	-51	-61
1964	20	19	1	84	38	46	-64	-19	-45
1965	3	3	0	42	24	18	-39	-21	-18
Total	63	58	5	429	205	224	-366	-147	-219

Source: Demographic yearbook *Демографска статистика* for 1961–1965

This calculation is applicable up to the end of 1965 in order to make comparisons with the accounted data derived from the 1965 census. A similar projection could be accomplished for the next four generations born in 1962, 1963, 1964 and 1965. The comparison with the 1965 end-of-year age distribution reveals huge differences as shown in Table 2.

It is hard to believe that within one single year, a single age group and a single generation (this is, the generation born in 1965) there could be 7,322 immigrants more than emigrants. For the generations 1964, 1963, 1962 and 1961 the differences are positive, denoting that there should be more emigrants than immigrants. For the period 1961–65, statistical data for Bulgaria (Table 3) show very low migration and do not provide a reasonable explanation for the differences illustrated in Table 2.

A larger scale of comparison is also possible. Firstly, let us look at what should be expected if migration data could be incorporated in calculations. Starting with the end-of-year 1960 population (males and females separately, as published by NSI in *Демографска статистка*, 1960), add natural increase and net migration for the coming years thus obtaining the population balance at end of each year from 1961–1965. The result could be compared with the figures of population accounts made by NSI. This is summarized in Table 4.

As the population accounts appear to be quite precise for the years 1961–1964 and it is relatively possible that 5,541 (the total accumulated discrepancy between population balance and accounts as shown in Table 4; last column, bottom figure) have not been covered by the census (temporarily absent or present and erroneously recorded, etc.) and hence never included in the resulting accounts for the end of 1965.

Table 4: Population growth, both sexes, Bulgaria 1960–1965

Year	Accounted population at the beginning of the year	Life births	Deaths	Net external migration	Population balance at the end of year	End of year accounted population	Difference between balance and accounts
1960					7,905,502		
1961	7,905,502	137,861	62,562	-67	7,980,734	7,980,734	0
1962	7,980,734	134,148	69,640	-84	8,045,158	8,045,158	0
1963	8,045,158	132,143	66,057	-112	8,111,132	8,111,132	0
1964	8,111,132	130,958	64,479	-64	8,177,547	8,177,547	0
1965	8,177,547	125,791	66,970	-39	8,236,329	8,230,788	5,541
Total		660,901	329,708	-366			

Source: Demographic yearbook *Демографска статистика* for 1961–1965 and the specialized publication *Население по пол и възраст на НР България – общо и по окръзи през периода 1965–1967 г.*

Now take the single age and sex distribution of the population of Bulgaria at end of 1960, using several next few issues of the demographic yearbook to discover the number of life births by sex and detailed information on deaths; then apply the Lexis algorithm and project the age/sex groups till the end of 1965. These time migration data are difficult to use because of the lack of detailed information by age and generation. Now compare the result with the corresponding data in the specialized publication of National Statistical Institute (1968). If there is a difference, it should come only from not taking migration into account. Table 5 shows some details of this population projection by age and sex starting from 1960, but not counting migration.

At this point, it is now possible to have a deeper look at Table 4 figures which could change the conclusions dramatically. Unfortunately, the NSI has published only summary data of population accounts for the years 1961, 1962, 1963 and 1964 so the only detailed comparison by age and sex could be with the data for the end of 1965 derived from the census.

The total 5,907 (Table 5, all ages, both sexes) totals 5,541 exactly (the discrepancy for 1965) plus the net migration 366 for 1961–1965 (Table 3 and Table 4); i.e. precisely what it would have been if there was no migration whatsover during this 5-year period. However, the sex distribution of this total is out of any proportion. Table 3 data proove that this can not be related to migration. There are also huge discrepancies between ages. It was already shown (Table 2) that the difference of 339 for the age group 0–4 is very unevenly distributed among single ages. Table 5 reveals that this is true for all ages.

Table 5: Projected minus accounted population, end of 1965

Age	Total	Males	Females
0–4	339	303	36
5–9	-2,324	-1,101	-1,223
10–14	-2,052	-1,879	-173
15–19	9,923	5,525	4,398
20–24	8,276	6,094	2,182
25–29	683	-30	713
30–34	-2,850	-3,759	909
35–39	-68	-1,288	1,220
40–44	-1,911	-1,542	-369
45–49	916	-301	1,217
50–54	527	-1,194	1,721
55–59	-3,247	360	-3,607
60–64	-4,598	-1,464	-3,134
65–69	-2,825	-3,069	244
70–79	489	389	100
80+	4,629	1,924	2,705
Total	5,907	-1,032	6,939

Note: Calculations made using data from the demographic yearbook *Демографска статистика* for 1961–1965. Population accounts from the specialized publication *Население по пол и възраст на НР България – общо и по окръзи през периода 1965–1967 г.*

All this demonstrates that the basis for life table calculations may be fragile if migration registration is loose and resident population accounts are not suffic reliable. Furthermore, including census data appears to be of little help and there are some doubts about the life tables produced by the Institute in the sixties.

Continuing the population projection by age, sex and generation into the 70's, 80's and 90's does not make much sense because of the massive emigration from Bulgaria of Turkish minorities particularly in the late 80's and significant accumulation of differences toward census dates which is difficult to apportion to single years. Instead, it might be more instructive to carry on Table 4 beyond 1965.

Table 4a: Population growth, both sexes, 1966–2001

Year	Accounted population at the beginning of the year	Life births	Deaths	Net external migration	Population balance at the end of year	End of year accounted population	Difference between balance and accounts
1966	8,230,788	123,038	68,366	-136	8,285,324	8,285,325	-1
1967	8,285,325	124,582	74,696	-85	8,335,126	8,335,126	0
1968	8,335,126	141,460	72,176	-330	8,404,080	8,404,080	0
1969	8,404,080	143,060	80,183	-2,693	8,464,264	8,464,264	0
1970	8,464,264	138,745	77,095	-11,031	8,514,883	8,514,883	0
1971	8,514,883	135,422	82,805	-9,594	8,557,906	8,557,906	0
1972	8,557,906	131,316	84,174	-10,555	8,594,493	8,594,493	0
1973	8,594,493	139,713	81,470	-5,296	8,647,440	8,647,440	0
1974	8,647,440	149,196	85,239	-1,348	8,710,049	8,710,049	0
1975	8,710,049	144,668	89,974	-346	8,764,397	8,731,400	32,997
1976	8,731,400	144,929	88,348	-2,252	8,785,729	8,785,763	-34
1977	8,785,763	141,702	94,362	-10,501	8,822,602	8,822,602	0
1978	8,822,602	136,442	92,445	-61,137	8,805,462	8,805,462	0
1979	8,805,462	135,358	94,403	0	8,846,417	8,846,417	0
1980	8,846,417	128,190	97,950	-5	8,876,652	8,876,652	0
1981	8,876,652	124,372	95,441	-2	8,905,581	8,905,581	0
1982	8,905,581	124,166	100,293	-122	8,929,332	8,929,332	0
1983	8,929,332	122,993	102,182	1	8,950,144	8,950,144	0
1984	8,950,144	122,303	101,419	186	8,971,214	8,971,214	0
1985	8,971,214	118,955	107,485	289	8,982,973	8,949,880	33,093
1986	8,949,880	120,078	104,039	543	8,966,462	8,966,462	0
1987	8,966,462	116,672	107,213	334	8,976,255	8,976,255	0
1988	8,976,255	117,440	107,385	326	8,986,636	8,986,636	0
1989	8,986,636	112,289	106,902	293	8,992,316	8,992,316	0
1990	8,992,316	105,180	108,608	277	8,989,165	8,989,165	0
1991	8,989,165	95,910	110,423	209	8,974,861	8,974,861	0
1992	8,974,861	89,134	107,998	209	8,956,206	8,484,863	471,343
1993	8,484,863	84,400	109,540		8,459,723	8,459,763	-40
1994	8,459,763	79,442	111,787		8,427,418	8,427,418	0
1995	8,427,418	71,967	114,670		8,384,715	8,384,715	0
1996	8,384,715	72,188	117,056		8,339,847	8,340,936	-1,089
1997	8,340,936	64,125	121,861		8,283,200	8,283,200	0
1998	8,283,200	65,361	118,190		8,230,371	8,230,371	0
1999	8,230,371	72,291	111,786		8,190,876	8,190,876	0
2000	8,190,876	73,679	115,087		8,149,468	8,149,468	0
2001	8,149,468	68,180	112,368		8,105,280	7,891,095	214,185

Source: Demographic yearbook *Демографска статистика* for 1966–1967 and *Население* for 1968–2001.

Table 4a shows that net external migration turned out to be so unpredictable that it was probably simply impossible to register external migration in a coherent way during 1966–2001. The huge figures for 1978 and the zero for 1979 look more like an adjustment in the statistical methodology to reflect past migration, while the positive numbers for 1983–1992 do not correspond to the massive emigration of Turkish minorities from Bulgaria in the 1980's. In addition, all figures for 1992 external migration precisely equal (by age and sex, immigrants and emigrants, etc.) those of 1991, which was so discouraging that NSI discounted publishing information on external migration.

The last four columns of Table 4 and 4a also reveal how NSI has attempted to rectify external migration situation with the help of census data (1965, 1975, 1985, 1992 and 2001). Obviously, a serious attempt has been made to compensate for the lack of reliable information on external migration between census years with different corrections and adjustments based on census data.

Unlike 1965, all differences for census years in the last column have quite normal sex distribution. Only for 1985 does this number contain almost three times more males than females, but this is not impossible. Following the repression of the Turkish minority in the early 80's, it is not unlikely that males first left the country before their families, a phenomenon which was perhaps caught by the census.

The figure for 1996 is difficult to explain but those for 1992 and 2001 show that Bulgaria lost about 700,000 of its population during the last two decades of the twenty century.

Concluding remarks

Population accounts in Bulgaria are far from perfect, although vital statistics are believed to be accurate. The lack of reliable migration registration and accounting seems to be the major source of incompatibilities and discrepancies in population information. However, the age groups should have been projected from year to year with engineering precision, whatever the quality and relevance of the data used. Census data obviously do not help too much to rectify the state of affairs. This is even more critical for adequate estimation of regional distributions and urban/rural structures. One may only wonder how age-specific urban/rural fertility rates look just prior and after census years. Both external and internal migration used to be significant over the past fifty years and sometimes the numbers are so big that age and sex specific fertility and mortality data look more like a rounding error. Ambiguous population accounts are a shaky basis for life table calculations and demographic analysis in general.

There are plenty of technicalities for the NSI to urgently improve. The correct Lexis method should be immediately adopted in population accounting and past data properly corrected and updated as often as possible. The population register needs a thorough inspection for records with invalid data. The census programs should be elaborated in such a way that helps in the monitoring and assessment of the functioning of the population register. Specialized sample surveys could be conducted for verification purposes of population accounting on one side, and census data on the other. All this requires close cooperation between the National Statistical Institute, the Ministry of Re-

gional Development and Public Works, border control offices and other institutions. In the era of computers and internet, fragmentation of information could no longer be tolerated. There may also be a need to update the legal framework for the statistical data and public information.

The concept of external migration statistics needs to be entirely revised. Border crossings should be registered by basic objective demographic indicators and with all personal information removed, the particulars should be directed to the NSI. Trying to categorize the crossings as migration or not is a complicated matter and should be left to dedicated sociological surveys. Furthermore, proper registration would certainly help in designing and performing useful secondary surveys to study migration propensity and behavior.

The National Statistical Institute could make a significant contribution to the process if more detailed and analytical information is published together with statistical programs, methodologies and procedures. The Internet provides unlimited possibilities for such an information exchange. The Institute will only benefit from a broader discussion of statistical facts and methods with the public and academic circles.

REFERENCES

COALE, A., DEMENY, P. (1966). *Regional model life tables and stable populations*. Princeton.

ЧОЛАКОВ, Н. (1975). "Итеративно построяване на таблици за смъртност", *Икономическа мисъл*, София, 9, pp. 52–60.

ГОЛЕМАНОВ, Н., ХРИСТОВ, Е. (2001). "Bulgaria: Mortality of the Elderly. (Trends and Perspectives)", *Workshop on Ageing, Intergenerational Solidarity, and Age-Specific Vulnerabilities*, Rome, April 20–21, 2001.

KEYFITZ, N. (1968). *Introduction to the Mathematics of Population*. Reading, Massachusetts: Addison-Wesley.

МАЛЕШКОВ, Х. (1983). "Изучаване на смъртността по причини", *Статистика*, София, 2, pp. 48–66.

PRESTON, S. (1970). "An International comparison of excessive adult mortality", *Population Studies*, vol. 24, no. 1, pp. 5–20.

PRESTON, S., KEYFITZ, N. and SCHOEN, R. (1972). *Causes of Death. Life Tables for National Population*. New York: Seminar Press.

РУСЕВ, Б. (1973). *Анализ на състава, естественото и механичното движение на населението*. Sofia: National Statistical Institute.

СУГАРЕВ, З. (1975). *Демографска статистика*. София: Наука и изкуство.

SOURCES

Демографска статистика, National Statistical Institute, Sofia, 1960–1967.

Население по пол и възраст на НР България – общо и по окръзи през периода 1965–1967 г., National Statistical Institute, Sofia, 1968.

Население, National Statistical Institute, Sofia, 1968–2001.

Таблици за смъртност и средна продължителност на населението на НР България за 1978–1980 г., National Statistical Institute, Sofia, 1983.

Таблици за смъртност и средна продължителност на населението на НР България за 1984–1986 г., National Statistical Institute, Sofia, 1988.

Население и демографски процеси 2001, National Statistical Institute, Sofia, 2003.

Nikola Čolakov

MIGRACIJSKA STATISTIKA, UKUPAN BROJ STANOVNIKA I TABLICE OČEKIVANOG TRAJANJA ŽIVOTA – METODOLOGIJA KOJA SE RABI U BUGARSKOJ

SAŽETAK

Smatra se da je vitalna statistika u Bugarskoj relativno detaljna, potpuna i precizna. Stručnjaci u Nacionalnom statističkom institutu upotrebljavaju te podatke zajedno s izravnijim metodama za izradu tablica očekivanog trajanja života prema pojedinačnim godinama starosti, posebno za muškarce i žene. Međutim, vanjske migracije ipak mogu predstavljati konceptualni problem u primjeni tih metoda, dok sofisticiranije metode mogu biti neprimjenjive jer su podaci o migracijama najčešće oskudni i vrlo nepouzdani. Migracija može utjecati na izradu tablica očekivanog trajanja života u dva oblika. Prvo, migracija utječe na mortalitet i mora se uključiti pažljivo, osobito godine s velikim brojem emigranata što je tipično za Bugarsku. Drugo, statističko praćenje migracije ili bilo kakvi drugi nedostaci ili iskrivljavanje podataka o migraciji mogu predstavljati ozbiljan problem pri uporabi ukupnog broja stanovnika u izračunu tablica očekivanog trajanja života i za statističku analizu općenito. Krajem šezdesetih godina prošlog stoljeća postojao je značajan odljev migranata u Tursku koji na statističke podatke nikada nije utjecao tako izravno kao na mortalitet. Tajna i nezakonita emigracija iz Bugarske nakon Drugoga svjetskog rata i masovan odlazak Turaka tijekom osamdesetih također su mogli biti činilac pogrešne procjene veličina dobnih skupina između popisa stanovništva. Otuda su proizašle pogrešne procjene razine mortaliteta. Kako je tadašnja politička vlast nerado otkrivala te činjenice javnosti, statističarima je bilo jako teško doći do relevantnih podataka i izračunati prosječno trajanje života na konzistentan način ili izvršiti pouzdanu statističku analizu. U članku se propituje metodologija Instituta i ukazuje na neka srodna pitanja proistekla iz praćenja službene statistike u Bugarskoj na tom polju.

KLJUČNE RIJEČI: vitalna statistika, migracija, prirodni prirast, prosječno trajanje života, stanovništvo, saldo stanovništva, projekcije stanovništva

Nikola Cholakov

STATISTIQUE MIGRATOIRE, NOMBRE TOTAL D'HABITANTS ET TABLES D'ESPÉRANCE DE VIE : LA MÉTHODOLOGIE UTILISÉE EN BULGARIE

RÉSUMÉ

On considère que la statistique vitale en Bulgarie est relativement détaillée, complète et précise. Les spécialistes de l'Institut national de statistique utilisent ces données conjointement à des méthodes plus directes de réalisation des tables d'espérance de vie selon les tranches d'âge, respectivement pour les hommes et les femmes. Toutefois, les migrations extérieures peuvent représenter un problème conceptuel dans l'application de ces méthodes, tandis que les méthodes plus sophistiquées peuvent être inutilisables car les données sur les migrations sont le plus souvent avares en détails et très peu fiables. La migration peut influer sur la réalisation des tables d'espérance de vie sous deux formes. Dans le premier cas, la migration influe sur la mortalité et doit être soigneusement prise en compte (englobée), particulièrement pour les années présentant un grand nombre d'immigrants, ce qui est le typique de la Bulgarie. Dans le second cas, le suivi statistique de la migration ou quelque autre déficience ou déformation des données sur la migration peuvent représenter un sérieux problème lors de l'utilisation des chiffres exprimant le nombre total d'habitants dans le calcul des tables d'espérance de vie et pour l'analyse statistique en général. La fin des années 60 du siècle dernier fut marquée par un notable courant de migration vers la Turquie, qui n'a jamais été reporté sur les données statistiques de façon aussi nette que la mortalité. L'émigration clandestine et illégale hors de Bulgarie après la Seconde guerre mondiale ainsi que le départ massif des Turcs au cours des années 80 auraient aussi pu être un facteur d'erreur d'évaluation des groupes d'âge entre deux recensements. D'où des évaluations erronées du taux de mortalité. Le pouvoir politique de l'époque ne désirant pas rendre publics ces faits, les statisticiens avaient des dificultés pour aboutir à des données fiables, calculer la durée de vie moyenne d'une façon cohérente ou réaliser une analyse statistique fiable. L'auteur de cet article scrute la méthodologie de l'Institut et met le doigt sur certaines questions relevant du suivi de la statistique officielle en Bulgarie dans ce domaine.

MOTS CLÉS : statistique vitale, migration, croissance naturelle, évolution moyenne de la vie, population, solde démographique, projections démographiques