ABOUT MORTALITY DATA FOR BULGARIA

by Dimiter Philipov
Revised by Domantas Jasilionis, Trifon I. Missov & Sergey Timonin
Last updated: 06 September 2022

GENERAL

The National Statistical Institute (NSI) is the governmental unit that provides statistical information. It was established in 1900 as the Central Statistical Office.

Demographic vital statistics are available in published form since 1881, but data from 1900 onwards have greater reliability than those for prior years. The first demographic yearbook was issued in 1960.

Data exist in both published and electronic formats. The latter typically provides more detailed information.

The first population census was conducted in 1880. Reliability improved with the census carried out in 1900. Population censuses are usually conducted every ten years. Those after the Second World War were carried out in 1946, 1956, 1965, 1975, 1985, 1992, 2001, and 2011.

Source of Data

Mortality data used in the Human Mortality Database (HMD) come from unpublished electronic sources provided by the National Statistical Institute (Sofia, Bulgaria), as well as from NSI's Infostat database (https://infostat.nsi.bg).

Specific Episodes in the Demographic History of Bulgaria

Since 1947 there have been two large emigration waves to Turkey. The first one took place in 1950 and 1951 when about 100,000 and 55,000 people, respectively, left the country. The second wave happened in 1989 and 1990. Just before and after the collapse of Communism in November 1989, there was massive emigration of an ethnic Turkish minority out of Bulgaria to Turkey following the mass protests against the Bulgarization campaign set by the Communist party in 1984-1985. Around 360 000 Bulgarian Turks were estimated to have left Bulgaria for Turkey starting from January until the border closure by Turkey in August 1989. However, about one-third of these Turkish emigrants returned to Bulgaria after the fall of the communist government and the lifting of the requirement to change one's Turkish name into a Bulgarian name in November 1989. The second massive emigration wave including ethnic Bulgarians and Turks occurred in 1990-1991 following the deep economic crisis related to the beginning of the transition to a market economy. Apart from these two waves, emigration rates until the beginning of the 1990s have been negligible.
Since then, events that most probably initiated large out-migration waves are the change of Bulgaria's political system (1989-1991), the Bulgarian banking and currency crisis (1996-1997), the opportunity of the visa-free entrance to the EU for Bulgarian

nationals (2001), and Bulgaria's entry into the EU (2007). It has been estimated that about 196,000 people left and 19,000 entered the country during the 1990s.

The 2000s were characterized by massive out-migration after Bulgaria joined the EU in 2007 when the NSI started providing out-migration numbers. Following the 2011 census, it has been estimated that from 2001 to 2011 Bulgaria recorded 160,897 out-migrants and only 14,347 registered in-migrants between 2007 and 2011. However, the official yearly out-migration numbers in 2007-2011 are questionable as the NSI was able to count only those who declared their departure to the Bulgarian authorities. The NSI reported a positive net migration flow during the COVID-19 pandemic of 43,421 people in 2020-21.

Birth rates were high towards the end of the 1940s, during the post-war compensation period with a total fertility rate of 2.8 in 1947 (Human Fertility Database). Significant fluctuations were observed with the introduction of a population policy in 1967, which contributed to a temporary rise in the number of births. The total birth count increased from a low of about 123,000 in 1966 to a high of 143,000 in 1969. A similar temporary rise was observed at the beginning of the 1970s with a total birth count of nearly 150,000 in 1974. It has regularly declined afterwards to reach a low of 64,000 first in 1997, and less than 59,000 in 2021, after fluctuating between 59,000 and 80,000 in the interim.

TERRITORIAL COVERAGE

There were no territorial changes in Bulgaria during the period covered by the HMD (1947-2021). However, numerous such changes took place prior to 1940.

DEATH COUNT DATA

Coverage and completeness

National death statistics encompass all deaths that occurred in the resident population. Until 2009, the concept of infant death (adopted in 1970) differed from that used by the World Health Organization (WHO). The definition used by Estonia National Statistical Institute (NSI) was more restrictive than the WHO definition. For NSI, newborns weighing less than 1,000 grams and surviving for less than six days were registered as spontaneous abortions (i.e. miscarriages) rather than as infant deaths. The latter definition leads to the underestimation of the infant mortality rate (IMR) compared to that based on the WHO concept (Kingkade, Sawyer, 2001; Aleshina, Redmond, 2003; Gantcheva, Kolev, 2001). From 2009 onwards, Bulgaria has been using a definition which is close to the one recommended by WHO (for more details see the section on "Birth count data").

Specific details

One major concern about the reliability of death statistics relates to the underestimation of infant mortality. According to Aleshina and Redmond (2003), Bulgaria was the only European country outside of the Commonwealth of Independent States in which the more restrictive "Soviet" definition of infant mortality persisted until 2008. Relying on indirect estimation techniques (e.g. comparing ratios of post-neonatal and neonatal mortality rates), Kingkade and Sawyer proposed an adjustment factor of around 1.5 (depending on the period) for the official counts of infant deaths.

A second issue relates to the quality of mortality data at older ages. Like in Russia and Lithuania (see the corresponding *Background and Documentation* files), the Bulgarian data exhibits some age heaping at 60, 70, 80, and 90 years between the end of the 1940s and the end of the 1950s (more details are provided below, in the section on "Data quality issues"). Furthermore, life expectancy at age 80 years appears to be suspiciously high in the 1950s and 1960s compared for instance to Sweden and might reflect significant age overstatement in the population.

Besides the aforementioned problems, there are other discrepancies in the data on deaths for particular years. For example, in 1986, there is an unexplained drop in the number of deaths within the 1900 cohort so that the mortality rate among 86-year-old people is substantially lower in this particular year compared to the surrounding years.

There are also problems with mortality estimates for 1994. Unfortunately, in this case, our method, which entails splitting death counts in the open age interval 80+ into Lexis triangles, yields implausible results. After applying this method, mortality rates at ages 80-89 years exhibit a sudden drop, while mortality for all ages above 90 years increases unexpectedly (Appendix 2, Figures 2.C and 2.D).

POPULATION COUNT DATA

Coverage and completeness

Until 1964, NSI reports the actual population numbers. The actual population consists of the *de facto* population, to whom are added residents who are temporarily abroad. For years since 1965, the NSI reports the resident population counts. The numbers refer to the end-of-the-year population. Data by single years of age are available in electronic form.

The population census counts the actual and the resident population. Relevant publications distinguish between the resident population, the *de facto* population, and the population actually present.

Specific details

Official population estimates are provided by age, ending with an open age interval at the highest ages. Until 1994, the open age interval was 100 and older. In 1995, it was 90 and older, whereas in 1996 and 1997 it was 80 and older. For 1998 and 1999, the data were provided with an even broader open age interval (75+ and 70+, respectively).

The latter aggregations may indicate errors in population estimates at older ages. Furthermore, all the above-mentioned data for the 1990s are post-censal population estimates, which have not been revised according to the census of 2001. Official population estimates for the end of the 1980s and 1990s show significant fluctuations and discrepancies due to unregistered emigration during the period. A notable discrepancy between the official post-censal estimates for December 31st, 2010 and the 2011 census counts (as of February 1st, 2011) results from significant unregistered emigration during the 2000s. This difference represents about 140,000 people.

Taking into account these data quality problems, we decided not to use the official estimates for the period 1989-2011 for further calculations of the mortality surface. Official population estimates are thus used for the period 1947-1988 only and three series of inter-censal estimates are computed following the HMD methodology for the period 1989-2010. The first set of estimates refers to the period from 1988 (a pseudo-census point) to 1992 and the second set of estimates covers the period from 1992 to 2000. For more details about this procedure, see below the section about "Data quality issues." The third set of estimates for 2002-2011 was derived using two pseudo-census points (December 31 population estimates for 2001 and 2011) instead of the 2001 and 2011 census counts. Population counts from 2011 onwards are the official NSI post-censal (December 31st) estimates.

BIRTH COUNT DATA

Coverage and Completeness

Births refer to the resident population. Until 2009 the definitions of a live birth, a stillbirth, and a spontaneous abortion notably differed from the internationally accepted concept. Since 2009, these definitions have become close to the WHO definition.

Before 2009:
Live births included all births of newborns weighing at least 1000 grams and exhibiting any sign of life, irrespective of the duration of the pregnancy. Newborns weighing less than 1000 grams were counted as live births only if they survived for at least six full days, otherwise they were registered as spontaneous abortions (if gestation lasted for less than 28 weeks or their height was no more than 35 centimeters) or as stillbirths (if gestation lasted for at least 28 weeks or their height was over 35 centimeters).

Since 2009:
Live births include all births of newborns showing any evidence of life irrespective of the duration of the pregnancy. If the weight of a newborn is less than 600 grams and/or the duration of the pregnancy is less than 22 weeks, these births are counted as live births only if the newborn survived for at least three days after delivery. Newborns weighing at least 600 grams, showing no sign of life, and born after at least 22 weeks of gestation are registered as stillbirths. Newborns weighing less than 600 grams, showing no sign of life and born before completing 22 weeks of gestation are recorded as spontaneous

abortions. Spontaneous abortions also include newborns weighing less than 600 grams and surviving less than three days after delivery.

DATA QUALITY ISSUES

Problems with the official population estimates for the period 1986-2010

Bulgarian data on death and population counts cover the period from 1947 to 2020. However, for the estimation of mortality surfaces in the HMD, the official annual population estimates were used only for 1947-1988 and 2011-2020.

The official population estimates for Bulgaria show a sudden drop in the total population count for census years 1985, 1992, and 2001 (Figure 1). For example, according to the official data, the total number of males decreased by about 252,000 between 1991 and 1992 (census year). A similar notable discrepancy is found when comparing the official post-censal estimates for December 31st, 2010 to the 2011 census counts (as of February 1st, 2011).

Discontinuities in the population trends are due to several factors. First, it seems that population estimates have not been recalculated backwards based on the corresponding latest censuses. Thus, the available data consist of post-censal estimates rather than inter-censal population estimates. Second, population estimates were greatly affected by unregistered emigration (especially during the period from 1989 to 2010) not accounted for in the official statistics. The NSI has been reporting migration numbers only since 2007. They do not seem to be realistic at least for 2007 and 2008: in 2007 NSI reports official out-migration of 2,958 whereas Table 1 (containing migration of Bulgarians to selected receiving countries) suggests there are at least 60,406 people who left the country; in 2008 the numbers are respectively 2,112 (NSI) and 52,025 (from selected foreign statistical offices) (Table 1). Eurostat, which cross checks migration statistics among European states, has published NSI's numbers (all marked as "provisional") only for 2007 and from 2012 onwards.

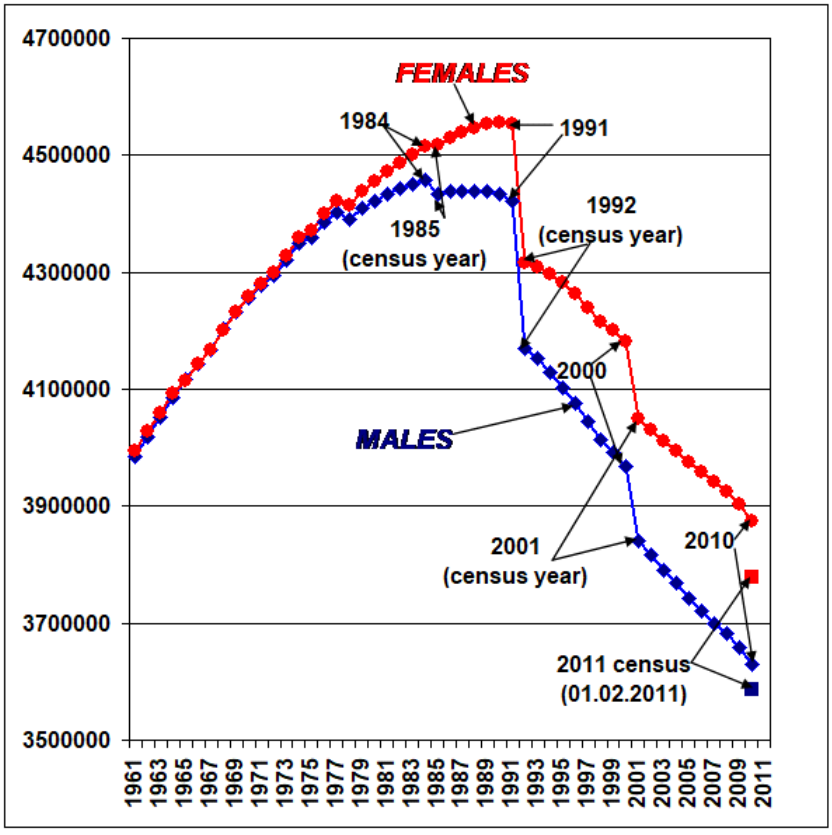
For 1992-2011, as we do not have additional data sources on Bulgarian migration, we calculated our own inter-censal population estimates, thus eliminating the discontinuities observed between 2000 and 2001 and between 2010 and 2011. However, this standard solution does not provide plausible results for the years from 1985 to 1992 because of an irregular pattern of out-migration. The inter-censal survival method described in the *Methods Protocol* of the HMD assumes that migration is evenly distributed between two census points, which thus substantially overestimates Bulgaria's out-migration in the 1985-1988 period.

After analyzing additional data (including indirect migration estimates), we decided to split the 1985-1992 inter-censal period into two sub-periods: 1985-1988 (a period of stable and negligible migration) and 1989-1992 (a period of substantial out-migration). The official population estimates were used for the former period, but new population estimates were calculated for the latter period. Specifically, the year 1988 was treated

as a “pseudo-census point” for the start of the interval and the 1992 census counts were used to end the interval. The standard inter-censal survival method described in the HMD *Methods Protocol* was then used to derive January 1st population estimates for years 1989 to 1992.

Population estimates for the years 2002-2010 should be considered with caution as they result from a uniform redistribution of the number of individuals needed to match the 2001 and 2011 census figures. Most probably, such a redistribution does not reflect the actual emigration pattern that might have exhibited peaks in 2007 and perhaps 2008, the first two years after Bulgaria joined the European Union. Such evidence emerges, for example, from the official statistics on registered immigrants in Germany, the UK, and the Netherlands though the numbers do not match the out-migration to these countries reported by NSI. As a result, there is neither consensual nor reliable year-by-year migration statistics for the years 2001-2011. In this period, Eurostat reports the NSI out-migration figure only for 2007. The number 2,958 underestimates by far the real emigration flow implied by statistics from other European countries on the number of foreign born from Bulgaria. For instance, Germany alone reports 24,093 newly registered Bulgarian citizens in this one year (see Table 1).

Figure 1. Trends in the official population by sex, Bulgaria, 1961-2011



Underreporting of infant deaths

Before 2009, under-reporting of infant mortality was the most serious problem related to the quality of mortality statistics in Bulgaria. Kingkade and Sawyer (2001) suggested that an unexpectedly high ratio of post neonatal to neonatal probabilities of dying indicated some under-reporting of infant deaths. The ratio was found to be about 0.85 for Bulgaria, while the corresponding figure was only 0.43 for Western Europe and about 0.50 for countries in Central Europe and for the Baltic States (Aleshina, Redmond, 2003). Taking into account the discrepancies in the reported count of infant deaths, Kingkade and Sawyer (2001) estimated that the infant mortality rate based on the internationally accepted definition was about 57% higher than that reported by the Bulgaria NSI. However, it is difficult to prove whether such a substantial portion of infant deaths is really being excluded (Aleshina, Redmond, 2003).

Age heaping and overstatement in deaths

Age heaping at older ages is often considered as one of the most serious problems with mortality statistics in countries of the former USSR (Anderson & Silver, 1997). For example, Zakharov (2002) has showed that age heaping is very pronounced at 70, 80 and 90 years in the Russian data for the period before 1970 (see the Russian *Background and Documentation* file in the HMD). Using the same procedures, we found possible age heaping for the years before 1960. Figure 2 (Appendix 2) shows that mortality at ages 60, 70, 80, and, especially, at age 90 is often higher than at ages 61, 71, 81, and 91, respectively.

In addition, Figure 3 (Appendix 2) demonstrates that life expectancy at age 80 is probably overstated for both males and females until the end of the 1960s. For example, the estimates suggest that life expectancy for Bulgarians exceeded that of their Swedish counterparts by more than one year in 1958. These implausible values can be attributed to age exaggeration in both deaths and population estimates.

Given the aforementioned data quality problems, the data for Bulgaria should be used with caution for the period before 1970 and, regarding the population numbers, for the years 2002-2010.

Table 1: Number of registered immigrants from Bulgaria by the official authorities in the receiving countries (2001-2015)

	GER	IRE	ESP	ITA	NLD	SWE	UK	NOR	SWI	AUT	TOTAL
2001	13156		11771		325	119	481			891	**26743**
2002	13191		15872		433	149	2369	121		1472	**33607**
2003	13369				473	152	3784	100	370	1728	**19976**
2004	11586		20997		402	123	1392	103	364	1677	**36644**
2005	9057		18377		415	100		114	368	1388	**29819**
2006	7749	98	21748		450	123		86	369	1214	**31837**
2007	20919	92	31331		4908	1159		193	336	1468	**60406**
2008	24093	105	10771	8436	5231	962		393	372	1662	**52025**
2009		160	7686	6244	3097	767		366	536		**18856**

2010		144	8389	5877	3059	629		623	621		**19342**
2011		156	9893	5101	3326	600		697	654	2082	**22509**
2012		156	6224	4794	2946	764		945	711	2246	**18786**
2013		144	4947	3734	2662	862		1171	767		**14287**
2014		182	4699	2932	3541	934	9153	949	703	3847	**26940**
2015		93	4921	2794	3534	877		879	671	3830	**17599**

Source: Eurostat, 2019.

Table 2: Reported in- and out-migration by NSI (2007-2021) and comparison to total out-migration according to Table 1

	In-migration (NSI)	**Out-migration (NSI)**	**Out-migration (Table 1)**
2007	1561	2958*	60406
2008	1236	2112	52025
2009	3310	19039	18856
2010	3518	27708	19342
2011	4722	9517	22509
2012	14103	16615*	18786
2013	18570	19678*	14287
2014	26615	28727*	26940
2015	25223	29470*	17599
2016	21241	30570*	
2017	25597	31586*	
2018	29559	33225	
2019	37929	39941	
2020	37364	6649	
2021	39461	26755	

*Numbers published by Eurostat

Revision NOTES

Changes with the December 2017 revision:

Life tables: All life tables have been recalculated using a modified methods protocol. The revised protocol (Version 6) includes two changes: 1) a more precise way to calculate a0, the mean age at death for children dying during the first year of life and 2) the use of birth-by-month data (where and when available) to more accurately estimate population exposures. These changes have been implemented simultaneously for ALL HMD series/countries. For more details about these changes, see the revised Methods Protocol (at http://v6.mortality.org/Public/Docs/MethodsProtocol.pdf), particularly section 7.1 on Period life tables and section 6 and Appendix E, on death rates. The life tables calculated under the prior methods (Version 5) remain available at v5.mortality.org but will not be further updated in the future.

Changes with the September 2019 revision:
Population: New inter-censal population estimates for 2002-2010 have been calculated according to the HMD methodology. New post-censal population estimates (based on the 2011 population census) for 2011-2017 added to the series.
Deaths: Death counts for 2011-2017 have been added to the series.
Live births: Annual births by sex have been extended to the year 2017.
Monthly live births: A new input file containing the total numbers of live births by month for 1900, 1905, 1910, 1920, 1925, 1930, 1935, and 1939-2017 has been added to the input database.

ACKNOWLEDGEMENTS

We would like to thank Yordan Kalchev, Krasimira Krusteva, Magdalena Kostova, Detelina Popcheva and their colleagues at the National Institute of Statistics for their help and cooperation in obtaining the Bulgarian data. We also thank Dora Kostova for assistance in getting these data.

REFERENCES

Aleshina, N., Redmond, G. (2003). *How high is mortality in Central and Eastern Europe and the CIS?* Florence, Italy: *UNICEF Innocenti Working Papers*, No.95, 44p.

Anderson, B.A., Silver, B.D. (1997). Issues of Data Quality in Assessing Mortality Trends and Levels in the New Independent States. Pp.120-154 in: *Premature Death in the New Independent States*. Washington, DC, USA: National Academy Press.

Gantcheva, R., Kolev, A. (2001). *Children in Bulgaria: growing impoverishment and unequal opportunities.* Florence, Italy: *UNICEF Innocenti Working Papers*, No. 84, 69p.

Kingkade, W., Sawyer, C. (2001*). Infant mortality in Eastern Europe and Former Soviet Union before and after the breakup*. Paper presented at the IUSSP conference. Salvador de Bahia, Brazil: 23p.

NSI (National Statistical Institute) (2012). *Population - Methodological Annotations*. Retrieved 13 April 2012. (http://www.nsi.bg/otrasal-metodologiaen.php?otr=53).

IOM (2003). *Bulgaria: The social impact of seasonal migration*. Vienna, Austria: *Migration Trends in Selected Applicant Countries, Vol. 1,* 106 p.

Zakharov (2002). Some data problems revealed (research note). in: *About mortality data for Russia*. Retrieved 08 November 2004 (http://www.mortality.org).

APPENDIX 1

DESCRIPTION OF DATA USED FOR LEXIS DATABASE

DEATHS

Period	Type of Data	Age grouping	Comments	RefCode(s)
1947-1992	Annual number of deaths in the resident population, by sex, year of birth, and single year of age (Lexis triangles)	0, 1, 2, …100+, unknown		1
1993	Annual number of deaths in the resident population, by sex and single year of age (1x1 squares)	0, 1, 2, …100+		1
1994	Annual number of deaths in the resident population, by sex and single year of age (1x1 squares)	0, 1, 2, …80+		1
1995	Annual number of deaths in the resident population, by sex, year of birth, and single year of age (Lexis triangles)	0, 1, 2, …100+, unknown		1
1996	Annual number of deaths in the resident population, by sex and single year of age (1x1 squares)	0, 1, 2, …100+		1
1997-1998	Annual number of deaths in the resident population, by sex, year of birth, and single year of age (Lexis triangles)	0, 1, 2, …100+, unknown		1, 8

DEATHS (CONTINUED)

Period	Type of Data	Age grouping	Comments	RefCode(s)
1999	Annual number of deaths in the resident population, by sex, year of birth, and single year of age (Lexis triangles)	Triangles: 0, 1, 2, …89. 1x1: 90, 91, …, 100+		8
2000-2001	Annual number of deaths in the resident population, by sex, year of birth, and single year of age (Lexis triangles)	0, 1, 2, …100+, unknown		8
2002-2003	Annual number of deaths in the resident population, by sex, year of birth, and single year of age (Lexis triangles)	Triangles: 0, 1, 2, …100. Period-cohort: 100, 101, maxi-mum age attained.		9
2004-2021	Annual number of deaths in the resident population, by sex, year of birth, and single year of age (Lexis triangles)	Triangles: 0, 1, 2, …89. 1x1: 90, 91, …, 100+		11, 17

POPULATION

Period	Type of Data	Age grouping	Comments	RefCode(s)
1946	Census counts of population by sex and single year of age as of December 31st. Actually present (de facto)	0, 1, 2, 3, …,90+		2

	population.			
1956	Census counts of population by sex and single year of age as of December 1st. Actually present (de facto) population.	0, 1, 2, 3, …,90+		2
1960-1964	Annual population estimates by sex and single year of age as of December 31st. Actually present (de facto) population.	0, 1, 2, 3, …,100+	Post-censal estimates	3
1965-1988, 1992	Annual December 31st resident population by single year of age and sex	0, 1, 2, 3, …,100+	Post-censal estimates	3
2001, 2011	Annual December 31st resident population by sex	0, 1, 2, 3, …,100+	Post-censal estimates	7, 10
2012-2021	Annual December 31st resident population by sex	0, 1, 2, 3, …,100+	Post-censal estimates	10, 18

BIRTHS BY SEX

Period	Type of Data	Age grouping	Comments	RefCode(s)
1947-2021	Annual live birth counts by sex in the resident population.			4, 5, 6, 12, 15

BIRTHS BY MONTH

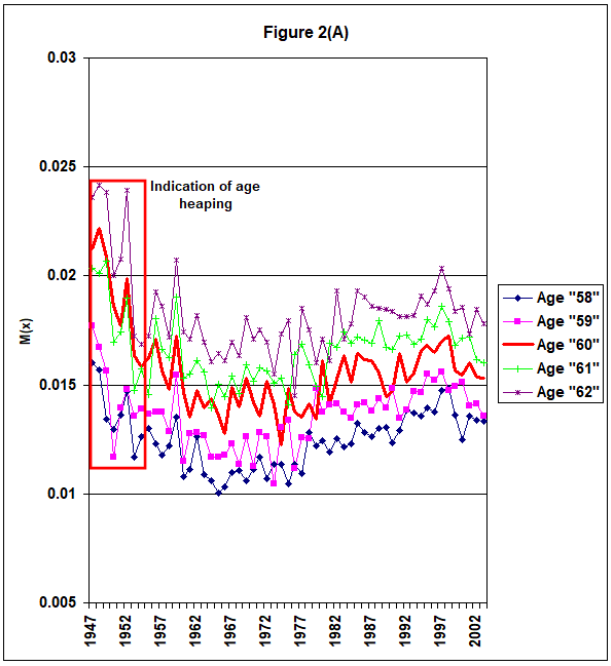
Type of data: Quinquennial and annual live birth counts by month.

Period covered: 1900, 1905, 1910, 1920, 1925, 1930, 1935, 1939-2021.

RefCode(s): 13, 14, 16.

APPENDIX 2

Figure 2. Mortality rates for selected ages, Bulgaria, total population, 1947-2003



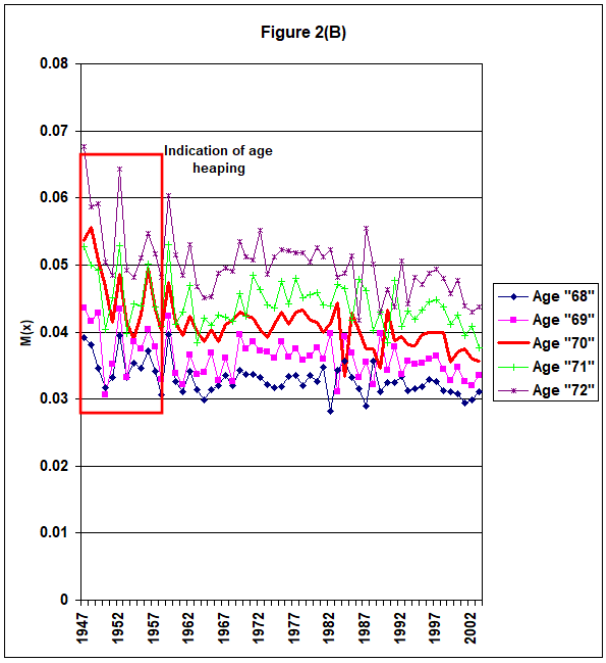
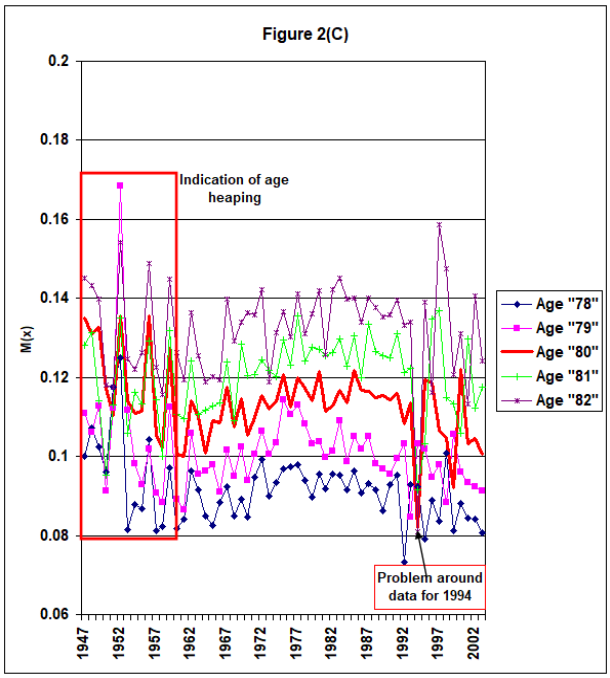


Figure 2 (continued). Mortality rates for selected ages, Bulgaria, total population, 1947-2003



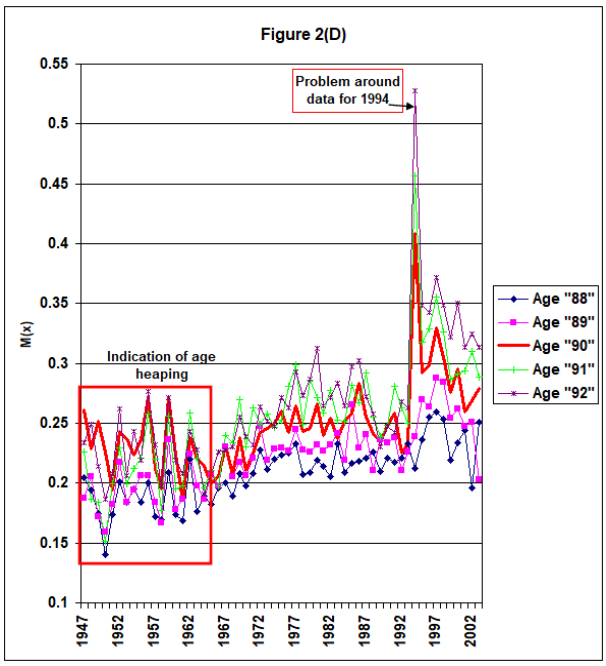
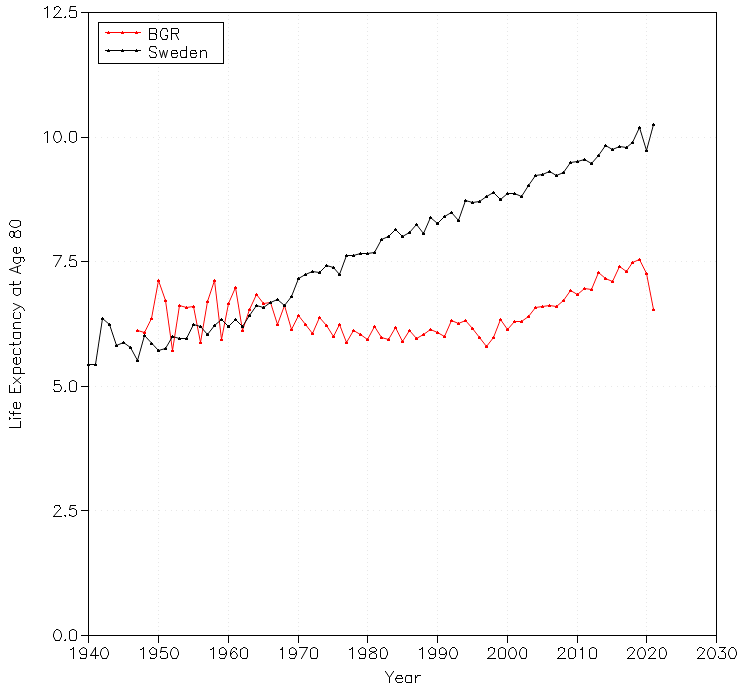


Figure 3. Life expectancy at age 80 in Bulgaria and Sweden, 1947-2021

External Plausibility: Life Expectancy at Age 80, Females, BGR



External Plausibility: Life Expectancy at Age 80, Males, BGR