Max-Planck-Institut für demografische Forschung
Max Planck Institute for Demographic Research
Konrad-Zuse-Strasse 1 · D-18057 Rostock · GERMANY
Tel +49 (0) 3 81 20 81 - 0; Fax +49 (0) 3 81 20 81 - 202;
http://www.demogr.mpg.de

MPIDR WORKING PAPER WP 2005-010
APRIL 2005

Official population statistics and the Human Mortality Database estimates of populations aged 80+ in Germany and nine other European countries

Dmitri Jdanov (jdanov@demogr.mpg.de)
Rembrandt Scholz (scholz@demogr.mpg.de)
Vladimir M. Shkolnikov (shkolnikov@demogr.mpg.de)

TO DR's "SC4" (special collection suggested by V.Shkolnikov and J.Wilmoth)

Official population statistics and the Human Mortality Database estimates of populations aged 80+ in Germany and nine other European countries

D. Jdanov[1]*, R. Scholz*[1]*, and V. M. Shkolnikov*[1]

[1]***Max Planck Institute for Demographic Research, Rostock, Germany***

Abstract

A systematic comparison of the Human Mortality Database and official estimates of populations aged 80+ is presented. We consider statistical series for East and West Germany and also series for Denmark, England and Wales, France, Finland, Hungary, the Netherlands, Russia, Sweden, and Switzerland. The Human Mortality Database (HMD, www.mortality.org) methodology relies on the methods of extinct and almost extinct generations. HMD estimates are precise if the quality of death data is high and the migration among the elderly is negligible. The comparisons between the HMD and the official populations are not fully appropriate for the 1990s since the HMD calculations are related to official population estimates. A significant overestimation of the male population aged 80+ and especially 90+ between the censuses of 1970 and 1987 was found in West Germany. The relative surplus of men aged 90+ increased from 5 to 20 percent, which expressed in absolute numbers indicates an increase from 2 to 10 thousand. In 1971-1987 the official death rates have fallen dramatically to implausibly low values. In 1987-88 death rates based on the official populations suddenly jumped to the HMD death rates due to the census re-estimation. In the 1990s an accelerated decrease in male death rates has resumed. For other countries, the relative and absolute deviations from the HMD estimates were especially high in Russia, Hungary, and England and Wales. Regression analysis reveals common factors of the relative deviation from the HMD populations. The deviation tends to decrease with time, increase with age, be higher during inter-census periods than in census years, and to decrease after the introduction of population registers.

Key words: aging; elderly; population estimates; quality of statistics.

Introduction

Since the 1950s, population aging has become a central demographic trend, leading to numerous and important socio-economic consequences. Accurate estimates of old-age populations are necessary for correctly assessing the current scale of the process and predicting future changes. People at advanced ages over 80 frequently need serious assistance from society. The spectacular growth of the absolute number of people aged 80+ suggests that at present the same relative error in a population estimate can result in a greater absolute error than 20 or 30 years ago (see Figure 1).

Official population estimates are especially problematic at advanced ages. This is related to a general tendency to overstate ages, sometimes related to difficulties with an imprecise character of information about the date of birth. Registration of births became compulsory in 1872 in Japan, in 1874 in England and Wales, in 1876 in Germany, and in 1911 in Portugal (Kannisto, 1988). In addition, the population estimates can be affected by errors in census counting or induced by immigrants from countries with no registration of births, and/or by errors in data entry or copying many years ago. Note that errors in population size are relatively higher at advanced ages, where absolute numbers of survivors are lower than those at younger ages.

In some developed countries the situation is worse than in other countries. For example, registration of birth became universal in the USA only in 1940 (Kannisto, 1988). In Russia, the obligatory registration of birth was introduced as late as in 1918 (Andreev, Darski, Khorkova, 1999). It had become universal only by the early 1930s and was distorted during the Great Patriotic War of 1941-45 due to the military occupation of large parts of the country's territory.

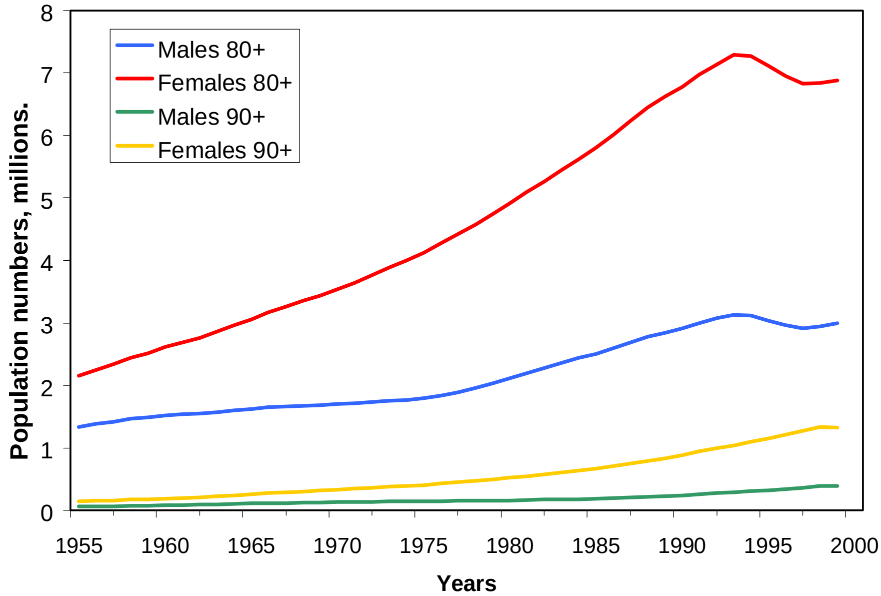

Figure 1. A growth of absolute numbers of people aged 80+ and 90+ in 1956-2000: the total population of Denmark, England and Wales, Finland, France, East Germany, West Germany, Hungary, the Netherlands, Norway, Sweden, and Switzerland.

Source: Kannisto-Thatcher Database on Old Age Mortality at the Max Planck Institute for Demographic Research (available at http://www.demogr.mpg.de/databases/ktdb).

Owing to these reasons, prior demographic analyses found it difficult to obtain precise estimates of populations at advanced ages when using routine estimates of population based on census counts, current registration of deaths, and migrations.

One can avoid several sources of errors in the estimation of elderly populations by using Paul Vincent's method of extinct generation (Vincent, 1951). This method uses death statistics only for a retrospective re-estimation of population at ages, at which international migration is very low and can be ignored. It avoids using population distributions with imprecise age structures and has been repeatedly used for the re-estimation of populations of elderly and for checking the quality of official population figures (Rosenwake, 1979, Kannisto, 1988, Thatcher, 1992, Hill et al., 2000, Thatcher, 2001). The method was extended for application to "almost-extinct" cohorts, e.g., cohorts older than 90 or 95, with only a small proportion of individuals being alive (Thatcher, 1988).

The Human Mortality Database methodology for the re-estimation of populations aged 80+ includes a powerful combination of methods (Wilmoth et al., 2004) built around P.Vincent's idea of retrospective estimation of population from deaths, further developed by V. Kannisto and R. Thatcher. The present study performs a systematic comparison of HMD population estimates with current population estimates for eleven countries of Europe. First, we provide a brief review of the HMD methodology, outlining its methods for the estimation of populations aged 80+ by sex, age, and year of birth. In Section 2, we compare the HMD and the official population estimates for East and West Germany from the 1950s onward. In Sections 3 and 4, we analyze relative and absolute differences between the two series of estimates for other nine European countries. In Section 5, we identify factors influencing variation in the absolute values of the relative differences between the HMD and the official populations by means of regression analysis. Finally, we briefly summarize the empirical findings.

1. Methods and data

1.1. Methods for re-estimation of the population size at ages 80+. A brief summary of HMD methodology[1]

In most countries, years-population counts by one-year age group are available up to age 90 and are followed by the total of population aged 90+. For this common case, Figure 2 shows the zones of the Lexis diagram, in which different methods for the estimation of the number of survivors aged 80+ are applied.

---

[1] Material of section 1.1 is entirely based on the HMD Methods protocol by Wilmoth et al. (2002), available at http://www.mortality.org/Public/Docs/MethodsProtocol.pdf. Figures 2, 3, 4, and all formulae originate from this source.

Zone A corresponds to people aged 80-89 at the beginning of the last year t_n. The HMD methodology assumes that officially registered sizes of respective birth cohorts at the beginning of the last year of observation are reliable. The HMD populations on the 1st of January of years t_n-1, t_n-2, …, t_n-9 are taken as equal to the official current population estimates or the inter-census survival estimates (for more details, see the section "Intercensal survival methods" in the HMD "Methods protocol", 2002). These estimates correspond to a conventional approach for estimating populations within inter-census periods from information on demographic events. The fact that during the last ten years of observation some part of HMD population estimates originated directly from official statistics, suggests a force pushing the HMD estimates closer to the official ones. This should be taken into account when interpreting the results of comparisons between the HMD and official estimates.

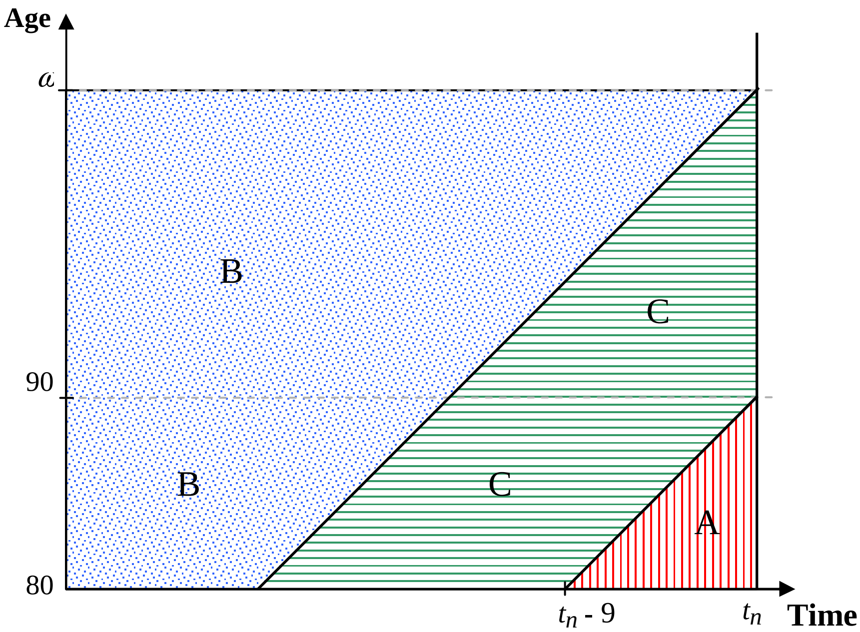

Figure 2. Zones of application of different methods for the estimation of populations aged 80+ in the HMD.
Notes. A - official and inter-census survival estimates; B - extinct cohort estimates; C - survivor ratio (SR90+) estimates of population at the beginning of the last year.

Zone B in Figure 2 corresponds to the extinct cohort method. The method is applicable to cohorts that have reached a certain age of extinction ω by the beginning of the last year t_n. Age ω is defined as age, after which there are virtually no deaths. More precisely, for a candidate age x we count the average number of deaths over cohorts, aged x, x+1, …, x+4 on January 1st of the year y_n, over the years from y_n-5 to y_n-1 (Figure 3):

$$\tilde{D}(x,t_n) = \frac{1}{5}\sum_{j=1}^{5}\sum_{i=0}^{j-1} D^c(x+i, t_n - x - j) \ , \qquad (1)$$

ω is defined as the lowest age *x*, such that $\tilde{D}(x,t_n) \le 0.5$.

According to the method of extinct generation, population size at age *x* for a cohort in question equals the total of all deaths in this cohort occurred at age from *x* and older ages:

$$P(x,t) = \sum_{i=0}^{\infty} D^c(x+i, t-x) \qquad (2)$$

Here, $D^c(x,t)$ is the number of deaths in the cohort born in year *t* recorded among those aged $[x, x+1)$; $P(x,t)$ is the population size on January 1st of year *t* at age *x*.

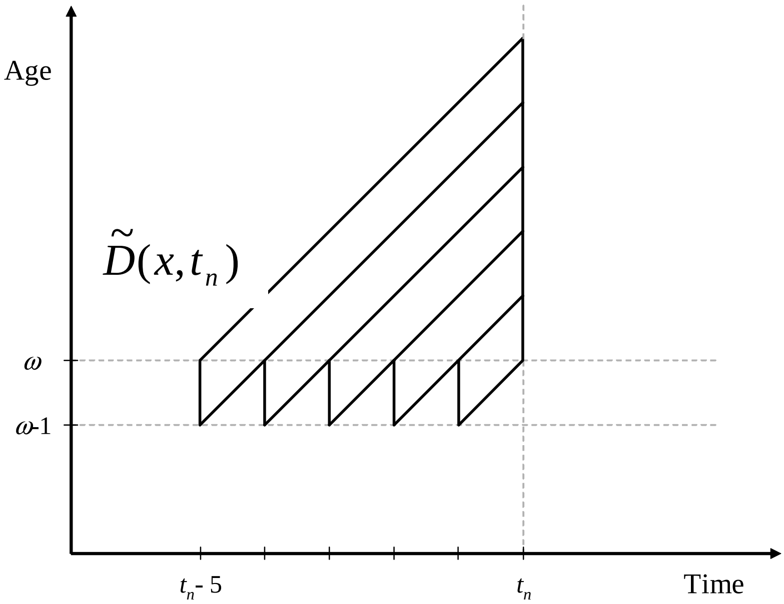

Figure 3. Illustration of the extinction rule (with *x* = ω - 1).

The extinct generation method assumes no international migration after age *x*. In the absolute majority of cases, this is a very reasonable assumption for advanced ages.

Zone B in Figure 2 corresponds to cohorts that can be considered as almost extinct on January 1st of the year y_n. At this time point, these cohorts are aged from 90 to ω and therefore only small fractions of their initial size remain alive.

In some countries with population registers, the official data are available by one-year age groups up to the highest age. These data are reliable. In such easy situations, the HMD population estimates for almost extinct cohorts can be calculated similarly to formula (2):

$$P(x,t) = \sum_{i=0}^{t_n - t - 1} D^c(x+i, t-x) + P\left(x + (t_n - t),\ t_n\right) \qquad (3)$$

However, for the majority of countries and years, only an aggregated population count for age 90+ is available or population estimates at single ages from 90 to ω -1 are available, but they are not reliable. In these situations, the survival ratio method is applied for obtaining the population estimates at ages from 90 to ω-1 at the beginning of the last year. Various modifications of this method have been proposed (see Andreev,1999, for a comparative analysis). The HMD exploits the version that proved to be the most robust (Thatcher et al., 2002).

The idea behind the method is simple. As with the extinct cohort method, one has to assume that the whole change in population size is determined by deaths only. In addition, it is assumed that a 5-year survival in the oldest non-extinct cohort aged ω-1 on January 1st of the year y_n is the same as the average survival in 5 prior extinct cohorts. This allows estimating the population size of the cohort ω-1 on January 1st of the year y_n. The same procedure then is applied to the next younger cohort and so forth down to the cohort aged 90 on January 1st of the year y_n (see the section "Survival ratio" in the HMD "Methods Protocol" for more details).

The HMD methods for population estimates require reliable data on deaths. Unfortunately, that is not true for all data available in the HMD. For instance, deaths can be undercounted in statistics of the 18th and 19th centuries and age at death can be misreported. Below we will give an example of problems in death data.

Difficulties with death counts. The extinct generation and the survivor ratio methods described above require data on deaths by Lexis triangle (age and year of birth). In the HMD, death counts are collected at the finest level of detail available – ideally, by sex, completed age, year of birth, and calendar year. This is, however, not the case for several countries, where such data are unavailable. Among the countries under consideration in this study, England and Wales and Russia have death counts classified by the Lexis squares (age at death) without an additional split by year of birth. For splitting deaths by cohort from a single age group by birth cohort, we use an equation obtained from regression analysis of deaths by Lexis triangle data of Sweden (1901-1999), Japan (1950-1998), and France (1907-1997) (see the section "Splitting 1x1 death counts into Lexis triangles" in the HMD "Methods Protocol" for more details).

Another even more serious difficulty is related to death counts in open-age intervals. The HMD methodology includes a method for the distribution of deaths within open-age group. It relies on the assumption that deaths follow the pattern of a stationary population with an age-specific pattern of death rates given by the Kannisto model of the mortality age curve (Thatcher, 1999) (see the section "Splitting deaths in an opened-age interval by Lexis triangles" in the HMD "Methods Protocol" for more details).

Finally, the quality of death statistics may be poor in some countries during certain time periods. This makes problematic the whole HMD methodology for the population re-estimation on the basis of death count data. Users are informed about problems of this sort by special warning messages on the HMD country-pages. It also makes problematic comparisons between the HMD and official population estimates. Indeed, one can not be sure that the HMD populations are preferable compared to the official ones.

According to Kannisto (1988), a slow decrease of deaths with age, a high male to female death ratio, and age heaping around ages ending by 0 and 5 can indicate the quality problems in the data on deaths at advanced ages. Figure 4 shows age-specific death rates for Russian males at ages around 85 and around 90, with a clear age heaping in the years before 1970. In the present study, we consider Russian data beginning from 1970 onwards.

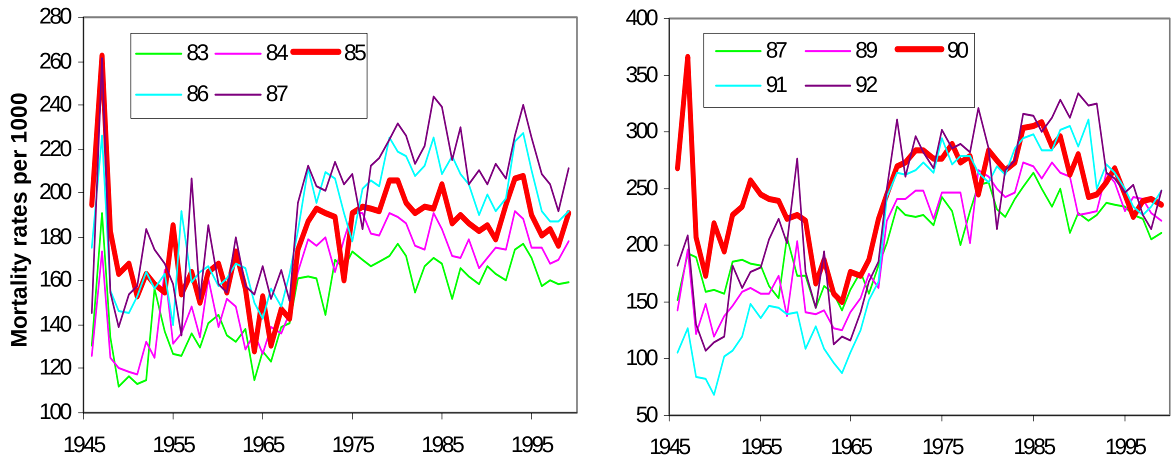

Figure 4. Mortality rates at selected ages, Russia, Males, 1946-1999.
Source: Zakharov (2001).

1.2. Country data series

We selected 11 western and eastern European countries for comparative analyses. These are Denmark, England and Wales, Finland, France, West Germany, East Germany, Hungary, the Netherlands, Russia, Sweden, and Switzerland. We aimed to select countries representing a variety of statistical systems and shapes of data on population and deaths. These are countries with national (Denmark, Finland, and Sweden) or municipal (the Netherlands) population registers, countries with residential registers (Switzerland, East and West Germany), with some residential registration (France, Hungary, and Russia), and countries with no regular residential registration (England and Wales).

Within the time periods under consideration, all countries are assumed to have reliable death statistics. In order to satisfy this condition, for several countries we analyze shorter data series than those available in the HMD. For countries with long data series beginning in the 19th century, such as Denmark, Finland, England and Wales, and Switzerland, we decided to begin from 1910. For Sweden, the time period is even shorter since HMD does not include annual official population estimates for years before 1960. Finally, due to problems with the quality of the death data indicated in the previous section, we excluded years prior to 1970 from the Russian series.

Table 1 presents important characteristics of the data used in the present study.

Table 1. Review of the data

Country	Time period	Census years	Years with opened age interval (90+-100+) in data on deaths	Opened age interval in the population data of the last year	Shape of the eleme-ntary cell on the Lexis diagram	Population Register[#]
Denmark	1910-2003	1911, 1921, 1930, 1940, 1945, 1950, 1955, 1960, 1965, 1970, 1976, 1981, 1991, 1998	1921-1942	No	Triangle, Square in 1910-20	Since 1968
Finland	1910-2003	1951, 1961, 1971, 1976, 1981, 1985, 1990, 1995, 2001	No	No	Triangle, Square in 1910-16	Since 1969
Sweden	1960-2003	1961, 1965, 1970, 1975, 1980, 1985, 1990	No	No	Triangle	Since 1970
The Netherlands	1950-2002	1947, 1960, 1971	No[##]	Yes	Triangle	The whole period
England & Wales	1911-2001	1951, 1961, 1971, 1981, 1991, 2001	No	Yes	Square	No
France	1910-2002	1911, 1921, 1926, 1931, 1936, 1946, 1954, 1962, 1968, 1975, 1982, 1990, 1999	1934-1935, 1947[####]	Yes	Triangle	No
Switzerland	1910-2003	1910, 1920, 1930, 1941, 1950, 1960, 1970, 1980, 1990, 2000	1910-1949	Yes	Triangle	No
Germany, West	1956-2000	1951, 1961, 1970, 1987	1956-1963	Yes	Triangle	No
Germany, East	1956-2000	1950, 1964, 1971, 1981	1956-1989	Yes	Triangle	1985-89
Hungary	1950-2002	1949,1960,1970, 1980, 1990, 2001	1950-2002	Yes	Triangle	No
Russia	1970-2003	1970,1989, 2002	1970-2003	Yes	Square	No

Notes.
[#]Since this year, the population register provides a basis for compiling annual statistics on the population and its composition by age and sex.
[##] The NSO of the Netherlands provides data on deaths with an open-age interval 108+.
[###] Population data for France for 1910-1990 are taken from Tableau I-C-1 of Vallin and Meslé (2001).
[####] 105+ in 1998-2001.

2. German population estimates since 1956

When looking at empirical results, one should keep in mind that, according to the HMD methodology, HMD population estimates must be equal to the official population estimates at the beginning of the last year. In addition, during the last 10 calendar years HMD population estimates for ages from 80 to 89 include official population numbers: for ages from 80 to 89 at the beginning of the last year, for ages from 80 to 88 at the beginning of the previous year etc. This means that the sense of comparison of the HMD and the official population estimates is different for the last decade due to the “gravitation” to the official population numbers. It suggests also that the comparison will be less problematic for ages 90+ than for ages from 80 to 89.

Figure 5 shows relative differences between the HMD and the official population estimates by age and calendar year on the Lexis diagram[2] in East and West Germany. The proportional difference is calculated as

$$\delta(x,t) = \frac{P^{HMD}(x,t) - P^{official}(x,t)}{P^{HMD}(x,t)}.$$

In Figure 5 the magnitude of the difference $\delta(x,t)$ corresponds to color of the Lexis trapezoid connecting points (x, t) with (x+1, t) and (x+1, t+1) with (x+2, t+1). Figure 5 contains a large amount of information. The white space in the upper part of the panels corresponds to open-ended age intervals, where official population numbers by single-year age group are unavailable. The relative differences tend to form diagonal structures corresponding to birth cohorts. Differences tend to increase with age and this pattern seems to be more pronounced for men than for women. Continuous color patterns are periodically interrupted by vertical lines corresponding to the census points. At these dates, official populations are re-estimated and new inter-census periods begin. Immediately after the census years, the relative differences tend to be lower but increase again shortly after.

In Figure 5, the blue colors correspond to negative relative differences (e.g., HMD estimates are lower than the official ones), while the red colors correspond to positive relative differences (e.g., HMD estimates are higher than the official ones). Relative differences over 10 percent are mostly observed at ages over 90.

---

[2] The program for building Lexis maps is freely available (Andreev, 1999).

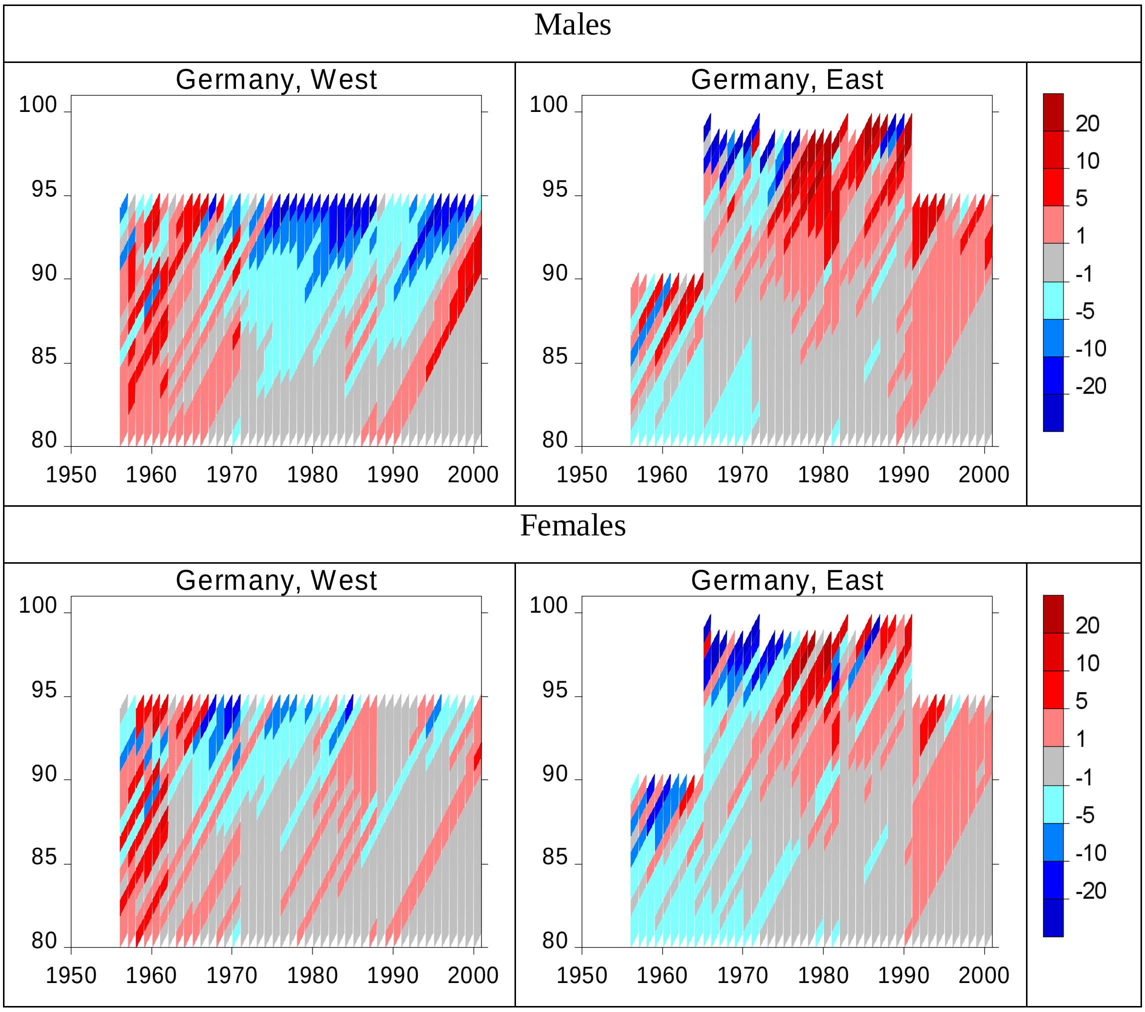

Figure 5. Relative difference between the HMD and the official population estimates by age and calendar year: East and West Germany. (In percent).

One can identify a period-type distortion on the Lexis surface between 1990 and 1991 in East Germany. This is comparable to the effects the population censuses cause. It could be a result of the re-estimation of East German population for its integration into the West German statistical system just after the unification.

Figure 6 presents a more aggregate pattern of relative differences between the HMD and the official populations aged 80+ and 90+. Although these differences generally decrease at census years, the West German census of 1970 is an exception since both for age 80+ and age 90+ the differences increased to negative values at the beginning of 1971 when the new census population replaced the estimated one. Comparison of the upper and the lower panels of Figure 6 shows the extent to which the relative difference is greater in the older age group. Comparison of the right and the left panels reveals larger relative differences in the 1970s-80s in West Germany compared to East Germany.

Before the beginning of the 1960s, official estimates of population aged 80+ in West Germany are 3-4% lower than the HMD ones. During the 1960s the relative difference becomes greater for men than for women and decreases to a very low positive value. During the 1970s-80s the relative difference is quite stable across time with a very low value for women. However, in 1971, after the census of 1970, it becomes negative and its magnitude grows of about 1% for men. During the 1970s-80s it remains stable at level of about 1.5%. At the beginning of 1988 it drops dramatically to almost zero due to the population re-estimation after the census of 1987.

For the West German population in the 1970s-80s aged 90+ the magnitudes of negative relative deviations of official estimates from the HMD ones is about 5% for women and increases from 5.5% in 1971 to over 20% in 1987.

Variations in the relative differences during the 1990s are strongly influenced by an attraction of the HMD estimates to the official population of the last year as conditioned by the HMD methodology. In the early 1990s a small shift upwards for age 80+, both in West and East Germany and a significant shift downwards for age 90+ for men in West Germany is visible.

In East Germany in the 1950s and the early 1960s, official population estimates for age 80+ exceeded the HMD estimates by about 3% for women and about 2% for men. The difference fell to 1% due to the 1964 census and remained at this level until the next census of 1970. During the 1970s, the negative relative difference decreased and in the second half of the 1970s it reached the zero level for women and became positive, but stayed low for men (0.3-0.4 percent). By the time of the census of 1981, the relative difference dropped to almost zero both for men and women. Throughout the 1980s, the relative difference increased to about 1.5% in 1990.

Temporal changes in the relative difference between the two population estimates for the East German population aged 90+ are quite similar to those for age group 80+. Large fluctuations observed in the older age group in the early 1960s constitute an exception.

The positive deviations from the HMD data in the 1950s-60s in West Germany suggest a population underestimation by official statistics. This could be due to some problems in the census counting in the West German censuses of 1951 (Dinkel and Meinl, 1991) and 1961. By contrast, in East Germany we find a population surplus compared to the HMD. Taking into account the coincidence of the upward and the downward shifts in the two parts of Germany as seen in Figure 6, one can expect some influence of an east-west migration.

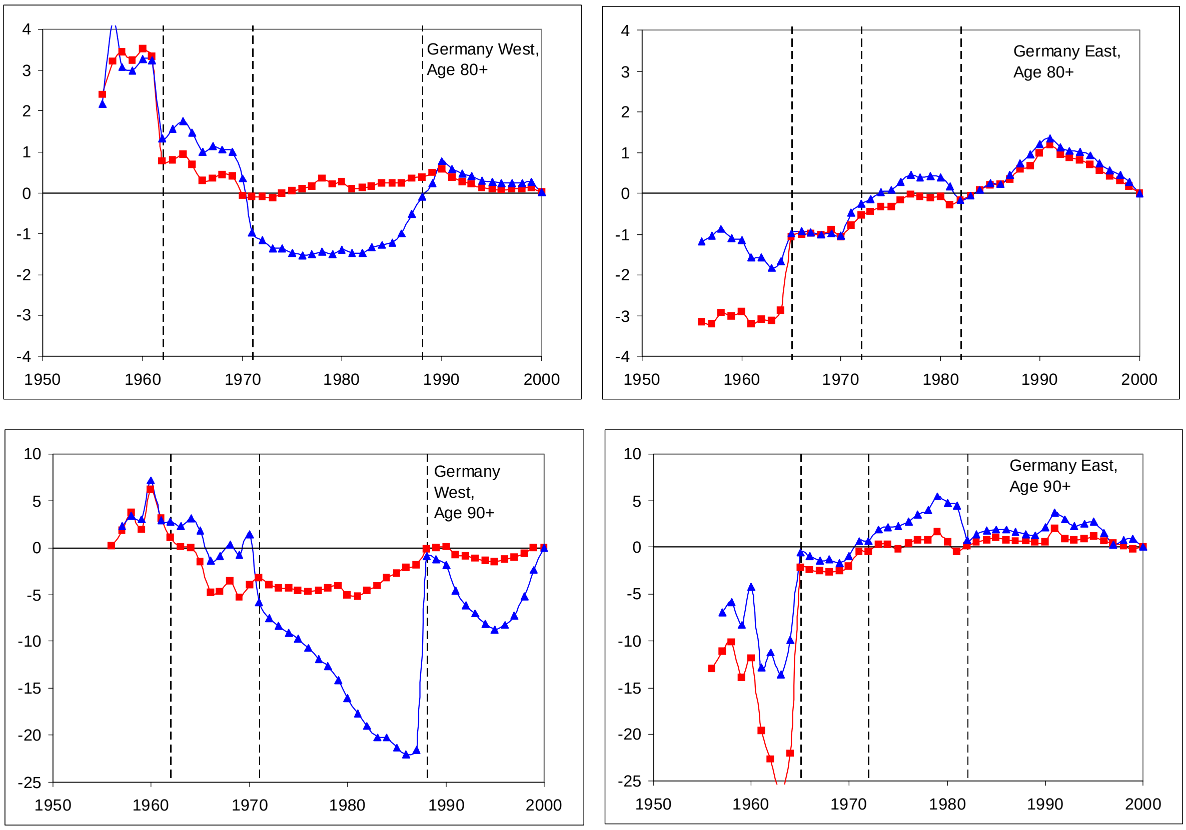

Figure 6. Relative difference between the HMD and the official population estimates aged 80+: East and West Germany. (In percent).

Note. Red - females, blue - males. Vertical lines correspond to the point when the effect of the population censuses is visible.

Figure 7 shows the same trends for the total population of Germany. They largely replicate the West German trends in Figure 6. A smaller magnitude of variations in Figure 7 confirms a partial compensation among the two German populations.

In the following we will focus on data of the last three decades not considering earlier periods. In the 1970s-80s, there is a clear pattern of overestimation of the official population especially for men. Between 1971 and 1987 official estimates of the male population aged 80+ are 1-1.5% higher than the HMD estimates. This corresponds to an increase in the absolute difference between the two population estimates from about 5 thousand men in 1971 to about 7 thousand men in 1987. For age 90+ the relative difference between the two population estimates for men is higher than that for age 80+ and tends to increase with time. From 1971 to 1987 the excess of men in the official statistics compared to HMD rises from 5 to 20 percent. This corresponds to an increase in absolute numbers from 2 to 10 thousand men.

The new census of 1987 resulted in a sharp drop of the relative difference between 1987 and 1988. As we mentioned earlier, it is difficult to evaluate dynamics of the difference in the 1990s due to a provisional character of HMD population estimates and their dependence on the official population at the beginning of the year 2000. However, a preliminary update of the HMD population with the 2003 data shows that even in 1988, when the last West German census was implemented, the HMD population estimates were somewhat lower than the official ones (Left panel of Figure 7). The same comparison for age 90+ (Right panel of Figure 7) shows that the pattern of a growing population overestimation in the official statistics, similar to that in the 1970s-80s, is likely to be present in the 1990s too.

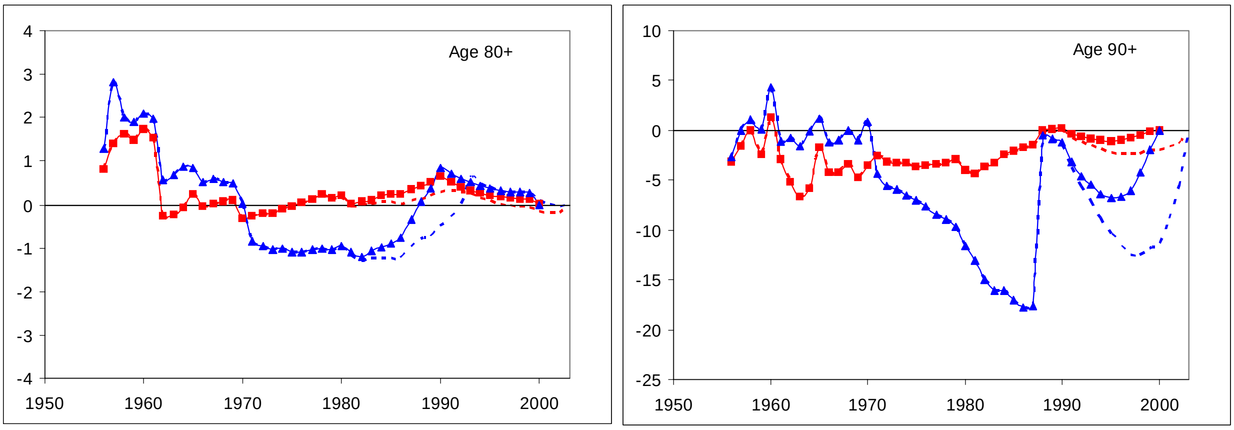

Figure 7. Relative difference between the HMD and the official population estimates aged 80+ and 90+: population of the whole of Germany. (In percent).
Notes. Red - females, blue - males. Dashed lines correspond to preliminary HMD update up to the beginning of the year 2003.

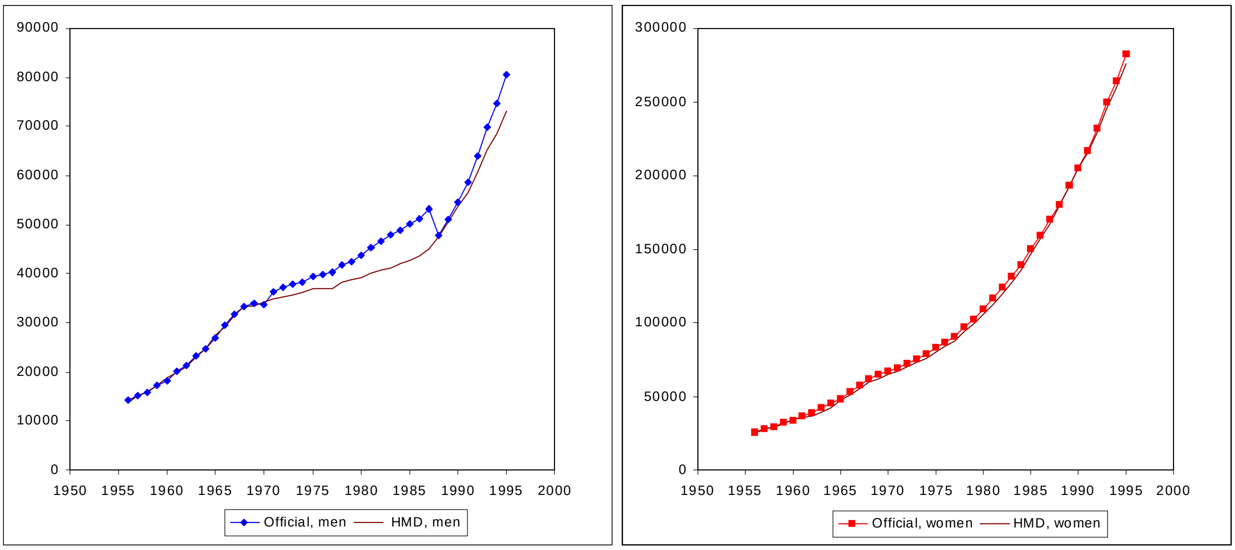

Figure 8. Trends in the official and HMD populations aged 90+ for the total population of East and West Germany.

Figure 8 shows temporal changes in the absolute population sizes for the total population of East and West Germany aged 90+. Here we use preliminary HMD estimates with 2003 as the last year. The figure confirms that the widening gap between the official and the HMD male populations over the 1980s is the most remarkable phenomenon. It seems also that this pattern is replicating in the 1990s.

For women the absolute differences are smaller than for men. For example, in 1995 they are estimated as 7.5 and 6.1 thousand for men and women respectively. However, in agreement with Figure 7, the relative difference is much greater for men than for women: 19% *vs*. 2%. Consequently, mortality effects of the population overestimation must be also greater for men.

Figure 9 presents a comparison of death rates at age 90+ based on the official and on the HMD population estimates. Both time series reveal a general mortality decline. In the late 1960s the HMD and the official death rates for men are almost identical. However, there is a surprisingly steep decrease of the official male mortality over the inter-census period in West Germany in the 1970s-80s. As a result, "official" death rates for men in 1982-87 are almost as low as those for women. Just after the West German census of 1987, male death rates calculated from the official population suddenly jump to HMD values. However, later in the 1990s a reduction of the male-female mortality difference resumes.

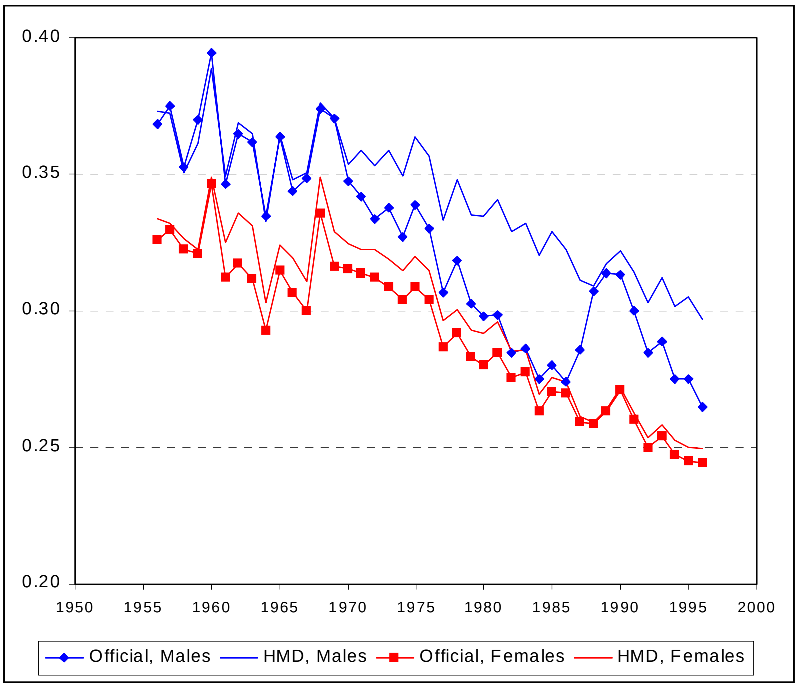

Figure 9. Trends in death rates at age 90+, calculated from the official and the HMD population estimates, for the combined population of East and West Germany.

The analysis of this section points at an overstatement of population at advanced ages in the current population statistics of Germany. In this connection, three important features should be mentioned. First, the relative differences are much greater for men than for women. Second, the problem is mostly characteristic for West Germany and relatively minor in East Germany. Third, deviations from the HMD estimates drop just after the West German census of 1987.

It is clear that the current mortality statistics are characterized by certain numerator-denominator bias. Men experience more problems with the denominator probably because they are more mobile than women. The indirect population estimates by the SR-method (Thatcher et al., 2002) show an overestimation of the official population of around 20 percents for men and 1-2 percent for women in 2003. It appears as if some fraction of the population is not exposed to mortality at all. This fraction moves up to higher ages without any reduction, while the remaining population rapidly diminishes due to high mortality at advanced ages. If this is true then the relative weight of the special fraction would increase with age steeper for men than for women due to the sex differences in mortality.

So far, the mechanism of the overestimation of the elderly population in Germany remains unknown and has to be identified.

The traditional problem of age overstatement (Coale and Kisker, 1986, Kannisto, 1988) could also be present in Germany. As one could expect, overstatement of age is more frequent among men who emigrated from countries with poor quality of civil registration in the past (for example, Turkey and the former Soviet Union). According to the Federal Statistical Office, in 1996-99 the percentage of foreign men in the population was about 7-8% at age 90+ *vs*. only 3% at ages from 80 to 89. For women respective percentages were 4% *vs*. 2%. Death rates among people, whose real age is lower than their "statistical" age, would be artificially low. However, this type of bias can not diminish due to a census (as it happens in our data), since people with overstated age are *really present* in the population and shall be counted.

It seems that the huge immigration into Germany can not directly contribute to the difference between the HMD and the official population estimates. Indeed, immigration should affect the HMD population estimates in all years including census years. We know, however, that the differences between the two population estimates dropped at the beginning of 1988 when the official population was taken from the West German census of 1987. This means that immigration in the period from 1988 to the beginning of the last year (either 2000 or 2003) did not strongly influence the HMD population at the beginning of 1988.

This study lacks detailed data for a thorough investigation on the mechanism underlying the overestimation of the elderly. A comparison of pre- and post-census characteristics of population for the census of 1987 and analyses of micro-data of the 1990s could help to find a real explanation for the phenomenon and (perhaps) would allow to undertake some measures to improve the present situation.

3. A comparison of the population estimates since 1950 for other countries

Figure 10 presents relative differences between the HMD and the official populations aged 80+ for nine European countries. Three Nordic countries, placed in the upper row, have population registers since 1968-1970. Correspondingly, there is almost no difference between the official and the HMD estimates during the last decades. The case of Finland is especially illustrative. This country had a significant undercount of the elderly in the 1960s. The error was continuously decreasing over the 1970s and has become extremely small afterwards.

Countries presented in the second row of Figure 10 experience somewhat greater relative deviations from the HMD estimates during the last three decades. The magnitudes of the differences are comparable to those in Germany, with most of the deviations being below one percent. In Switzerland, the error is slightly lower than in France and the Netherlands.

Since the mid-1980s, official statistics overestimate population numbers in the Netherlands. In France, there is some overestimation for men and underestimation for women in the population statistics. Although the comparison reveals some overestimation of men and women in Switzerland, its magnitude is small.

Countries placed in the third row of Figure 10, i.e., England and Wales, Russia, and Hungary, experience significantly greater relative differences between the HMD and the official population estimates. In these countries, official population estimates often differ from the HMD estimates by 2 percent and more during the inter-census periods and the difference tends to diminish at the census points.

In England and Wales, official population numbers were understated in the 1950s-60s and overstated in the 1980s-90s. The magnitude of possible errors in this country is generally higher than in Germany and France. One should keep in mind, however, that population of England and Wales is not really a national population. It is exposed to more intensive internal migration flows and in this sense is incomparable to other countries.

In Russia and Hungary, a significant overestimation of the elderly in the 1980s has been replaced by significant underestimation in the 1990s. These countries also experience high levels of potential errors.

4. Absolute differences

Table 2 shows what the relative differences considered so far mean in absolute numbers. The latter tell about sizes of potential problems in national-level planning.

In some countries and time periods, both negative and positive differences are observed. Differences of different signs can partly compensate each other when averaging them over 10-year periods. Therefore, two types of differences are shown in Table 2: average annual differences and average annual absolute-value ones.

In the 1980s and the 1990s, Russia had the greatest difference between the HMD and the official populations, close to 40 and 30 thousand people, respectively. England and Wales had the second greatest difference: about 24 and 25 thousand in the 1980s and the 1990s respectively. Germany has a greater number of people aged 80+ than England and Wales, but a much smaller deviation from the HMD than in England and Wales: about 12 thousand both in the 1990s and the 1980s (FRG + former GDR). However, the number is high. France comes next, with about 10 and 12 thousand in the 1980s and the 1990s, respectively. Hungary is characterized by large differences for its relatively small population size: 7 and 8 thousand in the 1980s and the 1990s, respectively.

In other countries, the differences in the 1980s-90s were smaller and even negligible in Denmark, Finland, and Sweden.

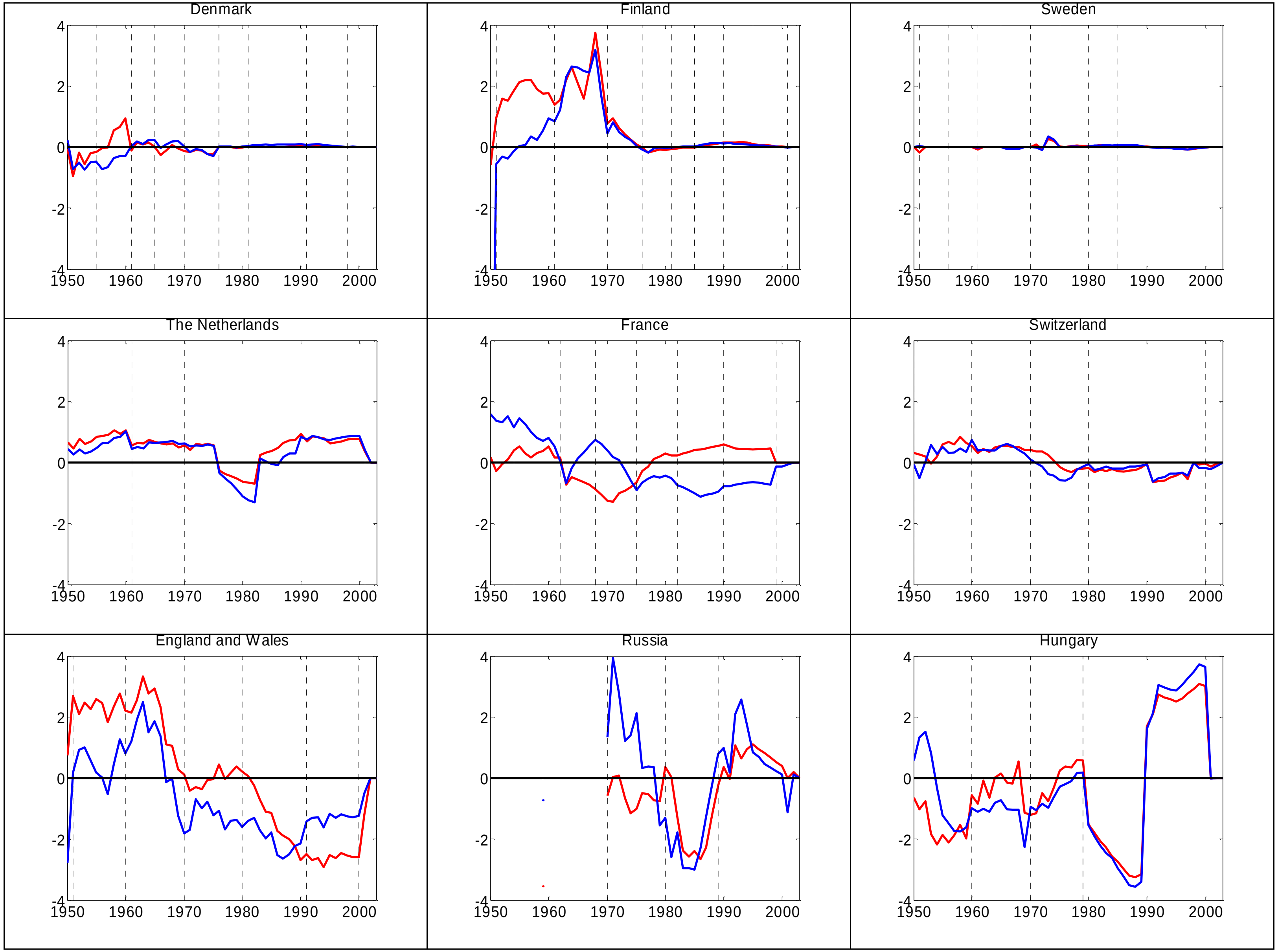

Figure 10. Relative difference between the HMD and the official estimates of the population aged 80+ in nine countries of Europe. Red - females, blue - males. (In percent).

Table 2. Average annual populations aged 80+ by calendar decade according to the HMD and official sources. Average annual differences and average absolute differences between the two estimates.

Country	1950-59			1960-69			1970-79			1980-89			1990-99		
	HMD	Δ (Δ%)	$\vert\Delta\vert$ (%)	HMD	Δ (Δ%)	$\vert\Delta\vert$ (%)	HMD	Δ (Δ%)	$\vert\Delta\vert$ (%)	HMD	Δ (Δ%)	$\vert\Delta\vert$ (%)	HMD	Δ (Δ%)	$\vert\Delta\vert$ (%)
Denmark	66838	-169 (-0.25)	273 (0.41)	52657	75 (0.14)	158 (0.30)	130814	-118 (-0.09)	124 (0.09)	180189	49 (0.03)	51 (0.03)	223467	69 (0.03)	70 (0.03)
Finland	38124	278 (0.73)	613 (1.61)	50154	1086 (2.17)	1086 (2.17)	68007	138 (0.20)	200 (0.29)	117893	25 (0.02)	66 (0.06)	175246	161 (0.09)	161 (0.09)
Sweden	-	-	-	215516	-124 (-0.06)	124 (0.06)	245141	146 (0.06)	191 (0.08)	332167	122 (0.04)	122 (0.04)	444251	-186 (-0.04)	195 (0.04)
Netherlands	137046	933 (0.68)	933 (0.68)	205332	1330 (-0.65)	1330 (0.65)	285799	280 (0.10)	1515 (0.53)	405021	242 (0.06)	2121 (0.52)	522268	4054 (0.78)	4054 (0.78)
England and Wales	835137	12804 (1.53)	14534 (1.74)	1106211	18733 (1.69)	19712 (1.78)	1342159	-4778 (-0.36)	7005 (0.52)	1725129	-23929 (-1.39)	24519 (1.42)	2244724	-29794 (-1.33)	29794 (1.33)
France	838642	4589 (0.55)	4924 (0.59)	1126274	-2547 (-0.23)	6364 (0.57)	1414298	-7068 (-0.50)	8085 (0.57)	1949113	214 (0.01)	10328 (0.53)	2496811	2374 (0.10)	12367 (0.50)
Switzerland	70922	266 (0.38)	297 (0.42)	103145	483 (0.47)	483 (0.47)	141346	-143 (-0.10)	387 (0.27)	218569	-482 (-0.22)	482 (0.22)	300486	-1218 (-0.41)	1218 (0.41)
Germany West	987748	30661 (3.10)	30661 (3.10)	1074353	13796 (1.28)	13796 (1.28)	1420054	-4694 (-0.33)	6960 (0.49)	2130085	-1389 (-0.07)	9853 (0.46)	2866288	7062 (0.25)	7062 (0.25)
Germany East	287854	-7450 (-2.59)	7450 (2.59)	381306	-6414 (-1.68)	6414 (1.68)	450634	-1168 (-0.26)	1649 (0.37)	548636	1168 (0.21)	1666 (0.30)	611876	4608 (0.75)	4608 (0.75)
Germany	-	-	-	-	-	-	-	-	-	-	-	-	3478164	11670 (0.34)	11670 (0.34)
Hungary	100724	-1125 (-1.12)	1473 (1.46)	143562	-868 (-0.60)	1007 (0.70)	192257	-428 (-0.22)	1084 (0.56)	254810	-6762 (-2.65)	6762 (2.65)	311157	8311 (2.67)	8311 (2.67)
Russia	-	-	-	-	-	-	1823065	-3454 (-0.19)	14907 (0.82)	2464278	-37874 (-1.54)	40051 (1.63)	3450007	26839 (0.78)	27084 (0.79)

5. Factors of deviation from the HMD estimates

In previous sections, we described differences between the HMD and the official population estimates over the period since 1950. After looking at these data, one could notice a few instructive regularities. In particular, it seemed that the deviation tended to increase with age, tended to be lower at the moment of population censuses and also has diminished much in Nordic countries after the introduction of the population registers. A regression analysis allows for the identification of these regularities in a more objective and precise manner.

Table 3a presents the results of the ordinary least squares regression connecting the absolute values of relative differences between the HMD and the official populations for five countries with long data series available. The results generally confirm the prior observations. Deviations between the population estimates are significantly associated with sex, early calendar periods, and high ages. The relations with the census points are statistically significant for England and Wales only and actually increase the error. This paradoxical result can be possibly explained by a problematic quality of early English censuses in respect to the elderly. For other countries, this relation goes into expected direction, but appears to be insignificant. The effects of the open-age interval and of the register are insignificant, too. The latter is clearly related to small fractions of the total number of years lived with population registers.

Table 3a. Outcomes of the OLS regression of the absolute value of the relative difference between the HMD and the official estimates on characteristics of the data since 1910.

Variable		England & Wales 1910-2001	France 1910-2002	Switzerland 1910-2003	Finland 1910-2003	Denmark 1910-2003
Sex	M	0	0	0	0	0
	F	**-1.1/0.2/0.000**[#]	**-1.3/0.4/0.000**	**-3.2/1.0/0.002**	**-14.8/4.9/0.002**	0.1/0.6/0.855
Years	1910	**1.2/0.5/0.032**	**6.2/1.5/0.000**	**18.3/5.7/0.001**	25.8/25.6/0.315	3.9/3.3/0.228
	1920	**1.9/0.6/0.002**	-0.6/1.5/0.704	**16.6/5.7/0.004**	27.5/25.6/0.284	3.0/3.3/0.359
	1930	**1.3/0.6/0.042**	-0.6/1.5/0.673	**12.2/5.7/0.033**	29.5/25.6/0.250	3.1/3.3/0.329
	1940	**3.7/0.5/0.000**	1.4/1.2/0.242	2.7/5.2/0.606	**92.7/25.3/0.000**	**9.8/2.2/0.000**
	1950	**1.8/0.5/0.000**	**3.1/0.7/0.000**	0.5/2.1/0.799	15.7/25.2/0.532	1.5/2.1/0.475
	1960	**1.3/0.5/0.000**	1.1/0.7/0.112	-0.9/2.1/0.669	15.5/25.8/0.549	1.5/1.7/0.384
	1970	-0.2/0.4/0.732	0.03/0.7/0.963	-1.0/2.1/0.637	1.4/25.1/0.953	1.7/1.2/0.149
	1980	-0.2/0.4/0.610	-0.3/0.7/0.645	-0.2/2.1/0.910	-0.5/25.1/0.953	-0.1/1.1/0.933
	1990	0	0	0	0	0
Ages	80-84	0	**0**	0	0	0
	85-89	**1.6/0.3/0.000**	0.4/0.5/0.391	1.1/1.5/0.462	4.1/6.4/0.523	0.8/0.8/0.279
	90-94	**6.3/0.4/0.000**	**3.4/0.5/0.000**	**5.7/1.5/0.000**	6.9/6.7/0.302	1.5/0.9/0.100
	95+	-	**8.2/0.6/0.000**	**23.4/1.5/0.000**	**68.9/7.8/0.000**	**11./0.9/0.000**
Census	Yes	**1.0/0.4/0.017**	-0.6/0.5/0.222	-0.2/1.7/0.908	-3.0/8.2/0.711	-1.3/0.8/0.120
	No	0	0	0	0	0
Opened-age interval in deaths	Yes	-	1.8/1.3/0.154	3.7/5.3/0.485	-	-1.7/2.3/0.468
	No	-	0	0	-	0
Register	Yes	-	-	-	-0.8/23.3/0.973	-3.2/1.7/0.075
	No	-	-	-	0	0

Constant	**0.9/0.4/0.017**	-0.2/0.6/0.702	**-4.3/1.8/0.015**	-11.0/24.7/0.655	-0.1/2.1/0.944
R^2	0.46/0.02/0.000	0.36/0.05/0.000	0.46/0.14/0.000	0.27/0.63/0.000	0.40/0.07/0.000

[#] Regression coefficient/standard error/p-significance.

One can see also that the situation differs among countries. For example, in France and Denmark relations with time periods are not pronounced. In Finland and Denmark, the differences are significantly greater in the war period of the 1940s.

Table 3b shows similar regression outcomes for a selection of six countries for shorter time series starting in the 1950s. In all countries, a significant effect of older age can be detected as well, together with a significant effect of sex in France, West Germany, and Hungary. There are also significant time periods effect in all countries except Denmark. In contrast with prior regression analysis, the effect of the introduction of the population register in Denmark is significant.

Summary of findings and discussion

During the last four decades, the number of people aged 80+ in European countries has tripled. At the same time, national statistical systems experience particular difficulties in producing precise population estimates at advanced ages. The present study reviews official current population estimates by comparing them to the equivalent population estimates from the Human Mortality Database. The latter estimates are based on the extinct and the almost-extinct generation methods. Population estimates at advanced ages are very sensitive to the quality of statistics. The same absolute error leads to much higher effect in older ages due to a small number of survivors. As a result, the precision of mortality estimates steeply decreases with age across advanced ages. The disagreement between the HMD and the official population estimates also increases with age and tends to be higher for men than for women.

HMD is a powerful instrument for obtaining better population and mortality estimates at old ages. The HMD methodology relies on the death data. The whole range of HMD population estimates can be split into two parts. The first and biggest part is based on pure summation of deaths and depends only on the quality of death statistics. As it is high enough in most developed countries, HMD population estimates are usually precise enough up to the mid- to late-1980s. The second part of the HMD population estimates covers the last 10-15 years and their precision depends on the quality of the last-year population estimates (the right edge of the observation region). Thus, the principal causes of the differences between the HMD and the official population estimates vary in time.
For the period before the last 10-15 years of observation, the difference is mostly explained by problems in the official data. In countries without Scandinavian-type centralized population registers (such as West Germany) the deviation of the official population estimates from the HMD in-

creases as time passes after the last census, reaches a maximum just before the next census and drops off just after it.

The comparative analysis for 11 European countries revealed especially significant relative and absolute differences between the official and the HMD estimates in Russia, Hungary, and England and Wales. Smaller, but considerable differences between the two population estimates were found in France, Germany, and (to a lesser extent) in the Netherlands.

The precision of the HMD population estimates during the last 10-15 years requires more attention. Once again, this is not a real problem in Sweden, Denmark, and Finland, where since the early 1970s the quality of population estimates is permanently high. For France, Switzerland, and The Netherlands the influence of the right edge is potentially more important. However, relatively small differences between the HMD and the official population estimates and the plausible shapes of the HMD mortality curves suggest good quality of the HMD data in these three countries. For England and Wales, Germany, Hungary, and Russia the differences are greater. The HMD population estimates for these countries are probably affected by the quality of the right edge. One can hope that the situation will improve in England and Wales after recalculations based on the census of 2001. In Russia one could also hope for improvements in connection with the census of 2002.

We found a clear pattern of population overestimation for West German men in the 1980s-90s. The relative difference between the official and the HMD estimates was especially high at age 90+ and tended to increase between 1971 and 1987 at age 90+. During this period male death rates based on the official population estimates have decreased to implausibly low levels close to female death rates. After the West German census of 1987 the difference between HMD and official estimates disappeared, but a widening of the gap seems to have resumed in the early 1990s. Based on the data available so far, one can conclude that an after-census population correction is highly desirable in Germany. We need to wait for another 7-8 years before the HMD data will provide a more conclusive evaluation of the official population data of the 1990s.

Although the HMD mortality data are much less problematic than the official data, the quality of the population denominator can influence them. The potential magnitude of the problem can be estimated by comparing differences between the HMD mortality estimates and the official mortality estimates for the last pre-census years with the equivalent differences for a year just after the census. For example, for West Germany differences between male life expectancies at age 80 based on the HMD populations and the official populations are negligible in 1985-86 and become even smaller in 1987. However, it is not the case for male life expectancies at age 90. In 1985-86 the difference between the two male life expectancy estimates is about 0.5 years (16%). In 1987 it drops to almost zero.

During the last 10-15 years, some problems in estimating life expectancies at age 90 coinciding with much less problematic estimates at age 80, were observed also for Hungary and England and Wales (analysis not shown here).

Therefore, one should be cautious in using male mortality estimates for the 1990s at ages 90+ for Russia, Hungary, Germany, and England and Wales if the time from the last population census is longer than 3-5 years.

Table 3b. Factors of the absolute value of the relative difference between the HMD and the official estimates: data series since 1950.

Variable		France, 1950-2002	Hungary, 1950-2002	Denmark, 1950-2003	Germany, West 1956-2000	England and Wales, 1950-2001	The Netherlands, 1950-2002
Sex	Males	0	0	0	0	0	0
	Females	**-1.1 /0.3/0.000**	0.2/0.4/0.622	0.3/0.4/0.461	**-6.2/1.5/0.000**	**-0.9/0.2/0.000**	-0.1/0.3/0.738
Years	1950	**3.1 /0.5/0.000**	**1.4/0.7/0.034**	1.8/1.1/0.092	-5.7/4.5/0.205	**1.9/0.4/0.000**	-0.6/0.5/0.201
	1960	**1.0 /0.5/0.032**	**2.8/0.6/0.000**	1.7/0.8/0.051	-2.3/2.4/0.322	**1.3/0.4/0.000**	0.3/0.5/0.535
	1970	0.1 /0.5/0.883	-0.3/0.6/0.576	0.9/0.6/0.094	**5.1/2.2/0.018**	-0.2/0.3/0.642	**1.1/0.5/0.025**
	1980	-0.3 /0.5/0.576	**1.3/0.6/0.029**	-0.01/0.5/0.991	**9.9/2.2/0.000**	-0.2/0.3/0.488	-0.8/0.5/0.122
	1990	0	0	0	0	0	0
Ages	80-84	0	0	0	0	0	0
	85-89	0.3 /0.4/0.509	**1.7/0.5/0.000**	0.1/0.5/0.803	0.3/2.1/0.886	**0.7/0.2/0.008**	0.4/0.4/0.400
	90-94	**1.6 /0.4/0.001**	**7.0/0.5/0.000**	0.7/0.5/0.172	2.1/2.1/0.303	**5.8/0.3/0.000**	**1.7/0.4/0.000**
	95+	**7.5 /0.5/0.000**	**-**	**5.5/0.5/0.000**	**24.7/2.1/0.000**	**-**	**8.3/0.4/0.000**
Census	Yes	0.1/0.5/0.795	0.4/0.7/0.548	-0.1/0.5/0.827	**-6.6/3.1/0.034**	**1.0/0.3/0.003**	-0.4/0.7/0.587
	No	0	0	0	0	0	0
Opened age interval in deaths	Yes	-	-	-	5.2/3.5/0.140	-	-
	No	-	-	-	0	-	-
Register	Yes	-	-	**-2.9/0.9/0.002**	-	-	-
	No	-	-	0	-	-	-
Constant		0.2/0.5/0.600	0.08/0.5/0.875	1.2/1.1/0.263	1.4/2.1/0.506	**1.3/0.3/0.000**	0.4/0.5/0.382
R^2/s.e./ sig.		0.50/0.03/0.000	0.47/0.03/0.000	0.40/0.04/0.000	0.43/0.14/0.000	0.43/0.14/0.000	0.55/0.03/0.000

Notes

1. In fact, the difference does not diminish after some of the old censuses of the 19th or the early 20th century, which produced problematic population counts. Since the 1950s the quality of the census counts has become high enough in all countries under consideration.

References

Andreev, E.M., Darski, L.E. and T.L.Kharkova. 1998. *Demografitseskaia istoriya Rossii 1927-1959. [Demographic History of Russia].* Moscow, Informatika.

Andreev, K.F. 1999. *Demographic Surfaces: Estimation, Assessment and Presentation, with Application to Danish Mortality, 1835-1995.* Ph.D. Dissertation, Faculty of Health Sciences, University of Southern Denmark.

Coale, A.J. and E.E.Kisker. 1986. Mortality crossovers: reality or bad data? *Population Studies*, 40: 389-401.

Dinkel, R.H. and E. Meinl. 1991. Die Komponenten der Bevölkerungsentwicklung in der Bundesrepublik Deutschland und der DDR zwischen 1950 und 1987. *Zeitschrift für Bevölkerungswissenschaft*, 17(2): 115-134.

Hill, M.E., Preston, S.H. and I.Rosenwaike. 2000. Age reporting among white Americans aged 85+: results of a record linkage study. *Demography*, 37(2): 175-186.

Kannisto, V. 1988. On the survival of centenarians and the span of life. *Population Studies,* 42(3): 389-406.

Kannisto, V. 1994. *Development of Oldest-Old Mortality, 1950-1990: Evidence from 28 Developed Countries.* Odense University Press, Odense.

Rosenwaike, I. 1979. A new evaluation of United States census data on the extreme aged. *Demography,* 16(2): 279-288.

Thatcher, A. R. 1999. The long-term pattern of adult mortality and the highest attained age. *Journal of the Royal Statistical Society Series A.*162:5-43.

Thatcher, A.R. 1992. Trends in numbers and mortality at high ages in England and Wales. *Population Studies*, 46(3): 411-426.

Thatcher, A.R. 2001. The demography of centenarians in England and Wales. *Population: An English Selection*, 13(1): 139-156.

Thatcher, A.R., Kannisto, V. and K.Andreev. 2002. The survivor ratio method for estimating numbers at high ages. *Demographic Research* 6(1): Volume 2, Article15. Available at http://www.demographic-research.org/.

Vallin, J. and F. Meslé. 2001. Tableau I-C-1: Population par sexe et âge (de 0 à 100 ans), au 1 janvier, de 1899 à 1998, avec deux estimations selon le territoire pour les années de changement de territoire [revised post-publication]. In: *Tables de mortalité françaises pour les XIXe et XXe siècles et projections pour le XXIe siècle*. Paris: Institut national d'études démographiques.

Vincent, P. 1951. La Mortalité des vieillards. *Population*. 6(2): 181-204.

Wilmoth, J.R. with the assistance of K. Andreev, M. Bubenheim, D. Jdanov, D. Philipov, V. Shkolnikov, and P. Vachon. (2002). *Methods Protocol for the*

Human Mortality Database. Retrieved 13 June 2004 (http://www.mortality.org/Public/Docs/MethodsProtocol.pdf).

Zakharov, S. 2001. Description of data used for Lexis database. Research note of April, 2001. Available at http://www.mortality.org/PrivRegistered/RUS/InputDB/RUScom.pdf