HMD: Data Quality Checklist

Country: Republic of Korea[1] (KOR) Range of Years Covered: 2003-2020

Reported Submitted by: Dana A. Glei* Date: 6/13/2022
* with assistance from Magali Barbieri, Jeeeun Kim, and Sooyoung Kim

Data Quality Charts at:
https://www.mortality.org/RunResults/KOR/CHECKS/HMDQC.html

Updated for 2019-20	**Date Completed**	**Comments**
Checks for Input DB		
• *Internal Errors*: Check that rows and columns add up correctly	6/13/2022	
• *Gross Errors*: Note any corrections or problems in published numbers in the "Notes" file	N/A	
Internal Consistency (InputDB vs. LexisDB)		
• Check outliers identified during procedures for splitting open age interval. [see Chart 4]	6/13/2022	I believe these outliers result from age heaping (e.g., Females in 2014 & 2015: too many deaths at age 100+ and too few at ages 93-97).[2]
• Check deaths and population estimates in LexisDB against the raw data in the InputDB [see Charts 5-8]	6/13/2022	Deaths match (except for differences resulting from redistribution of deaths of UNK age); our pop estimates match the official estimates very well.
Internal Plausibility		
• Inspect plots of death rates by age and time for any unusual patterns [see Charts 9-10]	6/13/2022	OK
• Check for evidence of age heaping [see Chart 11]	6/13/2022	Death rates for the cohorts of men and women aged 74 in 2020 (born in 1945-46) appear to be lower than successive cohorts.
• Check implied migration for unexpected patterns [see Chart 12]	6/13/2022	(see notes below under "Implied Net Migration")

[1] A.K.A. South Korea.
[2] Jeeeun Kim says, "Refer to the Excel file "/dir/0/danaglei/COMMONS/Countries/KOR/Lit/Lee.2007.Analyzing the death rate of Korea in 1994, 2014, 2015.xlsx". Lee Seok Min thinks that there are birth cohort effects on mortality rates by age. For example, in 2014, there were more deaths at age 91 compare to 90 or 92, because the population of age 91 is bigger than the other cohort."

Updated for 2019-20	**Date Completed**	**Comments**
External Plausibility		
• Check total annual (implied) migration and compare with external estimates if available [see Chart 13]	6/13/2022	HMD estimates of implied migration do not seem very consistent with KOSTAT estimates of international net migration, but it is not clear how KOSTAT is computing net migration (see notes below under "Implied Net Migration").
• Check e_0, e_{65}, and e_{80} compared to Sweden. Are estimates plausible? [see Chart 14]	6/13/2022	Prior to 2008, e_x was substantially lower in KOR compared with SWE (for both sexes at ages 0, 65 & 80), but values have been converging over time. Since 2013, e_0, e_{65}, and e_{80} among women in KOR has exceeded that for SWE. Among men, e_0, e_{65}, and e_{80} in KOR has converged with SWE.[3] (see also notes below under "External Plausibility - Compared with Other Countries").
• *External Comparisons:* Compare e_0 and q_0 estimates to some external source External source: KOSIS life tables (2003-2020); UNPD life tables for 5-year periods (2000-05,…2010-15)	6/13/2022	Our estimates for q_0 and e_0 since 2003 are similar to those produced by KOSIS and UNPD. (see notes below under "External Comparisons (With Other Estimates for South Korea)").
***IF DATA WAS UPDATED*: Old* vs. New**		
• Check ratio of "new" to "old" death rates. Investigate any suspicious cases (e.g., ratio <0.95 or >1.05). [see Chart 15]	6/13/2022	There are no notable differences in the adjusted mortality rates.
• *If any big differences between old and new death rates*: Check ratio of deaths and ratio of population estimates. What is the source of the differences? [see Charts 16 & 17]	6/13/2022	No discrepancies in the death counts, but there were some changes in the population estimates above age 80 (particularly for cohorts aged 90+ on 1/1/2019) probably because the most recent population estimates affect the results from extinct cohort and SR method.

* The "Old" data (https://www.mortality.org/RunResults/KOR_Old/STATS/) were based on the life tables run on 11/15/2019.

[3] Jeeeun Kim notes, "See attached file "Life tables in 2015", especially 'life expectancy at birth of major OECD member countries'." [File is stored in /hdir/0/danaglei/COMMONS/Countries/KOR/Lit/Life Tables in 2015 of Korea_KOSTAT.pdf]

Additional Notes

Contents

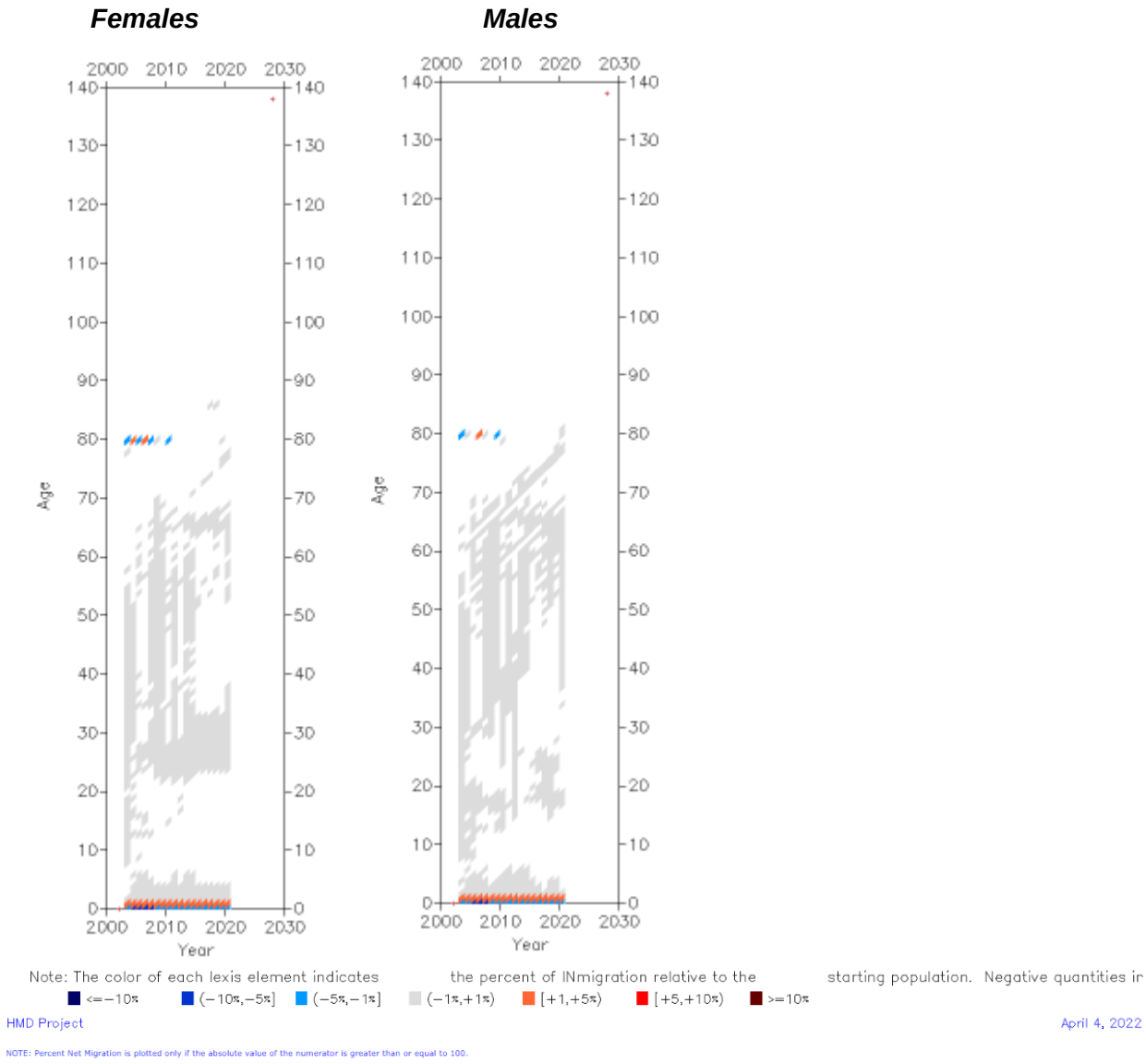
IMPLIED NET MIGRATION

Oddly, there appears to be evidence of out-migration of infants during their year of birth followed by in-migration for those same cohorts in the following year (Figure 1).

For example, the official statistics indicate that there were 227,270 girls born in 2010. There were 620 female deaths at age 0 in 2010, which the HMD methodology split into 493 deaths for the 2010 cohort (lower triangle) and 127 deaths for the 2009 cohort (upper triangle). Thus, we would expect to find 226,777 girls aged 0 on 12/31/2010 (227,270 births – 493 deaths)—assuming no in- or out-migration. The official population end-of-year population estimates indicate that there were 211,372 girls aged 0 on 12/31/2010 (approx. 16,000 fewer girls than expected), implying *out*-migration of infants between their birth in 2010 and December 31, 2010. Yet, KOSTAT's estimates of net international migration (see "/hdir/0/danaglei/COMMONS/Countries/KOR/RawData/Migration/NetMigration-by-sex-and-age-2003-2020.xlsx") show only -9,548 net out-migration among girls at aged 0 in 2010.

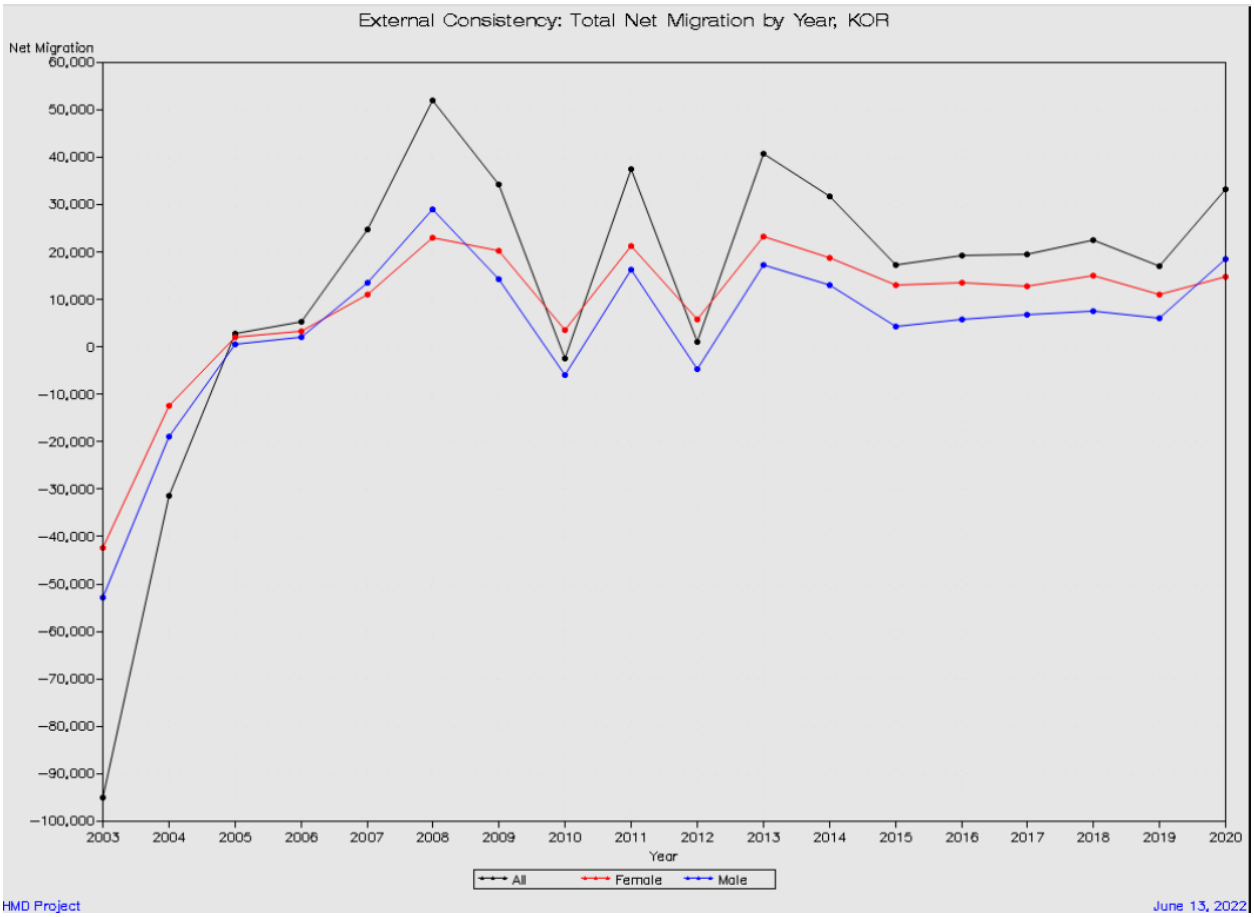
I have no idea how KOSTAT is computing their estimates of net migration, but they must not be computing it the way we would compute it. Perhaps the apparent out-migration of infants and in-migration of children aged 1 is a statistical artifact (e.g., differences in how KOSTAT and HMD compute population estimates for recent cohorts)?

Figure 1. Implied Net Migration, Estimates Run on 6/13/2022



In terms of overall net migration (all ages & sexes), KOSIS estimates[4] indicate out-migration of 57K-84K per year during 2003-2007, -37K in 2008, +21K in 2009, -15K in 2010, little net migration during 2011-14 (-7K to +5K), +10K in 2015, little migration in 2015-16 (-2K to +3K), and positive net migration in 2018 (+26K), 2019 (19K), and 2020 (+241K????).

In contrast, our estimates of implied migration are show below; they do not appear very consistent with those provided by KOSIS. For example, we find the highest in-migration in 2008 (+50K), whereas the KOSIS estimates suggested out-migration for that same year (-37K). It is not clear how KOSIS is computing net migration.



EXTERNAL PLAUSIBLITY (COMPARED WITH OTHER COUNTRIES)

Death Rates Compared with Other HMD Countries

When we compare the age-specific death rates for South Korea with those in Sweden, Japan, and the USA, we see that South Koreans exhibit higher mortality than Sweden or Japan at most ages prior to 2000 (Figures 2 and 3). Compared with the US, mortality before 2000 was higher primarily below age 15 and at older ages (40+ for men, 70+ for women).

By 2016, mortality in South Korea had fallen to levels similar to or below those in the US at virtually every age. Compared with Sweden, South Korean men exhibit lower mortality primarily in midlife (age 20-30), while their female counterparts have lower mortality than the Swedish

[4] See ~\My Data\HMD\KOR\RawData\Migration\NetMigration-by-sex-and-age-2003-2020.xlsx

primarily at ages 50-70. Mortality in KOR remains similar to or higher than in Japan at virtually all ages.

Figure 2. Comparison of death rates, South Korea vs. Sweden, Japan, and the USA, Men

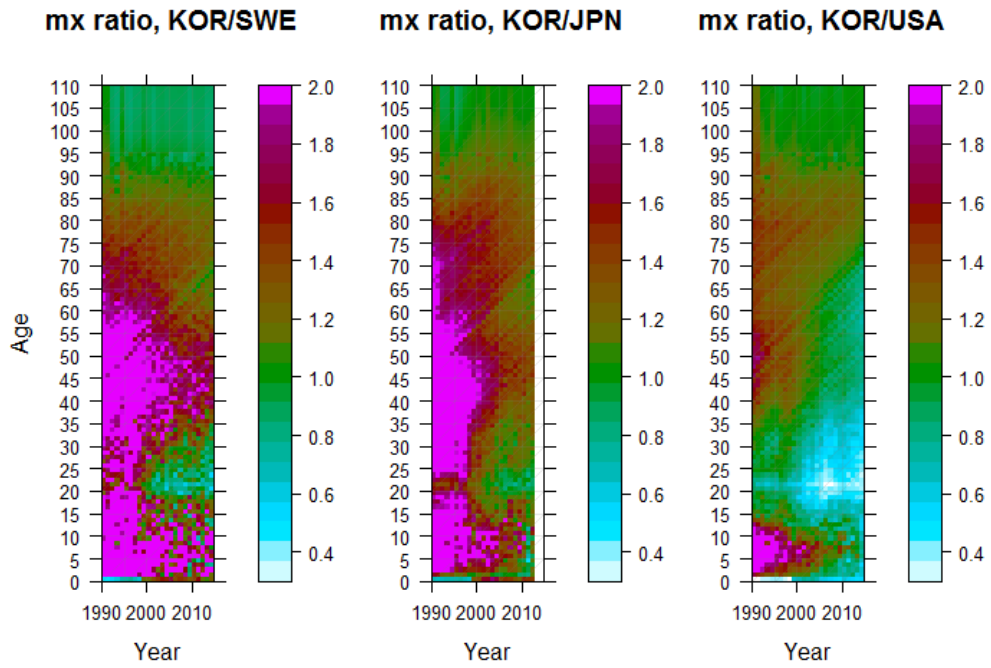
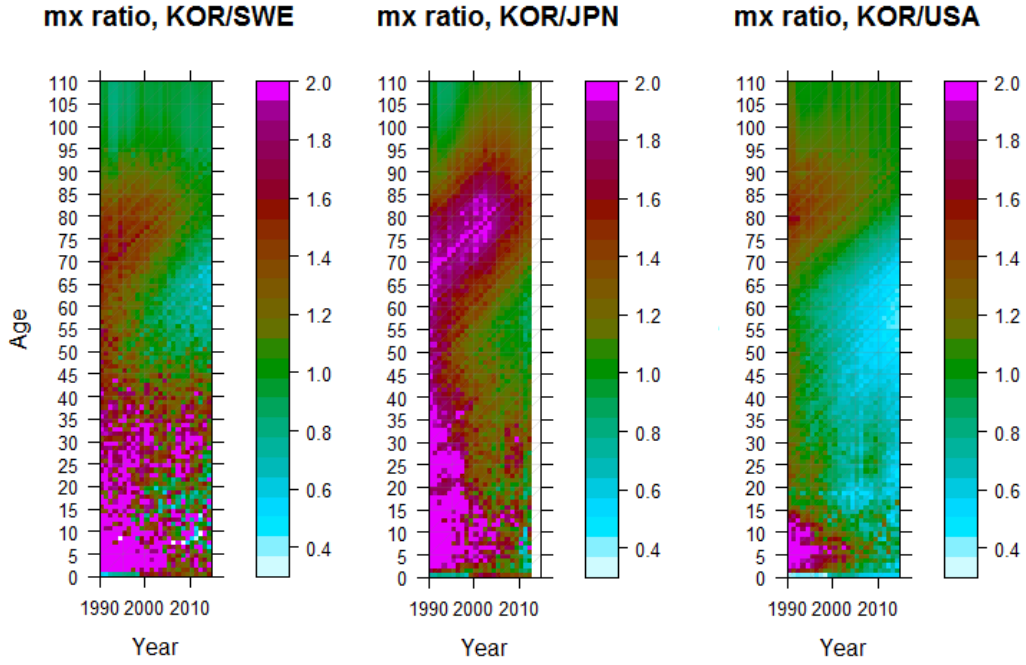


Figure 3. Comparison of death rates, South Korea vs. Sweden, Japan, and the USA, Women



Note: Figures 2 and 3 were produced using scripts provided by Dima (for instructions on running these scripts, see ~danaglei\COMMONS\MortX\DQ\From Dima\README.docx).

Mortality Pattern Compared with Other HMD Countries

To get a sense of the plausibility of the KOR mortality pattern, we compared some basic aggregate mortality indicators with the range of possibilities represented by all HMD country-years.

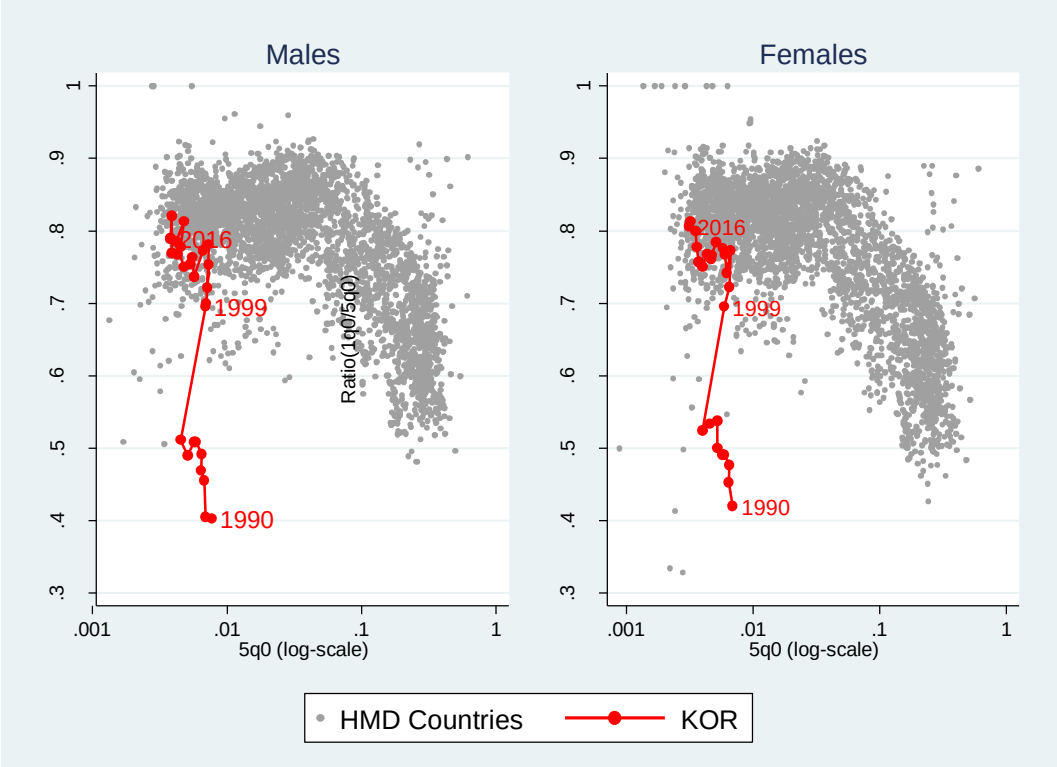
Figure 4 focuses on the relationship between infant and early childhood mortality. In the previous examination of KOR mortality (through 2005), infant mortality appeared to be under-estimated until at least the late 1990s. Some literature provides additional evidence of the severe under-reporting of infant deaths during those years. For instance, a study by Park (1998[5]) indicated that only 32% of infant deaths were registered in 1995. This is reflected in the suspiciously low ratio of infant to under-five mortality for 1990-98 (Figure 4). Between 1998 and 1999, this ratio jumped from 51% to 70% for males and from 53% to 69% for females (which resulted from a suspicious doubling of the infant mortality rate between 1998 & 1999, as noted later on p. 10; see also Figure 7).

The ratio of infant to under-five mortality continued to increase over time, albeit more slowly; by 2016, it had reached 79% and 81%, respectively. Though increased concentration of child mortality towards the early days of life is not implausible in a context of rapid survival progress, a ratio below 70% when early child mortality has reached low levels (${}_0q_1$<0.01) appears to be decidedly on the low side. Together with the analysis below (see "External comparison with other estimates for South Korea"), this finding suggests that it would be safer to begin the HMD

[5] Park, Kyung-Ae. (1998). *Recent trends and patterns of mortality in Korea.* Development and Society, Volume 27, Number 2.

series in the 2000, or even a couple of years later, as there is another jump in the ratio of $_1q_0$ to $_5q_0$ from 72% in 2002 to 78% in 2003 for males (74% to 77%, respectively, for females).

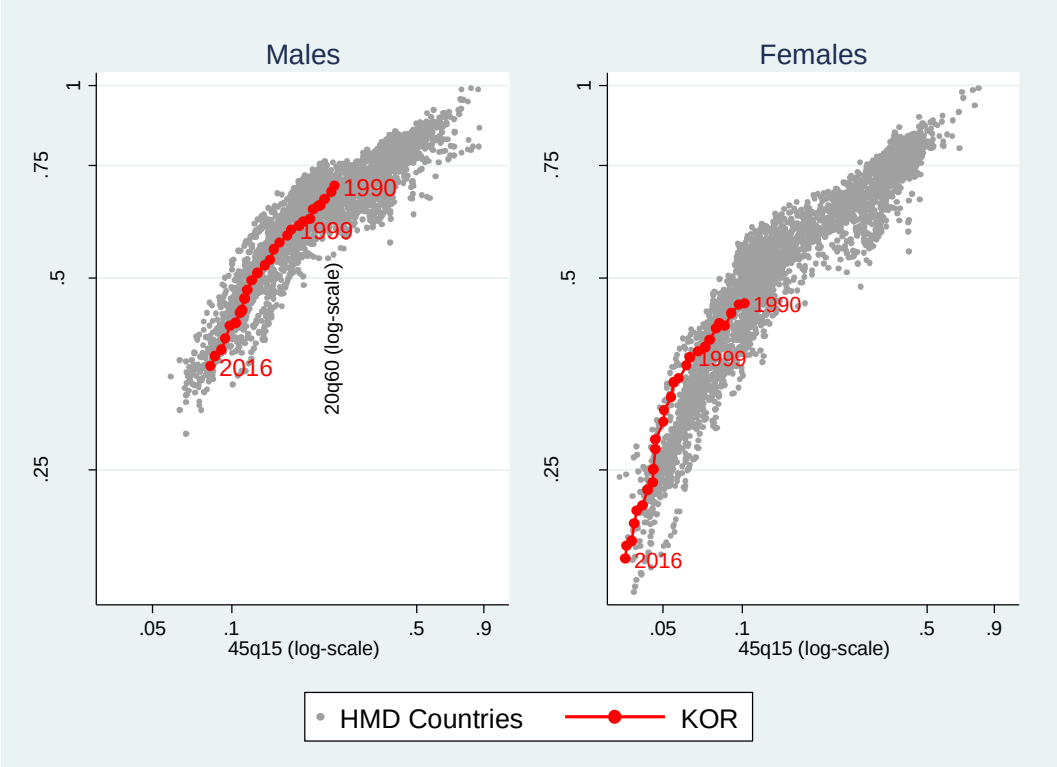
Figure 4. Ratio of infant ($_1q_0$) to under-five ($_5q_0$) mortality plotted by $_5q_0$, separately by sex, South Korea (1990-2016) versus HMD countries (excluding subpopulations)



Produced by ~\My Data\HMD\KOR_Old-2018-07-09\Do-Files\Chk_allHMD.do

Figure 5 compares mortality at older ages (60-80) relative to mortality in midlife (ages 15-60). Estimates for South Korea are within the range of defined by HMD countries, although in recent years women in South Korea fall towards the upper end of the range (i.e., suggesting high mortality at ages 60-80 relative to midlife mortality). It is not inconceivable that South Korean women might have low midlife mortality relative to older age mortality as fertility levels are very low (reducing the risk of maternal mortality), women in Korea are much less likely than men to smoke, and mortality from external causes may be low. Considering the rapid increase in life expectancy and differences between younger and older cohorts in terms of the mortality conditions to which they have been exposed (and in terms of women's status), this is not an unlikely scenario.

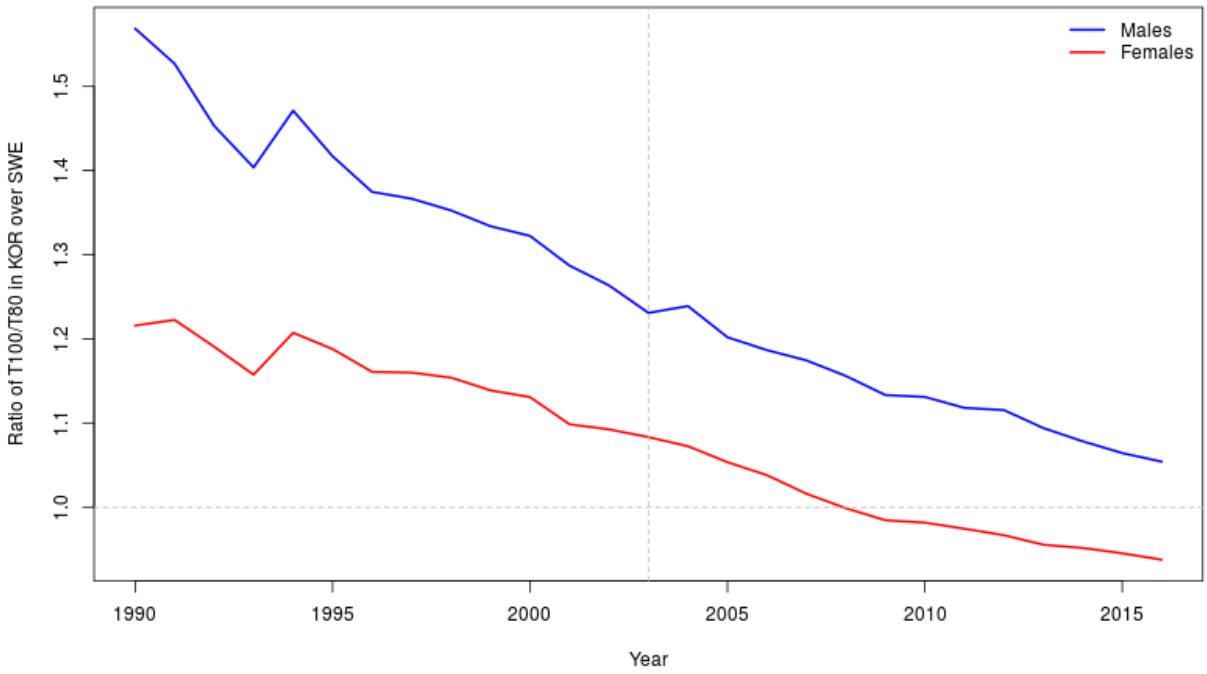
Figure 5. Probability of dying between ages 60 and 80 plotted by the probability of dying between ages 15 and 60, separately by sex, South Korea (1990-2016) versus HMD countries (excluding subpopulations)



Produced by ~\My Data\HMD\KOR_Old-2018-07-09\Chk_allHMD.do

Figure 6 shows the ratio T100/T80 in KOR relative to T100/T80 in Sweden. According to the paper by Jdanov et al. (2008), suspicions of age overstatement correspond to values of the ratio that are above 1.5. For all years since 2003, we are well below this level for KOR, which is another good sign that the data are of good quality.

Figure 6. Ratio T100/T80 for KOR relative to SWE



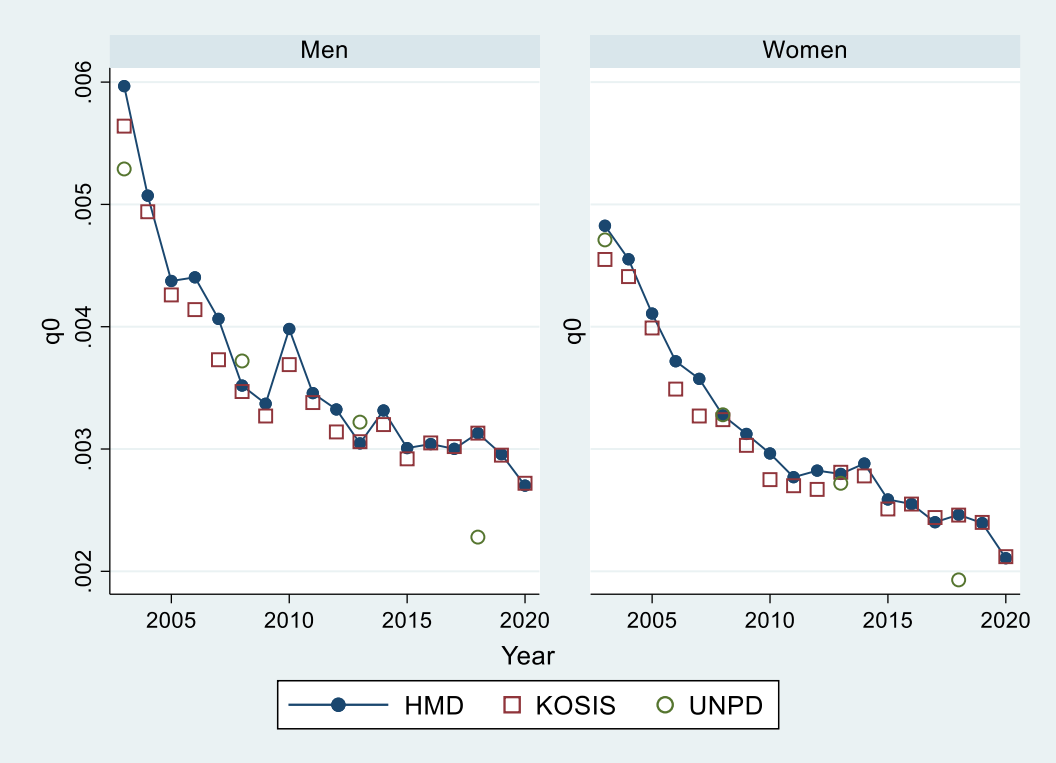
Produced by Magali (e-mail 1/4/2018)

EXTERNAL COMPARISONS (WITH OTHER ESTIMATES FOR SOUTH KOREA)

Probability of dying in the first year of life (q0)

As shown in Figure 7, most of our estimates of the probability of dying in the first year of life (q_0) are similar to those provided by KOSIS and UNPD after 2003. One exception is that our estimates are higher than the UNPD estimates for 2015-20. [Earlier analysis demonstrated that our estimates were much lower prior to 2000. Our estimates of infant mortality (as represented by q_0) approximately doubled between 1998 and 1999, increasing sharply from 0.0022 in 1998 to 0.0042 in 1999 for females and from 0.0024 to 0.0049, respectively, for males. The recorded death counts for age 0 also nearly doubled between 1998 and 1999 (from 656 to 1208 for girls; from 796 to 1556 for boys). This suspicious pattern suggests data quality problems (e.g., completeness of death coverage improving over time). Our (unadjusted) estimates prior to 2000 are much lower than those given by KOSIS and UNPD. Sooyoung Kim (KOSTAT) confirmed that they adjusted the infant mortality rate prior to 2000 based on the relationship between the death rate in the first year of life and the death rates at ages 1-4. As noted in the Background and Documentation, infant deaths were vastly under-reported prior to 1995 and are not considered reasonably complete until 2000. There is a close match between HMD and KOSIS estimates only since 2001.]

Figure 7. Comparison of q0 with external estimates, 2003-2020

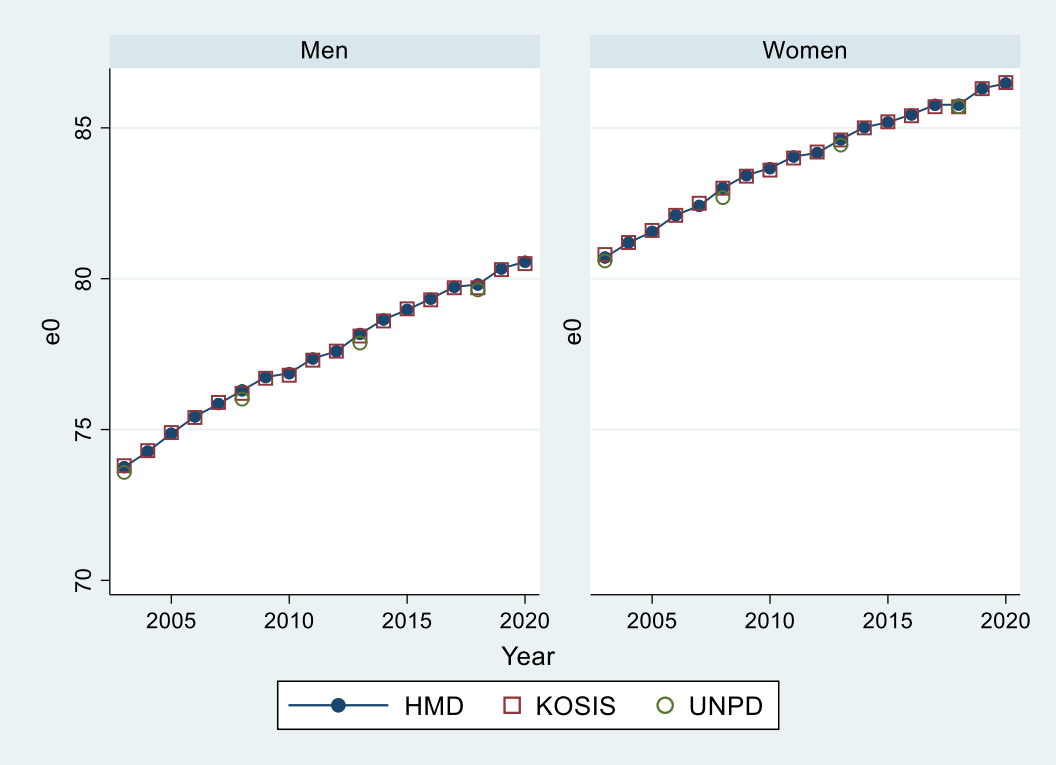


Produced by ~\HMD\KOR\Do-Files\Chk_LifeTables.do
Sources: KOSIS estimates (Statistics Korea, 2020 (file="~\Google Drive\HMD\RawData\Life Tables\KOSIS\KOR-Complete_life_tables_1970-2020.xlsx")); UNPD estimates (UNPD, 2019 (file="~\Google Drive\HMD\My Data\UN\Life Tables\UNPD-lifetable-estimates-2019.xlsx")).
Note: The UNPD (2019) estimates are for 5-year periods (2000-05,…2010-15) and have been plotted at approximately the mid-point of each period (2003,…2013).

Life expectancy at Birth (e0)

Figure 8 shows our estimates of life expectancy at birth (e_0) compared with estimates produced by KOSIS and UNPD. They are very similar. For example, in 2018 our estimate of e_0 among males was 79.8 compared with 79.7 from KOSIS (and 79.6 from UNPD for 2015-20).

Figure 8. Comparison of e0 with external estimates, 1980-2020

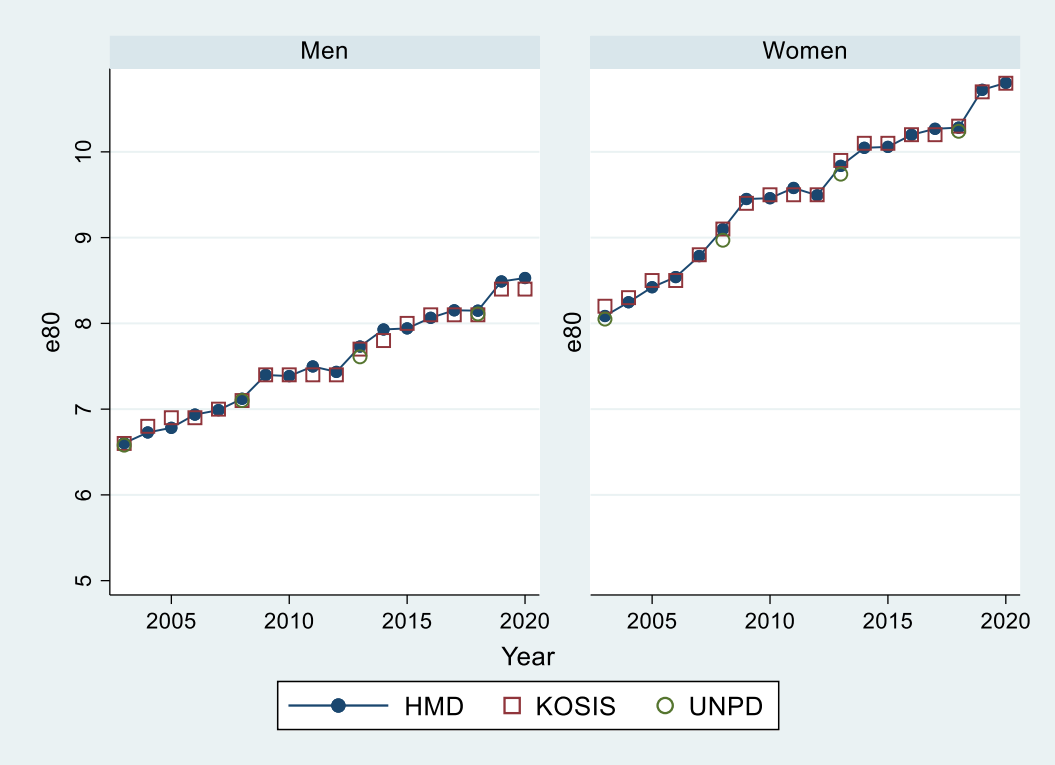


Produced by ~\HMD\KOR\Do-Files\Chk_LifeTables.do
Note: The estimates from KOSIS were downloaded 4/11/2022 from http://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B41&language=en&conn_path=I3 (file="~\Google Drive\HMD\RawData\Life Tables\KOSIS\KOR-Complete_life_tables_1970-2020.xlsx"). The UNPD (2019) estimates are for 5-year periods (1980-85,…2010-15) and have been plotted at approximately the mid-point of each period (1983, 1988,…2013); they were downloaded 5/4/2020 from https://esa.un.org/unpd/wpp/Download/Standard/Mortality/ (file="~\Google Drive\MortX\Data Sources\UN\Life Tables\UNPD-lifetable-estimates-2019.xlsx").

Life expectancy at Age 80 (e80)

Figure 9 shows a similar graph for life expectancy at age 80. Again, the HMD estimates closely match the estimates from KOSIS and UNPD.

Figure 9. Comparison of e80 with external estimates, 2003-2020



Produced by ~\HMD\KOR\Do-Files\Chk_LifeTables.do

Note: The estimates from KOSIS were downloaded 10/7/2019 from http://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B41&language=en&conn_path=I3 (file="~\Google Drive\HMD\RawData\Life Tables\KOSIS\KOR-Complete_life_tables_1970-2017.xlsx"). The UNPD (2017) estimates are for 5-year periods (1980-85,…2010-15) and have been plotted at approximately the mid-point of each period (1983, 1988,…2013); they were downloaded 6/28/2017 from https://esa.un.org/unpd/wpp/Download/Standard/Mortality/ (file="~\Google Drive\MortX\Data Sources\UN\Life Tables\UNPD-lifetable-estimates.xlsx").

REFERENCES

Jdanov, D. A., Jasilionis, D., Soroko, E. L., Rau, R., & Vaupel, J. W. (2008). *Beyond the Kannisto-Thatcher database on old age mortality: An assessment of data quality at advanced ages.* MPIDR working paper WP 2008-013. Rostock, Germany: Max Planck Institute for Demographic Research. Unpublished manuscript.

Statistics Korea. (2017). Life Tables by Province, 1970-2014. Retrieved 6 September 2017 (http://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B41&language=en&conn_path=I3).

United Nations, Department of Economic and Social Affairs, Population Division (2017). *World Population Prospects: The 2017 Revision*, DVD Edition. New York: United Nations. Retrieved 28 June 2017 (https://esa.un.org/unpd/wpp/Download/Standard/Mortality/).