Rijksuniversiteit Groningen - University of Groningen
Faculty of Spatial Sciences
Population Research Centre

Master Thesis

Investigating differences in Healthy Life Expectancy (HLE) between migrants and natives in Germany using Multistate Lifetables

Supervisors:

Dr Adrien Remund (University of Groningen)

Professor Dr Karsten Hank (University of Cologne)

Author:

Adrian Kunz

Double-Master's Programme "Demography and Social Inequality"

M.Sc. Population Studies

Student number: S5751977

E-Mail: a.kunz@student.rug.nl

Place, Date: Groningen, 10/07/2024

Acknowledgements

This work would not have been possible without the support of some people.

First, I would like to thank my two supervisors: Dr Adrien Remund (University of Groningen) and Professor Dr Karsten Hank (University of Cologne). Dr Adrien Remund shared my enthusiasm for the topic and the multistate methodology and made working on this topic possible. His supervision during the whole research process was very valuable and I am very thankful for collaborating with him. I am also very thankful for the supervision by Professor Dr Karsten Hank. Although supervision mainly took place at the University of Groningen, he gave very helpful comments for improving my thesis and facilitated the thesis process at the University of Cologne.

Second, I want to thank Dr Tim Riffe from the University of the Basque Country who gave very useful advice about multistate modelling, the decomposition method and their implementation in R. I am also thankful for the freely available R code facilitating my deeper entry into R. Without his support, the implementation of this project would have been much more difficult.

Besides the thesis, I am very grateful for the support of Dr Adrien Remund and Dr Billie de Haas during a short difficult time last autumn when I was dealing with a personal issue. Furthermore, I want to thank Dr Billie de Haas in the role of the coordinator of the master's programme in Population studies. As a highly engaged person, she facilitated many things in the organisation of the master. At the University of Cologne, I am very thankful to Xenia Elbrächter for managing everything around the double degree programme and for quickly answering questions.

Finally, I want to thank my parents and my brother who supported me throughout my entire lifetime. Without the support of my family, I would have never been so successful. From my parents, I learned my ambitious way of working helping me throughout my studies. My brother always supported me during my entire lifetime and our discussions always enriched my life.

I Abstract

A well-established demographic finding is that migrants tend to enjoy longer life expectancies (LE) than natives which is often referred to as the Migrant Mortality Advantage (MMA). However, it is also known that migrants tend to have a disadvantage in most health measures resulting in a shorter healthy life expectancy (HLE) than natives. This paradox has been confirmed in Belgium, the Netherlands, England and Wales, but all this research relies on the Sullivan method. This technique is not very data-demanding but can be biased if its assumptions are violated. To prevent these biases and investigate the causes of this paradox, I used multistate models. My research questions were to investigate to what extent LE and HLE at age 50 differ between migrants and natives, and how the different transitions between health statuses and death contribute to these differences. I used German data from the Survey of Health, Ageing and Retirement in Europe (SHARE) to model and predict transition probabilities between good and bad self-reported health (SRH) and death by age, sex and migration status. These probabilities were used to apply multistate models and obtain HLE and LE by sex and migration status. Finally, I applied a decomposition method to estimate the contribution of each transition to the differences in (healthy) life expectancies between migrants and natives. My results confirm that migrants in Germany enjoy a higher LE than natives (men: +0.74 years, women: +0.79 years) but suffer from a shorter HLE (men: -2.28 years, women: -2.78 years). The decompositions show that migrants' lower death probabilities, particularly for unhealthy people, explain migrants' higher LE while migrants' higher incidence and lower recovery probabilities mainly contribute to migrants' shorter HLE. Migrants' lower death probabilities cannot compensate for their disadvantage in HLE. This study supports previous research confirming migrants' health disadvantages, but future studies should investigate the causes of migrants' lower recovery and higher incidence probabilities such as specific diseases.

Keywords:

Migrant Mortality Advantage, Migrant Health, Healthy Life Expectancy, Multistate Models, Decomposition Method

Content

II Abbreviations

Destatis	Statistisches Bundesamt
dtms	discrete-time multistate models
LE	Life Expectancy
HLE	Healthy Life Expectancy
MMA	Migrant Mortality Advantage
na	missing value
sd	standard deviation
SES	socioeconomic status
SHARE	Survey of Health, Ageing and Retirement in Europe
SRH	self-rated health
ULE	Unhealthy Life Expectancy
vgam/VGAM	vector generalized linear and additive models
WHO	World Health Organization

III List of Figures and Tables

1 Introduction

There is a constant migration flow to Germany and Europe (Statistisches Bundesamt [Destatis], 2024b). Over the most recent years, peaks of this influx in Germany in the years 2015 and 2022 appeared due to refugees from the wars in Syria and Ukraine (Destatis, 2024b). The long-term integration of migrants is important because most migrants stay in Germany leading to an increasing stock of migrants (Destatis, 2024a).

When migrants age, maintaining good health and preventing diseases and chronic conditions becomes an increasingly important issue. The proportion of older people among migrants has already started to increase and is expected to further increase due to the general demographic development (Schimany et al., 2012, p. 208). Therefore, considering migrants' health is getting more and more relevant.

The International Organization for Migration's call for health for all in Europe meaning also the inclusion of migrants' health illustrates this need (International Organization for Migration, 2009). In the European context of increasingly diverse and ageing societies, knowledge about health and mortality patterns among migrants is necessary to better estimate the demand for healthcare systems (Rechel et al., 2013).

A well-established demographic finding is that migrants tend to enjoy longer life expectancies (LE) than natives which is often referred to as the Migrant Mortality Advantage (MMA) (Abraído-Lanza et al., 1999; Razum, 2008; Razum et al., 1998).

Studies found this effect for a diverse range of developed countries such as Australia (Huang et al., 2024, p. 9; Kouris-Blazos, 2002), Belgium (Patrick Deboosere & Sylvie Gadeyne, 2005), Canada (Bourbeau, 2002), France (Boulogne et al., 2012; Khlat & Courbage, 1996), Germany (Razum et al., 1998; zur Nieden & Sommer, 2016), the Netherlands (Bos et al., 2004), Switzerland (Tarnutzer & Bopp, 2012; Zufferey, 2014, 2016), the United Kingdom (Wallace & Kulu, 2014) and the United States (Ruiz et al., 2013; Zheng & Yu, 2022). If a migrant mortality advantage does not exist overall, there might be an advantage compared to native people in the same socioeconomic position (weaker version of the MMA) (Riosmena et al., 2013, p. 1041).

However, it is also known that migrants tend to have a disadvantage in most health measures. For instance, Solé-Auró and Crimmins (2008) found that migrants have worse health than native people regarding functional ability, disability, disease presence, self-rated health (SRH) and behavioural risk factors using data from the Survey of Health, Ageing and Retirement in Europe (SHARE) from eleven European countries.

A widely used measure combining life expectancy and morbidity is healthy life expectancy (HLE). This indicates the average number of years a person can expect to live in good health due to disease or injury depending on the definition used for health (World Health Organization [WHO], 2024).

The most common method to compute HLE is the not-very data-demanding Sullivan method. It was used to investigate differences in HLE at age 50 between migrants and natives in Belgium, the Netherlands and England and Wales (Reus-Pons et al., 2017). These studies found that older migrants could expect to live fewer years in good health than older native people in all three countries (Reus-Pons et al., 2017, pp. 533–535). This was particularly the case for non-Western migrants (Reus-Pons et al., 2017, p. 535).

However, the computation of HLE using the Sullivan method can be biased if the assumptions of no recovery, same death rates for healthy and unhealthy people and that the age-specific disability prevalence is constant over time (stationarity) are violated (Imai & Soneji, 2007). To prevent these biases and investigate further the causes of the paradox of migrants' longer total life expectancy but shorter healthy life expectancy, the use of multistate life tables/models is appropriate. These methods do not rely on these assumptions and allow a more comprehensive analysis of morbidity and mortality dynamics (Imai & Soneji, 2007). To the best of my knowledge, there is no study which investigated differences in HLE between migrants and non-migrants using the multistate approach.

Therefore, the research questions driving this study are:

To what extent do Life Expectancy (LE) and Healthy Life Expectancy (HLE) at age 50 differ between migrants and natives in Germany, and how do the different transitions between health statuses and death contribute to these differences?

The specific objectives are (1) to compute LE and HLE separately for migrants and natives and (2) to decompose the difference in HLE into the contributions of the transition probabilities between the different health states. The latter objective allows us to separate differences in mortality from differences in morbidity and to investigate the causes of differences in the life expectancies between migrants and natives.

The remainder of the paper is structured as follows: First, I will give an overview of Germany's history of immigration and describe existing explanations of the migrant mortality advantage and previous research. Then, I will explain the data and methods I used. Subsequently, I will

present my results, from describing the transition probabilities over an overview of the different life expectancies to the decomposition of the LE and HLE differences. Finally, I will discuss my findings and return to my research questions.

2 Theoretical Background and Previous Research

2.1 Background about Migration to Germany

Despite large numbers of emigrants, the Federal Republic of Germany's net migration has been positive in most years since 1950, making Germany a country of immigration (Destatis, 2024b). In 2023, 21.176 million people who either themselves or at least one parent were born without German citizenship (around a quarter of Germany's population) lived in Germany (Destatis, 2024a). Migration to Germany has taken place in four phases since 1950:

The first phase from the middle of the 1950s consisted of the recruited migration of guest workers, mainly from Turkey, Italy, Portugal, Greece and Yugoslavia (Steinbach, 2018, p. 287; van Mol & Valk, 2016, pp. 32–33). Despite the expectation that migrants would return to their home country, many of them stayed in Germany (Steinbach, 2018, p. 287; van Mol & Valk, 2016, p. 35).

A second phase from the oil crisis of 1973-74 to 1990 was characterized by the recruitment stop of guest workers and the immigration of guest workers' families (Steinbach, 2018, p. 287).

The third post-reunification phase consisted of the immigration of ethnic Germans (repatriates) coming from the USSR and its successor states, refugees and asylum seekers, particularly from Turkey and Yugoslavia or they were contingent refugees (Steinbach, 2018, p. 288).

The ongoing phase since 2010 has been shaped by immigrants from the new Eastern European Union member states and refugees from conflict regions such as Syria, Ukraine or Balkan countries (Steinbach, 2018, p. 288).

Due to this history, (older) migrants in Germany are a very heterogeneous group. Older migrants in Germany have less economic capital than natives, suffer from worse housing conditions (Steinbach, 2018, p. 300) and live in larger households (Steinbach, 2018, p. 296). The conditions of labour migrants are generally worse than those of the repatriates because of their different migration histories and legal statuses (Steinbach, 2018, p. 300). For instance, labour migrants often work in low-paid jobs while repatriates are better educated (Steinbach, 2018, p. 295). Furthermore, older migrants from Turkey and the former Yugoslavia live more often in areas with air pollution, noise nuisance or a lack of green spaces (Steinbach, 2018, p. 295). This heterogeneity might differently affect migrants' health.

2.2 Explanations of the Migrant Mortality Advantage

The demographic literature offers four different explanations for why migrants experience lower mortality than non-migrants in the host country in high-income countries.

2.2.1 The Healthy Immigrant Effect (in-migration selection effects)

The first hypothesis for the migrant mortality advantage is the "healthy immigrant effect". This means that migrants entering the destination country are positively selected regarding health compared to native people in the origin country (Guillot et al., 2018, p. 3). Following Chiswick et al. (2008), the selection effect is particularly relevant for people migrating due to educational or working reasons while it plays a minor role for family migrants and refugees.

Previous research showed mixed findings. Rubalcava et al. (2008) found only weak support for the hypothesis using the Mexican Family Life Survey about Mexicans in Mexico and Mexican migrants to the US before migration when looking at height, weight, blood pressure and general health. A comparable study by Bostean (2013) using the same Mexican data and the US National Health Study showed that Mexican migrants are positively selected in health limitations, negatively selected in self-rated health and not selected at all in chronic conditions. Using survey data collected at one specific point in time (cross-sectional) from the US, the UK, Canada and Australia, Kennedy et al. (2015) found evidence for the healthy immigrant effect relative to the origin population in self-reported health, chronic conditions, obesity and smoking even after controlling for age and education. This evidence seems stronger for migrants from developing countries (Kennedy et al., 2015).
For France, Ichou and Wallace (2019) used survey data and found the healthy immigrant effect in self-rated health, chronic diseases and health limitations among men but weaker for women, perhaps due to gendered socialisation and women migrating as accompanying spouses. For self-rated health, the authors even showed a migrant disadvantage among women. Furthermore, educational selectivity was a main contributor to the healthy immigrant effect.
Guillot et al. (2018) found evidence mostly consistent with the healthy immigrant explanation when investigating age variations of foreign-born versus native-born mortality ratios using data from France, the US and the UK. Remarkably, the age pattern of excess mortality at young ages, a large advantage at adult ages and mortality convergence with natives at oldest ages, looks similar for very different migrant populations in three different country and welfare state contexts (for instance Canadian-born and Mexican-born immigrants in the US or Tunisian- and UK-born migrants in France) (Guillot et al., 2018, 11).

If the healthy immigrant effect explains the MMA, the effect should be lower the longer the people remain in the country as most people arrive at young ages and as the frailty compositions

of both groups, migrants and non-migrants converge over time (Guillot et al., 2018, p. 4). Therefore, the difference in HLE at age 50 between migrants and natives might be small or inexistent even when a mortality advantage exists.

2.2.2 The Salmon Bias (out-migration selection effects)

The salmon bias as it is usually referred to in the literature implies that migrants in poor health return to their home country (Namer & Razum, 2016, p. 6). One reason is people's wish to die at their home place (Tezcan, 2019, pp. 7–8). Other reasons for unhealthy people returning to their home country might be the need for familiar support, the lower costs of living or more affordable healthcare (Arenas et al., 2015, p. 1856).
There might also be indirect effects when migrants from a lower socioeconomic origin which is related to higher mortality are more likely to return to their home country than other migrants (Guillot et al., 2023, p. 1337). As the people returning to their home country are less healthy, the migrants remaining in the countries are positively selected (Guillot et al., 2018, p. 4).

The literature about selection effects during return migration is inconclusive. One of the first studies about the MMA showed lower mortality among US-born Latinos relative to US-born non-Latino Whites making the salmon bias implausible (Abraído-Lanza et al., 1999, p. 1546). Particularly lower mortality among Cubans and Puerto Ricans contradicts the Salmon bias because a return to Cuba had remained unappealing due to the bad political situation and Puerto Rico's deaths are recorded in the National Death Index (Abraído-Lanza et al., 1999, p. 1546).
In contrast, results from another study from the US showed that migrants returning from the US to Mexico have worse health than migrants staying in the US (Arenas et al., 2015). This can be seen as evidence for the Salmon Bias (Arenas et al., 2015). In a similar study comparing Mexican return migrants to their counterparts staying in the US, Riosmena et al. (2013) found evidence consistent with the salmon bias in hypertension, smoking, self-rated health and height. Another study from the US about Mexican return migrants showed a higher return propensity for migrants reporting health limitations, stress or sadness but the salmon bias was not shown for self-rated health and chronic conditions (Diaz et al., 2016). However, compared to the previous studies, this research only used cross-sectional data (Diaz et al., 2016). Therefore, the return migrants were not observed before returning to Mexico (Diaz et al., 2016).
The varying results regarding the Salmon Bias of the study by Abraído-Lanza et al. (1999) and the research on return migration might be attributed to the different designs and measures (mortality versus health). The return migration studies relied on studies that repeatedly collect

data about the same respondents (panel/longitudinal) or cross-sectional survey studies while Abraído-Lanza et al. (1999) used panel surveys linked with official mortality data from the National Death Index.

For the UK context, Wallace and Kulu (2018) found the Salmon bias in three foreign-born populations, namely India, Pakistan and Bangladesh, and the Caribbean, but not for other populations when comparing migrants' mortality to natives'. However, the Salmon bias does not fully explain the MMA among these populations (Wallace & Kulu, 2018). Another study in the UK found substantial mortality differences at young adult ages when comparing the mortality age profile of immigrants to their counterparts staying in their country of origin (Wallace & Wilson, 2019). If the salmon bias explained the MMA, the mortality differences at young adult ages should be low (Wallace & Wilson, 2019). The mortality differences should increase over time as older people are more likely to be unhealthy and return to their home country (Wallace & Wilson, 2019). While the former study used longitudinal data linking census with event data, the latter study investigated the mortality differences based on macro-level data (Wallace & Kulu, 2018; Wallace & Wilson, 2019).

Little research outside the Anglo-Liberal context has been conducted so far making more research necessary.

However, when comparing age variations in foreign-born vs. native-born mortality ratios using macro-level mortality data from the US, UK and France, Guillot et al. (2018) found evidence inconsistent with the salmon bias as the risk ratio increases after age 45. This is implausible as among older migrants, there is a higher risk pool for unhealthy return migration leading to lower relative mortality (Guillot et al., 2018, p. 4).

If return migration effects were the main explanation for the MMA, migrants should both live longer and healthier as people in bad health with a higher death probability return to their country. The salmon bias could be stronger at younger ages if recently arrived migrants rely on family support in their home country. However, the older people are, the less healthy they are. This means that the people returning to the country should be rather old. Therefore, the health gap between return migrants and migrants should be especially pronounced among older ages leading to the assumption that the HLE at age 50 might substantially differ between migrants and natives in terms of both mortality and health.

2.2.3 The Cultural Explanation

The third explanation of the MMA refers to a healthier lifestyle/cultural behaviour caused by different cultural norms in the home country than that of the country of destination (Darmon & Khlat, 2001). In their review, Darmon and Khlat (2001) showed for instance that the diet of Mediterranean migrants living in France partly explains the MMA and low rates of chronic diseases. A second part of the cultural explanation is that migrants tend to benefit from dense social support networks (Palloni & Arias, 2004). These networks work for instance as a cultural buffer against the risk of coronary heart disease (Palloni & Arias, 2004).

When comparing the mortality age profile of migrants to natives, Guillot et al. (2018) found an age pattern of higher migrants' mortality in the youngest ages, a mortality advantage in adult ages and a convergence towards natives' mortality in the oldest ages. Besides the healthy immigrant effect, this could also be consistent with different cultural behaviours as they are most relevant in adults ages around migration (Guillot et al., 2018, p. 11). However, the cultural thesis is implausible because the same age pattern was found among diverse migrant groups that differ in norms regarding their health behaviours (Guillot et al., 2018).

The study by Kennedy et al. (2015) investigating the healthy immigrant thesis as discussed above found this selection effect also for culturally closer migrants from developed countries (for instance US American migrants in Canada) making the cultural explanation implausible.

If the cultural explanation holds, this means that migrants are healthier, particularly in the ages of migration. After this, they might assimilate and incorporate the cultural practices of their destination country the longer they stay there. Therefore, the health advantage might decrease over age.

As chronic conditions need time to develop their effect on mortality due to the long latency time (changing to an unhealthier behaviour), mortality might be less affected than health by assimilation into the lifestyle behaviours of the host society (Law & Wald, 1999). Due to this longer latency time, the effects of healthier behaviours might appear stronger at older ages. On the other hand, the assimilation into the host society's unhealthier behaviours might compensate for this. This would imply that, similarly to the healthy immigrant effect, the difference in healthy life expectancy between migrants and natives might be small or even inexistent as the health advantage gets compensated for by assimilation at older ages. The overall life expectancy (living unhealthy and healthy combined) might be higher among migrants due to the time lag effect of chronic diseases.

2.2.4 Data Artefacts

The last explanation frequently mentioned by the literature is data artefacts (Guillot et al., 2018, pp. 5–6; Markides & Eschbach, 2011, pp. 228–229; Palloni & Arias, 2004, p. 387). Three issues arise around the discussion of data artefacts: coverage of deaths, coverage of the population and age misreporting (Guillot et al., 2018, p. 6; Palloni & Arias, 2004, pp. 387–388).

The coverage of deaths is affected by a mismatch between the numerator and the denominator of mortality rates (Guillot et al., 2018, p. 6). Person-years of exposure are usually based on a "de jure" definition whereas deaths are usually counted based on a "de facto" definition (Guillot et al., 2018, p. 6). Therefore, the count of deaths excludes deaths of residents occurring outside the boundaries and includes deaths of non-residents occurring within the boundaries (Guillot et al., 2018, p. 6). While it is less relevant for the natives, it might produce problems regarding the coverage of foreign-born people (Guillot et al., 2018, p. 6). People spending a substantial proportion of their time abroad might experience their death outside the country reducing the counts of death in the destination country (Guillot et al., 2018, p. 6).

For the second reason of artefacts, people spending a substantial amount of time outside the country might be also more likely to be undercounted in a census (Guillot et al., 2018, p. 7). Regarding age misreporting, it needs to be noted that information about their date of birth for migrants from less-developed countries often lacks, leading to age misreporting on census records and death certificates (Guillot et al., 2018, p. 7). Research has shown that certain groups of migrants tend to exaggerate their ages, particularly people of older ages (Palloni & Arias, 2004, pp. 387–388). This misreporting leads to an underestimation of mortality rates, especially for the elderly (Guillot et al., 2018, p. 7; Palloni & Arias, 2004, p. 388).

As the processes regarding data artifacts are rather complex and work in different directions it is difficult to hypothesize how it might affect HLE. Guillot et al. (2018) argue that errors in death coverage are more consequential than errors in population coverage (Guillot et al., 2018, p. 7). If migrants spend more time abroad as they age, that implies that their deaths are more likely to take place abroad making these migrants invisible if their death is not recorded in their host country. This leads to a higher life expectancy among migrants due to the underestimation of mortality. For migrants spending less time abroad when ageing, it might be the opposite.

Regarding how it affects the expectation about living healthy, it is difficult to assess as it depends on the selectivity of the migrants who spend a larger proportion of time abroad. As written above, age misreporting might lead to an underestimation of mortality. If people

overstate their age, time lived in poor health is assigned to an older age than they have. This might decrease the time spent in poor health up to this age increasing the computed time people can expect to live healthy.

2.3 Previous Research

Research from different country contexts has investigated health and mortality differences between migrants and natives. Table 1 gives an overview of the discussed studies.
Cross-sectional research has investigated health differences between migrants and natives.
This research has shown that (older) migrants compared to natives have worse self-rated health in several European countries (Lanari & Bussini, 2012; Reus-Pons et al., 2017; Solé-Auró & Crimmins, 2008). According to Lanari and Bussini (2012) this gap was particularly relevant for Eastern European migrants in France, Germany and Sweden. Older migrants in Europe suffer from more chronic conditions in Germany, Spain, Sweden and Switzerland, but from fewer in Austria (Solé-Auró & Crimmins, 2008). They have more limitations in functioning as shown for Turkish migrants in Germany (Carnein et al., 2015) and diverse migrant groups in France, Germany, Netherlands, Sweden and Switzerland (Solé-Auró & Crimmins, 2008).A higher prevalence of depression (Aichberger et al., 2010; Lanari & Bussini, 2012) was found in several European countries.

Regarding health trajectory after migration, studies using longitudinal survey data found a deterioration in self-rated health (SRH) in Canada, the US, and several European countries (De Maio & Kemp, 2010; Gubernskaya, 2015; Kim et al., 2013; Newbold, 2005; Reus-Pons et al., 2018, p. 7), in depression for Western migrants and diabetes for non-Western migrants in several European countries (Reus-Pons et al., 2018, pp. 8–9). In Canada, the deterioration in SRH was particularly pronounced for women and ethnic minorities (Kim et al., 2013).

Table 1: Previous studies about morbidity and mortality differences between migrants and natives (own illustration)

Authors	Data	Period of analysis	Host country	Outcomes
Aichberger et al. (2010)	1st SHARE wave	2004-2005	Denmark, Sweden, Austria, France, Germany, Switzerland, Belgium, Netherlands, Spain, Italy, Greece	Higher *depression* prevalence among migrants than natives, migration status effect stronger in Western and Northern than in Southern Europe
Carnein et al. (2015)	Official Data and Generations and Gender Survey	2004-2005 and 2005-2006 for Turkish Foreigners	Germany	Turkish foreigners live a longer proportion of their remaining LE with *health limitations* but have a longer LE
De Maio and Kemp (2010)	Longitudinal Survey of Immigrants to Canada	2001-2005	Canada	Deterioration in *SRH* and *emotional problems* for migrants after arrival
Garcia and Chiu (2016)	Hispanic Established Populations for the Epidemiologic Study of the Elderly	1993-2013	United States	Female Hispanic migrants with shorter time living without *health limitations*, mid- and late-life male migrants with an advantage in *health-limitations*-free life expectancy

Gubernskaya (2015)	Health and Retirement Study	1992-2008	United States	At age 50, migrants with better *SRH*, health deterioration afterwards
Huang et al. (2024)	Different datasets from the Australian Bureau of Statistics	2016	Australia	Migrants with longer LE but shorter or similar HLE based on *health limitations* than natives after age 65, young adult migrants with a larger advantage in LE and HLE than younger and older migrants which decreases afterwards
Kim et al. (2013)	Longitudinal Survey of Immigrants to Canada	2001-2005	Canada	At arrival, only 3.5 % of the migrants in bad *SRH*, decline afterwards (particularly for women and ethnic minorities)
Lanari and Bussini (2012)	1st SHARE wave	2004-2005	Austria, Belgium, Denmark, France, Germany, Sweden, Switzerland, Netherlands	Most Migrant groups with worse *SRH* (particularly Eastern European migrants in Germany, Sweden and France) and with higher *depression* prevalence
Newbold (2005)	National Population Health Survey	1994-2001	Canada	*SRH* status not different between migrants and natives,

	(Statistics Canada)			but migrants with higher health deterioration risk
Reus-Pons et al. (2017)	Diverse datasets from the statistical offices of the countries	2001 (Belgium), 2001 and 2011 (all other countries)	Belgium, Netherlands, England, Wales	In all countries, migrants with shorter HLE based on *SRH*, general health rather than mortality explained these differences
Reus-Pons et al. (2018)	SHARE	2004-2015	Austria, Belgium, Denmark, France, Germany, Italy, the Netherlands, Spain, Sweden, Switzerland	Older migrants with higher *SRH* deterioration risk than natives, Western migrants with higher risk of becoming *depressed*, non-Western migrants with higher risk of developing *diabetes*
Solé-Auró and Crimmins (2008)	1st SHARE wave	2004-2005	Austria, Belgium, Denmark, France, Germany, Greece, Italy, Netherlands, Spain, Sweden, Switzerland	Migrants are in worse *SRH* and suffer from more health limitations and more chronic conditions in Germany, Spain, Sweden and Switzerland, fewer chronic conditions in Austria

Some studies have investigated differences in healthy life expectancy between migrants and non-migrants (Huang et al., 2024; Reus-Pons et al., 2017). Although migrants can expect to live longer than natives in general, the time spent in good health (measured as SRH and health limitations) is shorter compared to natives in Australia, Belgium, Netherlands, England and Wales (Huang et al., 2024; Reus-Pons et al., 2017). The difference in HLE between migrants and natives was mainly explained by self-rated health rather than mortality (Reus-Pons et al., 2017). Carnein et al. (2015) showed that Turkish migrants in Germany suffer from shorter expected time without health limitations

Another study found the HLE difference only for Hispanic women in the US using a measure of self-reported disability while Hispanic mid- and late-life male migrants even experienced an advantage in time spent in healthy conditions (Garcia & Chiu, 2016). However, these studies used the easy computable Sullivan method for HLE while there does not exist any study using the multistate models.

3 Data and Methods

3.1 Data

For the analysis, I draw on data from the Survey of Health, Ageing and Retirement in Europe (SHARE)[1]. SHARE is a comprehensive panel study collecting information about health, socioeconomic status and support networks for people aged 50 and above every second year in several European countries, Cyprus, and Israel. Up to 2019, data has been available from about 140,000 respondents in more than 375,000 interviews in 7 waves (Bergmann, Michael: Kneip, Thorsten et al., 2019, p. 6). The last and eighth wave was administered in 2020.

I use data from the four most recent consecutive waves unaffected by COVID-19, namely waves 4, 5, 6 and 7. Waves 8 and 9 are partially affected by COVID-19 and might be biased. For my analysis, I only used German data. The reason for choosing only one country was that it is easier to work on a homogeneous population rather than having to deal with different social, economic, political and welfare contexts. Additionally, I needed separate life tables for foreigners and natives to calibrate the death information. As there are no such life tables for all SHARE countries, I need this restriction of my analysis.

[1] An alternative data source would have been data from the German Socio-Economic Panel Study (GSOEP) which also includes information about self-rated health, migration status, sex and death. However, the first idea was to conduct the same analysis for the full sample of SHARE countries. Due to my mortality correction procedure (see section 4.3.3), I switched to an analysis of Germany.

Table 2: Sample size and size of the foreign-born sample in SHARE data, wave 1 [Source: Survey of Health, Ageing and Retirement in Europe (2024b)]

Country	Sample Size	Foreign-born people
Austria	1561	135 (8.6 %)
Germany	2993	555 (18.5 %)
Sweden	3049	259 (8.5 %)
Netherlands	2958	186 (6.3 %)
Spain	2315	54 (2.3 %)
Italy	2551	38 (1.5 %)
France	3098	482 (15.6 %)
Denmark	1704	65 (3.8 %)
Greece	2897	69 (2.4 %)
Switzerland	996	169 (17.0 %)
Belgium	3808	265 (7.0 %)
Israel	2433	1335 (54.9 %)

To choose the country, I looked at two aspects: the sample size and the size of the migrant sample. First, I needed a sample that is large enough to compute and predict all transition probabilities for all age groups separately. Secondly, I needed a foreign-born sample that is large enough to allow the reliable computation of the transition probabilities for migrants.

Table 1 shows the sample size and size of the migrant sample for the initial samples of all countries available in wave 1. Similar proportions of migrants emerge for the refreshment samples. There are several countries with a sample size larger than 3000. However, most countries only include a small number of foreign-born people. Countries with a large migrant sample are France, Germany, Israel and Switzerland. While Switzerland and Israel are special cases (highly diverse societies and the country-specific history in Israel) with a relatively small sample, France and Germany are the only countries with a large sample size and a large share of migrants. Due to the similar sample size in both countries, I chose the country with the larger migrant sample (Germany). The German refreshment samples in waves 2 and 5 include a similarly high proportion of foreign-born people (15.3 and 13.5 per cent) (SHARE-ERIC, 2024c, 2024d).

First, I started to use only German data from 6 and 7. Waves 4 and 5 were later included to increase the sample size. Therefore, I looked at transitions between waves 4 and 5, as well as

between waves 6 and 7. Transitions for people that are present in both wave combinations are considered independent observations. I do not include transitions between waves 5 and 6.

3.2 Measures

For the multistate models, I assumed three distinct health states: a healthy state, an unhealthy state, and death (see Figure 1). While people can change between the unhealthy and the healthy state, death is an absorbing state which cannot be left. The main decision was to choose the health measure as it may influence how people distribute among the health states. These need to be distinct and mutually exclusive so that no observation can be in two states at the same time.

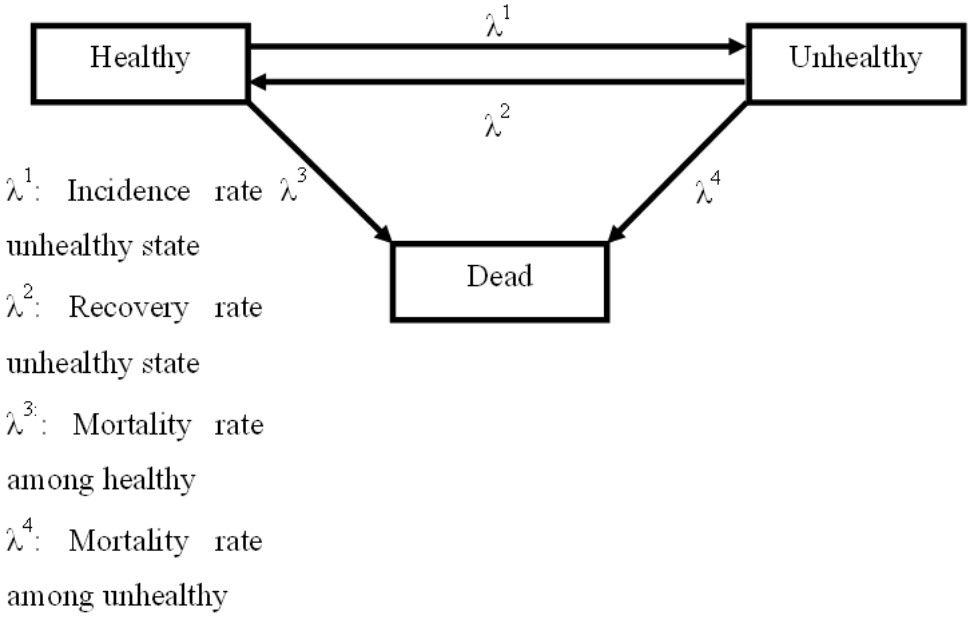

Figure 1: Multistate model with three states (Own illustration)

I used self-rated health (SRH) instead of other measures such as a self-report of limitations in daily activities or long-term diseases for the following reasons. While self-reported health measures the general health status in a relatively broad term, limitations in activities and long-term diseases only measure a specific severe health deterioration. Self-reported health is a well-known measure in public health and demographic research which showed a high degree of validity (in terms of construct, discriminant and concurrent validity) (Baćak & Ólafsdóttir, 2017; Cullati et al., 2018; DeSalvo, Fisher, et al., 2006) and reliability (Boardman, 2006; Cox et al., 2009). Its predictive power has been shown for mortality (DeSalvo, Bloser, et al., 2006),

use of health services (Halford et al., 2012) and healthcare expenditure (DeSalvo et al., 2009; Halford et al., 2012).
A problem which could appear when using it for a comparison between natives and migrants is comparability across different ethnic /cultural groups. According to Seo et al. (2014), variations in self-rated response patterns are not sensitive to migrant origin but to the survey language. As SHARE is conducted in the respective national language of the host country, this might be only a problem if people have a different understanding of a word across different cultures. Studies using other, more objective measures like functional ability, disability and disease presence found similar results (worse health among immigrants) (Solé-Auró & Crimmins, 2008). Chandola and Jenkinson (2000) found that SRH is related to morbidity in all ethnic groups and McGee et al. (1999) found that SRH predicts subsequent mortality relatively well for different ethnic groups.

Self-rated health (SRH) in SHARE was assessed with the question "Would you say your health is ...." on a five-point Likert scale (1- Excellent, 2 – Very good, 3 – Good, 4 – Fair, 5 – Poor). To define health states, I rely on the findings by Perneger et al. (2013) who showed that the largest gap is between good and fair health. Therefore, I defined the state as healthy if people stated their health as excellent, very good or good and as unhealthy if people stated a fair or poor health state. A second reason to assign fair to the unhealthy state is that it increases the number of observations in an unhealthy state.

Death is measured with end-of-life interviews with proxy people. As there are not always proxy people of dead people, the number of deaths is underestimated. A second reason for the underestimation of death is panel attrition. The longer a panel study runs, the more people get lost due to non-responding. Therefore, the proportion of people with unknown vital status in SHARE is substantial. According to my investigations of the original death data in the data preparation file, about 27.70 per cent of the panel respondents in wave 5 and 39.67 per cent of the respondents in wave 7 have an unknown vital status leading to an underestimation of death by a factor of around these values. These cases might be negatively selected, unhealthier people as they do not have a contact person. As I do not know an acknowledged method to deal with these unknown vital statuses, I excluded them from my analyses.

I define people as migrants if they were born outside of Germany. This might be problematic in cases when people were born abroad as descendants of Germans. However, the number of

such cases is likely very low. This operationalization excludes descendants of first-generation migrants and does not allow the investigation of differences in HLE between different generations of migrants. It was also not possible to distinguish between countries of origin or at least country groups of origin. The reason is that the sample size of migrants is already low when using the foreign-born status. A more fine-grained operationalization would make the computation even more difficult as the number of observations per origin for each group (distinguished by sex and age group) would be too small. Unfortunately, there is no better measure for migration. For sex, I rely on biological sex (male or female). Age is computed using the survey and the birth years (Age = survey year – birth year). I rely on two-year age groups from 50 up to age 100+. This grouping allows a more fine-grained analysis than five-year age groups without the problem of an insufficient number of observations per group as is the case with single age groups.

3.3 Methods

The approach to compute the different life expectancies is multistate modelling. This includes the computation and modelling of transition probabilities between the different health states as the first step. These probabilities are then used as input for multistate models computing the state-specific life expectancies. I then use a specific decomposition method to decompose differences in (healthy) life expectancies into the transition probabilities between all states by age. I explain in detail in sections 3.3.1-3.3.4 how I processed the data and the multistate model and decomposition methods I used to answer my research questions.

For my computations, I used R 4.4.0 / RStudio 2024.04.1+748. All my data preparations and analysis steps are available as R Markdown files and corresponding HTML files. All files can be executed from "0_Master_File.Rmd". Besides base R functions, I used the R packages VGAM (Yee, 2015), tidyverse (Wickham et al., 2019), dtms (Dudel, 2021; Dudel & Li, 2024), MortalitySmooth (Carmada & Riffe, 2015), flextable (Gohel & Skintzos, 2024) and gridExtra (Auguie, 2017).

3.3.1 Data Preparation and Sample Selection

The first step was the data preparation and sample selection (see markdown file 1*_Data_preparation.Rmd*). To this end, I created a full data set including all necessary variables for all people observed at two points in time. As I regard observations in waves 4 and 5 as

independent from observations in waves 6 and 7 to increase the sample size of transitions, the two points in time refer to 4 and 5 for the first sub-sample of transitions and 6 and 7 for the second sub-sample.

First, I merged the different files (physical health, death and demographic information) separately for waves 4, 5 and 6, 7. This results in two data sets, one for waves 4 and 5 and one for 6 and 7.

Then, I added a variable (*source*) to both data sets indicating whether the observation was from waves 4 and 5 or 6 and 7. The variable can take the values "45" and "67". This variable allows me to consider observations from 4, 5 and 6, 7 separately. For example, I can group by the *mergeid* and *source* to compute the destination health state (*disab_to*) which is the lead variable of *disab_from* (meaning the health state of the second wave).

After that, I bound both data sets together in one data set that includes the data from 6, and 7 below the data from 4, and 5. This results in 12,060 person-years. This long dataset includes two rows per case, one for each wave (waves 4 and 5 for 45 and waves 6 and 7 for 67).

Missing values on the self-rated health variable were then recoded as missings (na) and missing survey years were completed. For sample selection, I deleted all people who were still alive in each second wave (5 and 7) with a missing on the health variable. Due to that, 799 person-years were lost. All person-years without any information about the foreign-born status could not be used and were thus deleted. This resulted in an additional 154 deleted person-years. Then, person-years with missing self-rated health were only kept for dead observations. Therefore, an additional 311 person-years were deleted. Some people were younger than 50 in each first wave of observation (waves 4 and 6). As SHARE is only conducted for people aged 50 and above, I also deleted these observations. This resulted in an additional 112 person-years lost. The final sample selection includes the deletion of cases that only have one full observation. The reason for that is I needed information about the origin and the destination health state. An additional 1,060 person-years were thus deleted.

All in all, 2,436 person-years were deleted, and the analysis sample size is therefore 9,624 person-years. Table 3 gives a descriptive overview of the sample before and after sample selection. After selection, around 13.15 per cent (1,266) of these person-years belong to foreign-born people. 47.59 per cent (4,580) belong to men and 52.41 per cent (5,044) to women. 22.78 per cent are person-years of transitions between waves 4 and 5. Most person-years belong to transitions between waves 6 and 7 (77.22 per cent). On average, the person-years are 68 years old with a standard deviation (sd) of around 9.41 years. The proportions and mean-age of the analysis sample are similar to the values before selection. Around 1.4 per cent of the initial

sample had a missing on migration status and are not considered in the computation of the proportion of foreign-born people before selection.

Table 3: Descriptive overview before and after sample selection

Descriptive overview of demographics before and after sample selection

variable	mean before	sd before	na before	mean after	sd after
foreign-born	0.1374	/	0.0141	0.1315	/
female	0.5302	/	/	0.5241	/
45	0.2685	/	/	0.2278	/
age	67.8430	9.74	/	68.0114	9.41

Source: own computations; data from SHARE

3.3.2 Estimation of Transition Probabilities

In the second step of data analysis (see the markdown file 2*_Modelling_and_Adjusting_Transition_Probabilities.Rmd*), I use this dataset to create all necessary variables (age groups, destination and origin health states and a new migrant status variable) and to model and adjust the transition probabilities. The origin health state is always based on self-rated health from the first wave while the destination health state is defined as the lead variable of the origin health state for each case. Thereby, I group by mergeid and source which allows the data transformation for every observation distinguishing between observations from 45 and 67. It always uses the second row of each case grouped by mergeid and source to define destination health. This means that for each observation the value of the health state variable of the second wave (5 or 7) is used.

Transition probabilities in a multistate model are assumed to follow a Markov chain with the property to be memoryless (Markov property) (Dudel, 2021, p. 407). These probabilities are defined as $P(Z_{t+1} = s_j|Z_t = s_i)$ and mean the conditional probability that an individual in state s_i at time t will be in state s_j at time t+1 (Dudel, 2021, p. 406). This implies that these probabilities only depend on the state of the process at time t and not on the history before (Dudel, 2021, p. 407). Other assumptions of Markov chains are that they are homogeneous and they are

absorbing (Dudel, 2021, p. 407). Homogeneity means that transition probabilities do not change over time (Dudel, 2021, p. 407). Absorbing "[…] means that there is at least one state that will eventually be reached with probability one and that cannot be left once reached (Dudel, 2021, p. 407)." That is death in my case.

For the modelling of the transition probabilities, I ran an individual-level multinomial generalised additive model that regresses the destination health state on age groups (using penalized b-splines), sex, foreign-born status and the origin health state:

$$\text{logit}(P(\text{disab\_to} = k|X)) = ß_{0k} + f_{1k}(\text{age\_group}) + ß_{1k} \cdot \text{sex} + ß_{2k} \cdot \text{foreign} + ß_{3k} \cdot \text{disab\_from} \quad (1)$$

For that, I used the vgam-function from the R package VGAM (vector generalized linear and additive models) (Yee, 2015). The advantage of these models over normal regressions is that they allow the modelling of more complex trajectories over age. By jointly modelling migrants and natives, the trajectories for the small migrant sample are stabilized.

Subsequently, the model predictions are used as transition probabilities for all different groups (by sex, foreign-born status and origin health state) over all age groups. This means that probabilities can be additionally predicted for age groups without or with only a few transitions which is particularly relevant for the oldest age groups.

3.3.3 Mortality Adjustment of the Predicted Probabilities

Due to a missing link between SHARE data and national death registers, the number of unknown vital statuses is larger than in other studies increasing the importance of mortality correction (Bergmann, Michael: Kneip, Thorsten et al., 2019, p. 42). The mortality correction is similar to the method from Moretti et al. (2023a), but differently formulated and should lead to the same result. To apply the method, I prepared the data in a way that the predicted probabilities for healthy and unhealthy people are included in different columns instead of different rows. This allowed me to save all newly computed rates and probabilities in new columns. This method is based on the observation that mortality results from the mortality of unhealthy people weighted by the prevalence and of healthy people weighted by the complementary probability of prevalence (Moretti et al., 2023a, p. 8):

$$m(a) = pi(a) \cdot m^{\wedge u}(a) + (1- pi(a)) \cdot m^{\wedge h}(a) \quad (2)$$

For adjusting mortality, mortality rates separately by sex and foreign-born status from an official source are necessary. The most recent available source is a paper by zur Nieden and Sommer (2016). In their online appendix, they show life tables separately by sex and citizenship

(Foreign versus German) (zur Nieden & Sommer, 2016). This data has two shortcomings. Firstly, the data is from the 2011 census, thus older than the SHARE data, and secondly, it is based on citizenship instead of migration status. Since no alternative was available, an attempt was however made with this life table. I added up the life tables from continuous age into two-year age groups.

As mortality from life tables might be too irregular, I used the Mort1Dsmooth function from the package MortalitySmooth (Carmada & Riffe, 2015) to smooth with penalized b-splines the mortality rates whereby the life table number of person-years lived between ages x and x+1 (L_x) was used as an offset.

The mortality adjustment includes the following steps:

First, I needed to convert the death probabilities into rates:

$$r = - \log(1 - p_d^{old}) / t \quad (3)$$

r means the death rate, p_d the death probability and t time which is two in my case

Then, I compute a ratio of the initial mortality estimates (unhealthy mortality divided by healthy mortality:

$$R(a) = mu(a) / mh(a) \quad (4)$$

For the next step of the adjustment, I need the prevalence of bad health state. I derived this from a binary logistic generalised additive model like that above that regresses the origin health state on age group (using penalized b-splines; in R: bs = “ps”), sex and foreign-born status:

$$\text{logit}(P(\text{disab\_from} = k|X)) = ß_{0k} + f_{1k}(\text{age\_group}) + ß_{1k} \cdot \text{sex} + ß_{2k} \cdot \text{foreign} \quad (5)$$

I again used the vgam-function for that.

The new mortality estimates are then derived in the following way:

$$mh\hat{}(a) = m(a) / (1 - pi(a) + pi(a)\, R(a)) \quad (6)$$

$$mu\hat{}(a) = mh\hat{}(a)\, R(a) \quad (7)$$

m(a) means the life table mortality, pi(a) is the predicted prevalence and R(a) is the mortality rate ratio.

Following this, it is necessary to back-transform the rates into probabilities and to re-constrain the other transition probabilities. For back-transforming into probabilities, I use the following formula (the inverse of the formula above):

$$p_d^{new} = 1 - e^{-rt} \quad (8)$$

where t is again two.

I re-constrain the other transition probabilities (to healthy and unhealthy states) so that all three probabilities for both groups add up to 1. The intuition is that the adjusted mortality is more trustworthy, therefore I keep that fixed. The remainder (1 – death probability) is divided into the remaining two probabilities so that the unconstrained proportions of these two probabilities remain the same.

$$p_h^{new} = (1 - p_d^{new}) \cdot (p_h^{old} / (p_h^{old} + p_u^{old} \quad (9)$$

$$p_u^{new} = (1 - p_d^{new}) \cdot (p_u^{old} / (p_h^{old} + p_u^{old})) \quad (10)$$

Despite these efforts, the results section only shows results for the unadjusted probabilities. The reason is that the results derived from the adjusted probabilities are rather implausible. For instance, they show a higher total life expectancy for men than women and for natives compared to foreign-born people. Particularly, the total life expectancy for both male groups of more than 36 years at the age group 50-51 seems very unlikely and further investigations would be needed to make this procedure trustworthy, which was beyond the time available for this thesis[2]. However, the results can be found in Appendix B.

3.3.4 Multistate Models for Deriving the Life Expectancies

The transition probabilities serve as input for multistate models to derive the HLE, unhealthy life expectancy (ULE) and LE. The files 3*_Multi_State_life_tables_unadjusted.Rmd* and 4*_Multi_State_life_tables_adjusted.Rmd* include these computations. For the computation, I rely on Dudel's and Li's package for discrete-time multistate models (dtms) (Dudel, 2021; Dudel & Li, 2024).

The first step before using the dtms package included the data preparation in a way that allowed the right specification of the transition matrix P. I divided the full dataset of all probabilities into four sub-datasets for all different groups (female natives, female migrants, male natives and Male migrants).

Each data set needed to be constructed in a way that all different transitions for all age groups are included in different rows with only one column for all transition probabilities. It was necessary to shift the probabilities for the right specification of the transition matrix because transitions occur between age x and age x+1 but not between age x and age x. As the predictions were always for the same age groups, the probability for the destination age group was used.

[2] The reason for these discrepancies remains unclear. Perhaps, for men, mortality is corrected downwards because they are more likely to report a good health status. The lower prevalence decreases the weight of the rate ratios between both death probabilities.

The oldest age group was thus deleted. For the dtms package, it was necessary to combine the origin and destination health states with age group into respectively one variable.

Then, I constructed the transition matrix P which is block-off-diagonal using the function dtms_matrix (Dudel, 2021, p. 407). The block-off diagonality is because transitions take place between age x and age x+1 (Dudel, 2021, p. 407). The rows correspond to the different origin states and the columns include the different destination states.

$$P = \begin{pmatrix} 0 & P_{51} & 0 & \dots \\ 0 & 0 & P_{53} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (11)$$

(Dudel, 2021, p. 407)

Several measures of interest can be derived from the transition matrix P (Dudel, 2021, p. 407). The expected time people spent in state s_j starting from state s_i (the different expectancies), n_{ij} is computed with the following formula:

$$N = (I_s - U)^{-1} \quad (12)$$

(Dudel, 2021, p. 407)

U is the transition matrix excluding absorbing states (death), I_s is an identity matrix of dimension s × s with s as the number of transient states and N refers to the fundamental matrix (Dudel, 2021, p. 407). This equation assumes that transitions take place at the end of intervals and n_{ij} is overestimated by 0.5 (Dudel, 2021, p. 407). The corrected expectancies depending on state s_i are given by subtracting 0.5 from the row sums of N (Caswell, 2000; Dudel, 2021, p. 407). The function dtms_expectancy does exactly this computation of all expectancies.

3.3.5 Decomposition Method

The last step of my empirical analysis includes the decomposition of all different expectancies which allows estimating the age-specific contributions of all transition probabilities to the differences in the expectancies between migrants and natives. For this, I draw on the approach explained by and used the R code from Moretti et al. (2023a, p. 11)[3]. I decompose the changes in the different life expectancies "[..] by reparameterizing [Life expectancy] calculations in terms of only mortality and transitions between health states, and which omit transitions within states (Moretti et al., 2023a, p. 11)." Compared to matrix calculations, mortality is included in the calculation while self-transitions were excluded (Moretti et al., 2023a, p. 11).

[3] Their R code is available in Moretti et al. (2023b).

For that, I organise the four age-specific vectors of the incidence, recovery and both state-specific death probabilities into one single vector θ (Moretti et al., 2023a, p. 11). This vector includes all necessary parameters for computing the transition matrix, the fundamental matrix and the final life expectancy estimate (Moretti et al., 2023a, p. 11). As the different life expectancies are a function of the vector θ [HLE/ULE = $f(\theta)$], the difference in the respective expectancy can be decomposed implied by two versions of θ, one for German-born and one for foreign-born people (Moretti et al., 2023a, p. 11).

As Moretti et al. (2023a), I rely on the method from Horiuchi et al. (2008). The Horiuchi method assumes that the dependent variable y (in my case: the respective life expectancy) can be written as a function of covariates x, irrespective of whether x is causally related to y (Horiuchi et al., 2008, p. 787). In my case, the covariates are the age-specific transition probabilities. Then, the difference in the expectancies can be decomposed by describing it as the sum of the differences in the age-specific transition probabilities (Horiuchi et al., 2008, p. 787).

The difference in the different life expectancies between migrants and natives can thus be written as

$$y_2 - y_1 = \sum_{i=1}^{n} c_i, \text{ where } c_i = \int_{x_{i1}}^{x_{i2}} \frac{\partial y}{\partial x_{i1}} dx_i \quad (13)$$

(Horiuchi et al., 2008, p. 787).

Thereby, c_i means the total difference in respective expectancy y produced by differences in the i-th age-specific transition probability x_i (Horiuchi et al., 2008, p. 787). Further details of the Horiuchi method are described in Horiuchi et al. (2008).

This decomposition results in a vector of age-specific contributions from each difference in the vector θ to the total differences (Moretti et al., 2023a, p. 11). The age-specific contributions from the decompositions of the healthy and unhealthy life expectancy are then added up to derive the decomposition of total life expectancy.

3.4 Ethical Considerations

Research using individual data sources such as the *Survey of Health, Ageing and Retirement in Europe (SHARE)* should always be conducted in an ethically responsible manner. That is particularly important in my case because of the use of the sensitive self-rated health measure. Respondents to the *SHARE* only accept an anonymized use of their data for scientific purposes. The definition of anonymity is based on the rules of the *German Federal Statistics Act* and the *German Federal Data Protection Law* (SHARE-ERIC, 2020). Therefore, the names of respondents and other possible identifiers are not stored in the data set. Instead, individuals are

assigned unique individual and family identification numbers. These IDs allow merging data from different modules and waves.

Furthermore, it is impossible to simply download data from the *SHARE Research Data Center website* (SHARE-ERIC, 2024a), and users must submit a signed data statement. This statement must either state the user's scientific affiliation or users without any scientific affiliation are asked to describe their research project. The data collection and use must comply with the *European Union* and *national data protection laws*, particularly the *European General Data Protection Regulation (GDPR)* (SHARE-ERIC, 2020). Users accept the use according to these laws by signing the statement. I saved the original and my analysis data in Google Drive with my account at Groningen University. Therefore, it could only be accessed with my login credentials. Only my first supervisor who is affiliated with the university could also access it.

3.5 Use of Generative AI

Due to the emerging use of generative Artificial Intelligence tools (AI) such as ChatGPT, it is important to use them responsibly. ChatGPT can help understand and write statistical code. As I am a new user of R, I have sometimes used these functions to understand the R code and packages such as the dtms package or the decomposition code. I also asked ChatGPT how to adjust specific details in my code that I was unsure about. However, I always checked whether it resulted in what I wanted. ChatGPT was not always right. Then, I found a solution myself. I rarely asked ChatGPT for synonyms or replacements for single words but I did not use it to produce text, to adjust the writing style or to look for literature.

Furthermore, I used the free version of the typing assistant Grammarly to check spelling, grammar, punctuation and similar mistakes.

4 Results

In this section, I describe the most important findings based on the transition probabilities without mortality correction. The logged death probabilities and the decompositions of the unhealthy life expectancy based on the mortality-unadjusted probabilities are included in Appendix A. All results from the mortality-adjusted probabilities can be found in Appendix B. First, I start with plots showing the predicted transition probabilities. Then, results about the different life expectancies are presented. Finally, I will describe findings from the decomposition of the life expectancies.

4.1 Transition Probabilities

In this section, I describe plots of the transition probabilities over age. Figure 2 shows the transition probabilities over age. In Figure A 1 in Appendix A, the death probabilities are shown on a log scale. For all lines, it is important to consider that the prediction for the oldest age groups (starting around age 85 – 90) is based on a relatively small number of observed individuals. Therefore, the strongly changing picture of the lines for these age groups might be biased and should not be over-interpreted. However, it might be realistic that health deteriorates strongest in the oldest groups as these are ages over the usually computed life expectancy (Destatis, 2024c).

For all groups, it is visible that transitions from unhealthy (blue-green line) and healthy states (orange line) to healthy status are predicted to decrease over age. For people starting healthy, this decrease is stronger in the oldest age groups while the decrease for unhealthy people is linear over age. For all groups, it can also be noticed that more transitions from the healthy to healthy status are predicted than from unhealthy states over all ages showing the path dependency between origin and destination health.

In contrast, the transition probabilities to the unhealthy state (healthy: dark blue, unhealthy: purple) are first predicted to increase over age for all groups and both origin states. The slope of the increase in the transition probability is stronger for people originating from the healthy state. For people originating from unhealthy status, the probability is already high for the youngest age groups. After around age 90, the transition probability decreases over age for all groups and origin states. The line of people starting in good health always remains below the line for unhealthy people. Both lines converge only at age 100.

The final two lines show the death probability for healthy (green) and unhealthy people (yellow). Figure A 1 in Appendix A shows the logged death probabilities. The pattern of these lines is relatively similar for all groups. In the ages between 50 and 65, both probabilities increase only slightly. At these ages, mortality is already relatively high compared to the ages after. Then, a linear increase of both logged probabilities can be seen up to about age 95 when this increase diminishes. A linear increase on the log scale corresponds to an exponential increase in death probability with age.

For all foreign-born-sex groups, the death probability for people starting in bad health is always higher.

Two important differences regarding sex and foreign-born status should be noticed. First, the death probability for males is higher than for females. Second, the death probability is higher for natives compared to foreign-born people. Both findings connect to previous research about males' higher mortality and the MMA (Abraído-Lanza et al., 1999; Razum, 2008; Razum et al., 1998).

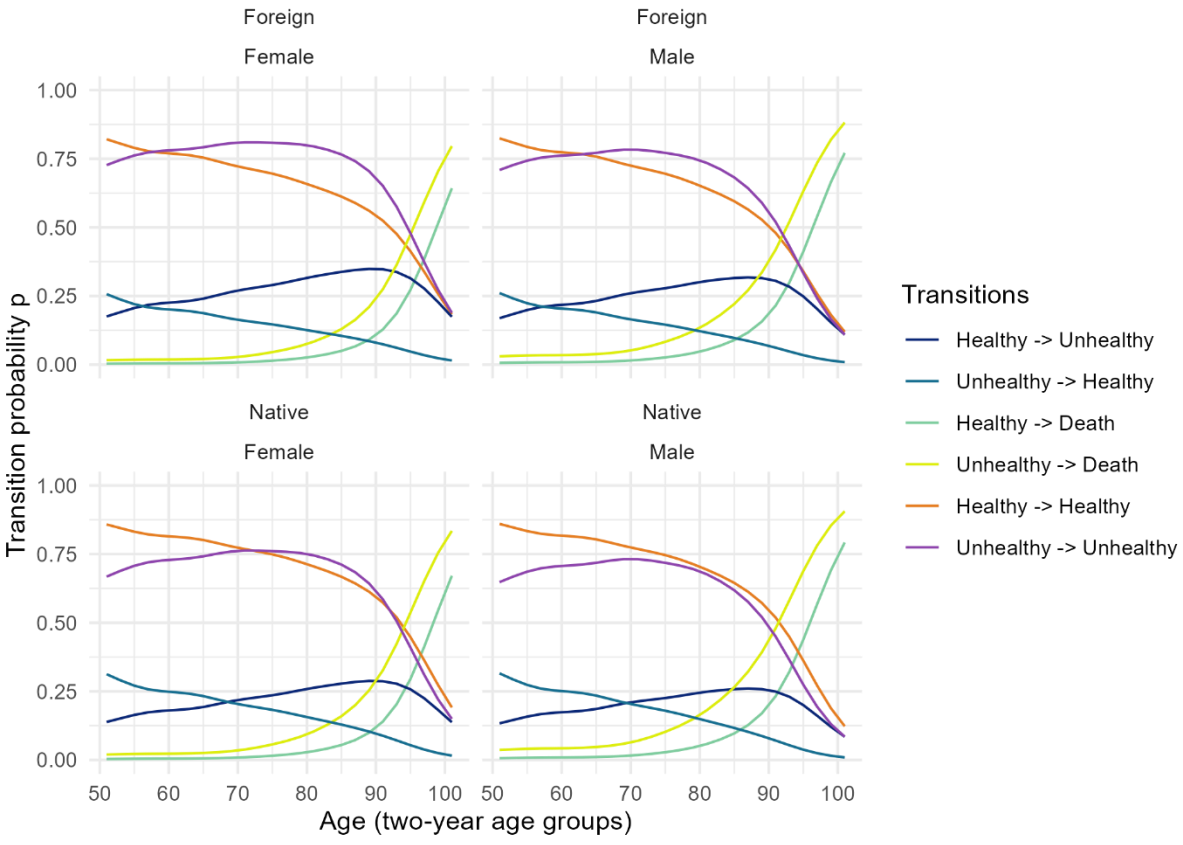

Figure 2: Predicted transition probabilities without mortality adjustment (multinomial-logistic generalised additive model with p-splines)
Source: own computations, data from SHARE

4.2 (Healthy) Life Expectancies

The following section describes the results for the healthy life expectancies derived from the Multistate Model with the dtms package.

Tables 2 to 5 show the life expectancies for the four different groups. The tables show the life expectancies depending on the origin health state. The average expectancies are computed based on the prevalence for the youngest age groups (the initial distribution of health states). For this prevalence, the predictions of the prevalence GAM model for the age group 50-51 are used.

Tables 2 and 3 show the different expectancies for native and foreign-born women.

Using the initial distribution of health states, foreign-born women are expected to live around 0.79 years longer at age 50-51 than women born in Germany. However, foreign-born women suffer from a shorter time in good health than Natives (-2.78 years) but are expected to live longer in bad health (+3.57 years). On average, both groups are expected to live longer in bad health than in good health.

For both groups, the people starting in bad health are expected to live longer unhealthy than healthy people (Natives: 2.87; Foreign-born: 3.00). However, foreign-born women starting in good health also live longer unhealthy than healthy (+2.11 years) while women born in Germany starting healthy are expected to live around 4.51 years longer healthy than unhealthy. While foreign-born women are expected to live longer in bad health than healthy regardless of their origin health state, for women born in Germany, this only applies to people starting unhealthy.

The results imply that foreign-born women live indeed longer but this additional time is spent rather in bad than in good health and this is even the case for people starting healthy.

Table 4: (Healthy) Life expectancy for German-born women

(Healthy) Life expectancy for Female Natives

Starting States	HLE	ULE	LE
Healthy	18.81	14.29	33.10
Unhealthy	14.78	17.17	31.95
Average	15.94	16.34	32.28

Starting States	HLE	ULE	LE

Source: own computations; data from SHARE

Table 5: (Healthy) Life expectancy for foreign-born women

(Healthy) Life expectancy for Female Foreign-born

Starting States	HLE	ULE	LE
Healthy	15.80	17.91	33.71
Unhealthy	11.86	20.91	32.77
Average	13.16	19.91	33.08

Source: own computations; data from SHARE

Tables 4 and 5 show all the life expectancies for Men.

Foreign-born men are expected to live around 0.74 years longer than German-born men at age 50-51 using the initial distribution of health states. Yet, foreign-born men suffer from a shorter expected time spent healthy than native men (-2.28 years) but a longer time in bad health (+3.02 years). Compared to women a notable difference arises: While both women groups are expected to live longer in bad health than healthy, for men, only foreign-born people are expected to live 3.97 years longer in bad health than healthy. In contrast, German-born men are the only group expected to live longer healthy than unhealthy (+1.33 years).

For both migrant statuses, people starting in bad health suffer from a longer time in bad health than people who start healthy (natives: 2.54 years; foreign-born: 2.67 years). Yet, while foreign-born men starting unhealthy remain much longer in bad health than healthy (6.18 years), this difference is small for natives (0.60 years). Men from both migrant statuses who originate from good health are expected to live longer healthy than unhealthy. However, this difference is much larger for natives (6.26 years) than for migrants (0.64 years).

This is different to the findings for women. For them, only German-born women starting in good health are expected to live a longer time in good health while foreign-born women from both origin health states live longer in bad health than healthy.

The results imply that foreign-born men have a higher life expectancy than German-born men, but this time is spent in bad health rather than healthy.

Table 6: (Healthy) Life expectancy for German-born men

(Healthy) Life expectancy for Male Natives

Starting States	HLE	ULE	LE
Healthy	17.40	11.14	28.54
Unhealthy	13.08	13.67	26.75
Average	14.29	12.96	27.25

Source: own computations; data from SHARE

Table 7: (Healthy) Life expectancy for foreign-born men

(Healthy) Life expectancy for Male Foreign-born

Starting States	HLE	ULE	LE
Healthy	14.82	14.18	29.00
Unhealthy	10.67	16.85	27.52
Average	12.01	15.98	27.99

Source: own computations; data from SHARE

Besides the differences between German- and foreign-born people, my findings also confirm the known sex gap. On average, women are expected to live much longer than men (natives: + 5.03 years; foreign-born: + 5.08 years). However, women from both migrant statuses suffer from a longer time in bad health than healthy on average while for men, this only applies to foreign-born people.

4.3 Decomposition of the Differences in Life Expectancies

The next step of my analysis was to decompose all differences between life expectancies into the contributions of the different transition probabilities. Unfortunately, the sums of the contributions do not fully correspond to the differences between the respective life expectancies in section 4.2. The reason for this discrepancy remains unclear. For HLE, the sums are 0.20 years or 7.19 per cent (women) and 0.25 years or 10.96 per cent (men) smaller than the HLE differences above. For ULE, the sums are 0.08 years or 2.24 per cent (women) and 0.06 years or 1.99 per cent (men) larger than the ULE differences above. For LE, the sums are 0.13 years or 15.19 per cent (women) and 0.19 years or 25.68 per cent smaller than the differences.

The Figures in this section use the direct results from the decomposition. The decompositions of healthy, unhealthy and total life expectancies are shown for both sexes. These analyses show how the transition probabilities excluding self-transitions contribute to the gap in the different life expectancies between natives and migrants. Positive values mean that the respective probabilities shape the difference in favour of a higher expectancy for migrants while negative values imply contributions in favour of natives. The value states how strong certain contributions are.

First, I show the results of the decomposition of healthy life expectancy. Figure 3 shows the results for women.

The upper panel A shows the sums of the contributions of death, incidence and recovery probabilities over age. The total sum states that healthy life expectancy is around 2.98 years lower for migrants than natives. Incidence and recovery probabilities are the drivers of migrants' lower healthy life expectancy. Migrants' higher incidence but lower recovery probabilities decrease their healthy life expectancy by around 1.92 years and 1.43 years respectively.

The positive, but much smaller contributions of both death probabilities cannot compensate for that. It remains unclear whether the positive contributions of death are statistically different from zero. The computation of confidence intervals is too complicated for this thesis[4].

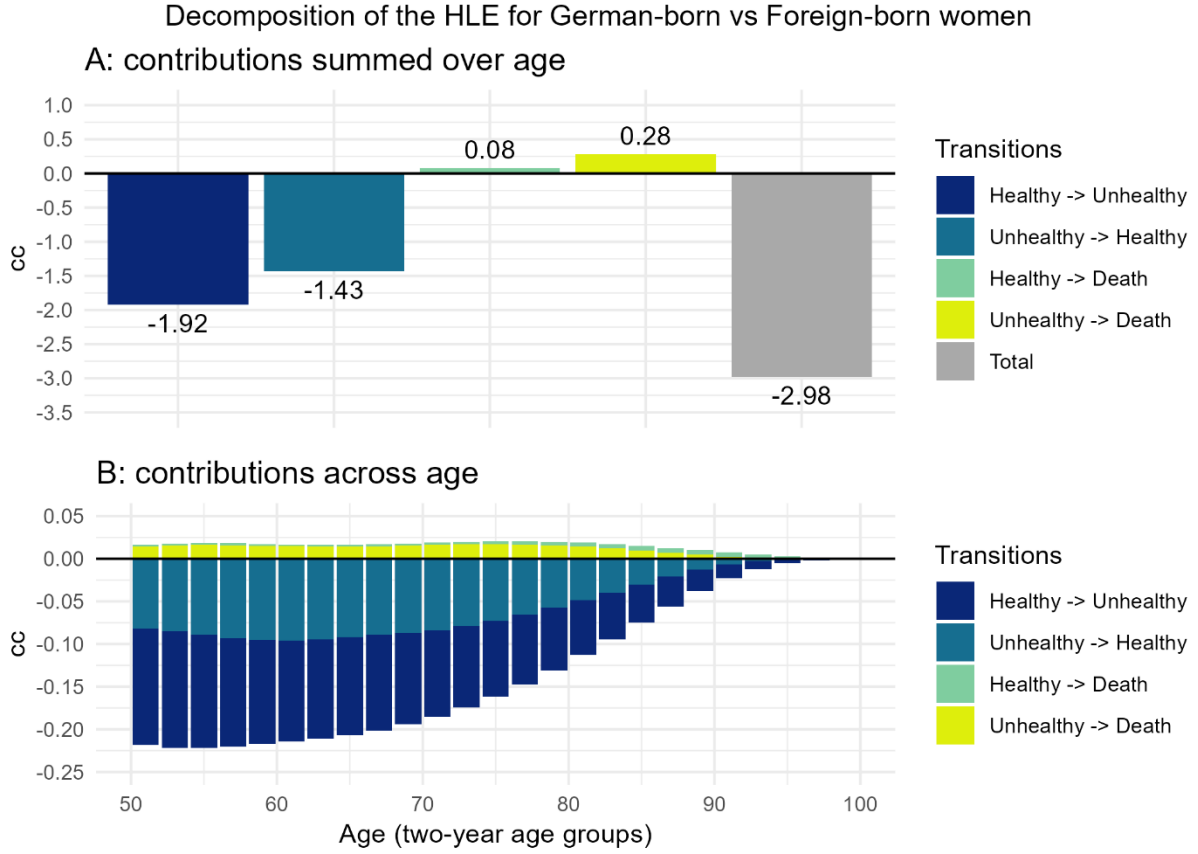

Figure 3: Decomposition of Healthy Life Expectancy for women
Source: own computations, data from SHARE

Panel B of Figure 3 shows the decomposition results over age. The magnitudes of both negative contributions (recovery and incidence) are highest in younger age groups and decrease over

[4] It would be necessary to use simulation methods to compute these confidence intervals such as bootstrapping methods or Monte Carlo Simulation. That goes beyond the scope of my analysis.

age. This shows that migrants' health advantage already develops during midlife while older ages only explain a smaller extent of the disadvantage in HLE. However, for transitions from healthy to unhealthy status, the contributions peak in the youngest age group in terms of the absolute values before they decrease monotonically over age. In contrast, the contributions of transitions from the unhealthy to the healthy state first increase in terms of absolute value over age up to the peak around age 60 (age group 58-59). Then, the absolute values of the contributions also decrease monotonically.

The compensatory contributions of migrants' lower death probabilities, particularly for healthy people, remain small in all age groups. The contributions of death probability for healthy people increase slightly after age group 66-67. After the age group 78-79, the contributions of mortality for unhealthy people decrease.

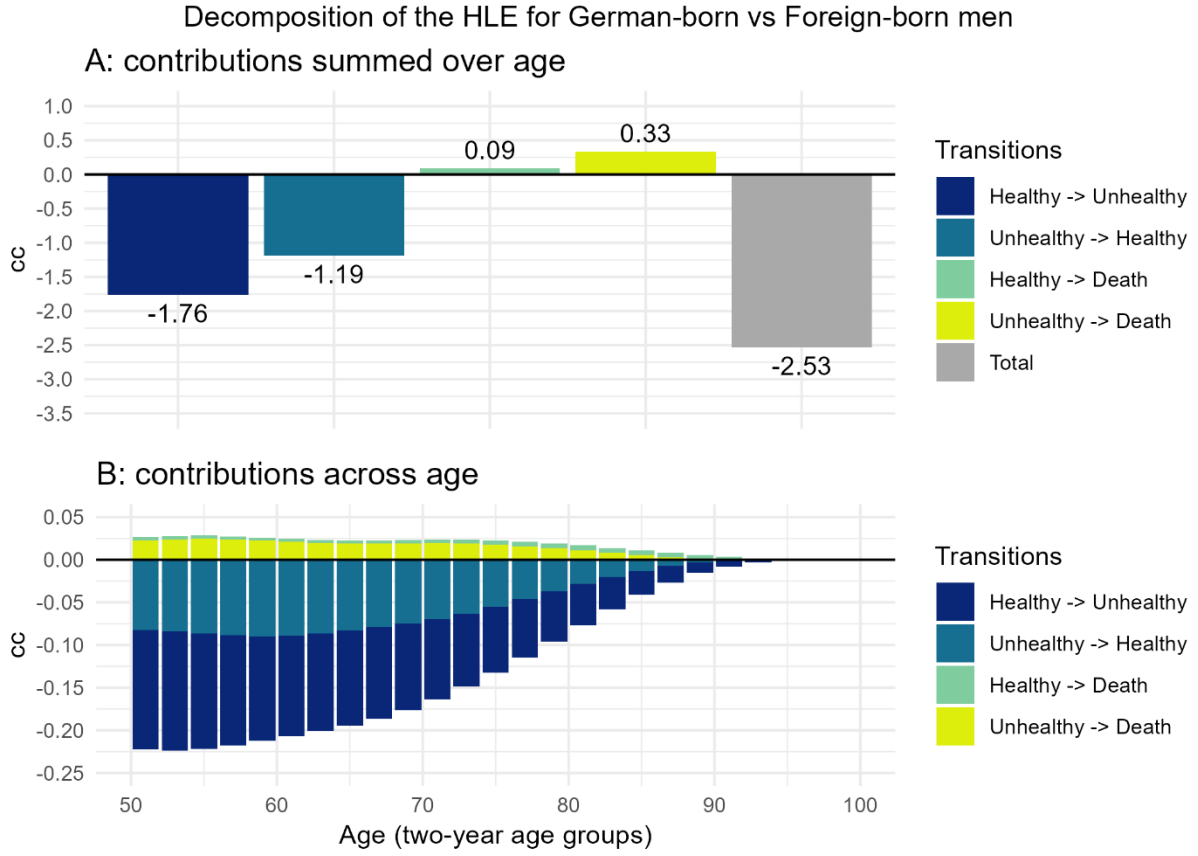

Figure 4: Decomposition of Healthy Life Expectancy for men
Source: own computations, data from SHARE

For Men, the results look similar. Figure 4 shows the decomposition of HLE for men.

In total, migrant men's healthy life expectancy is around 2.53 years lower than natives (Panel A). Again, migrants' higher incidence (-1.76 years) and lower recovery probabilities (-1.19) explain this gap. The positive contributions of both death probabilities are too small to compensate for the health disadvantage.

The contributions of all probabilities are again plotted over age in Panel B.
The plot generally shows the same picture as the women's Figure. The contributions of incidence and recovery probability are again strongest in the youngest age groups and decrease over age in terms of absolute value. The positive contributions of migrants' lower death probabilities are again very small and decrease across age.

The next step of my decomposition was to decompose the unhealthy life expectancy. I do not show the results here, but they can be found in Figures A 2 and A 3 of Appendix A. The reason for this decision is that all transition probabilities substantially shape migrants' higher ULE compared to natives (women: + 3.65 years; men: + 3.08 years). There is not one or two transition probabilities exceptionally explaining the expectancies gap, but migrants' higher ULE is the result of lower death (healthy and unhealthy) and higher incidence but lower recovery probabilities. Only the contribution of the death probability for healthy people is close to zero but positive. The bottom panels of Figures A 2 and A 3 show that the contributions of all probabilities are already strongest between ages 50 and 60 and decrease afterwards.

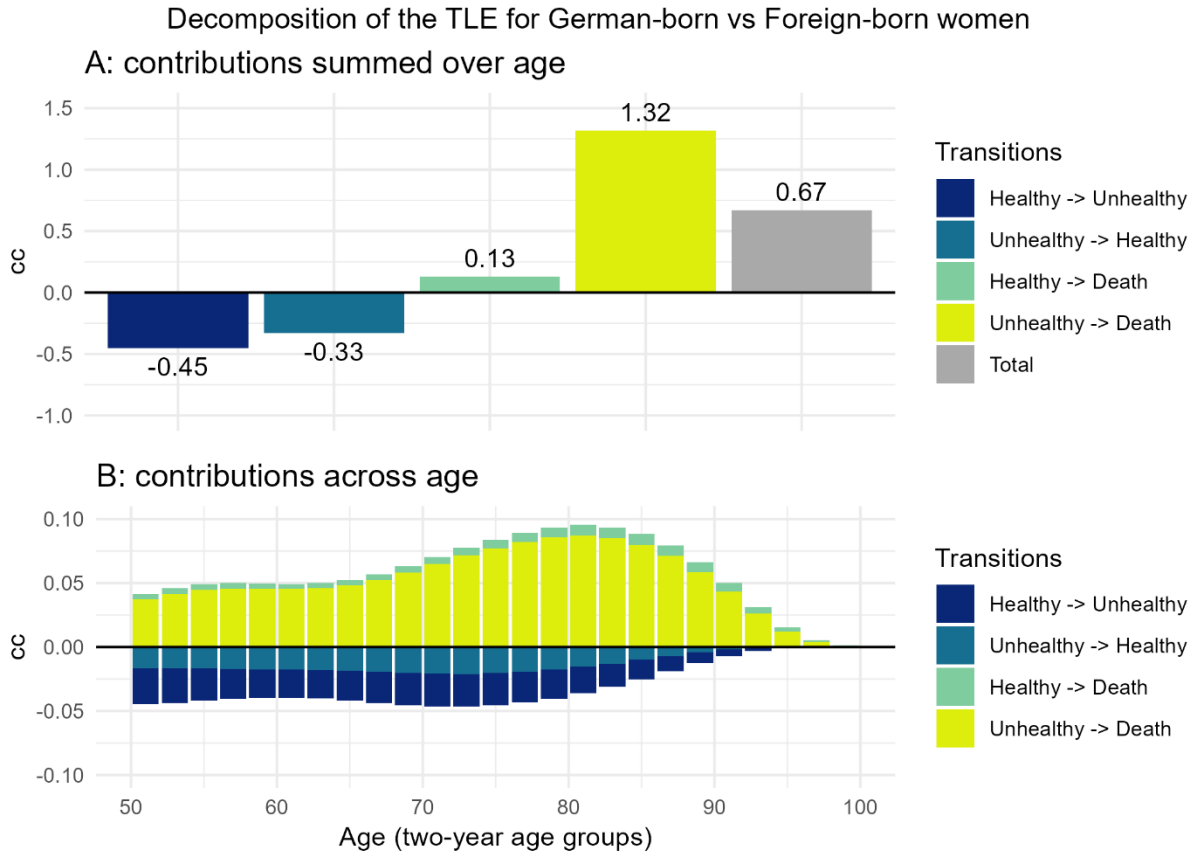

Figure 5: Decomposition of Total Life Expectancy for women
Source: own computations, data from SHARE

As the last step, I added up the results of the HLE and ULE decompositions to derive the decomposition for total life expectancy.
Figure 5 shows the results for women and Figure 6 for men. The results are very similar for both sexes.

The total life expectancy of those born abroad is 0.67 (women) and 0.55 years (men) higher than that of those born in Germany. This gap is almost exclusively driven by migrants' lower death probability for people starting in bad health (women: + 1.32 years; men: + 1.31 years) while differences between the death probabilities for healthy people only slightly contribute to the gap (women: + 0.13 years; men: 0.15 years). On the other hand, both, migrant's higher incidence (women: -0.45 years, men: -0.54 years) and lower recovery probabilities (women: -0.33 years, men: -0.36 years) substantially contribute to closing the gap.
Two conclusions can be derived from these findings. First, the results support previous findings of the migrant mortality advantage. Secondly, if recovery and incidence probabilities were in favour of migrants, or just similar to those of the natives, the MMA would be much stronger.

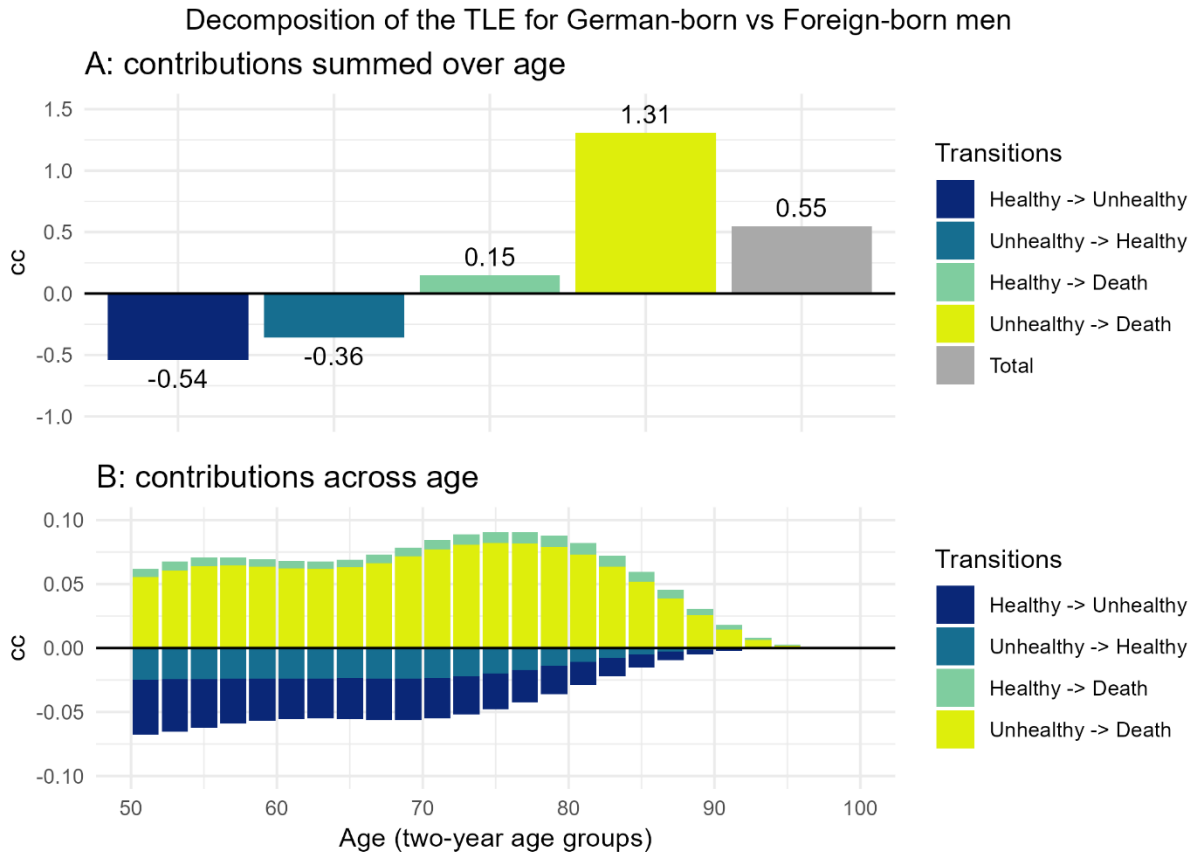

Figure 6: Decomposition of Total Life Expectancy for men
Source: own computations, data from SHARE

The bottom panels B of Figures 5 and 6 show the decomposition results across age. Compared to the Figures for HLE above with a different y-axis, the absolute values of all contributions for both sexes are much smaller mostly fluctuating between 0.05 and 0.10. The gap-widening contributions of the lower death probability for unhealthy migrant women increase from around 0.04 years in the youngest age group up to around 0.08 years in the age group 80-81. After this peak, the contributions again decrease until they disappear for the oldest age group. This implies that migrants gain their mortality advantage mainly during peak ages around 80.

For men, the contributions of the lower unhealthy death probability first fluctuate between 0.05 and 0.06 between ages 50 and 60 before they increase to around age 75 to a peak of around 0.075. After this age, they monotonically decrease and disappear in the oldest ages. On the other hand, the contributions of migrants' higher incidence and lower recovery probability fluctuate around -0.025 (incidence) and -0.015 (recovery) for women before the absolute values of these contributions decrease from age 70 onwards.

For men, the contributions of the recovery probabilities remain relatively constant around -0.025 before their absolute values decrease from age group 68-69 onwards. The absolute values of the contributions of the incidence decrease from its peak in the youngest age groups (-0.04) up to age 64-65 when it remains shortly constant. Then, the absolute values again decrease from age 70 onwards.

5 Discussion and Conclusion

5.1 Aims and Summary of the Methods

The migrant mortality advantage is a well-acknowledged demographic finding although its causes remain unclear. Several studies have shown that migrants can expect to live longer than natives (e.g. Abraído-Lanza et al., 1999; Boulogne et al., 2012; Wallace & Kulu, 2014). However, studies showed that migrants report worse health, more chronic conditions, more limitations of functioning and a higher depression prevalence than natives (Aichberger et al., 2010; Lanari & Bussini, 2012; Solé-Auró & Crimmins, 2008). A measure combining morbidity and mortality is healthy life expectancy (HLE). Previous research reported that older migrants in Belgium, the Netherlands, and England and Wales suffer from a shorter HLE (Reus-Pons et al., 2017). This study relied on the Sullivan Method which is biased if the assumptions of no-recovery, same mortality rates for healthy and unhealthy people and stationarity of age-specific disability prevalence are violated. Therefore, I investigated the following research question:
To what extent do Life Expectancy (LE) and Healthy Life Expectancy (HLE) at age 50 differ between migrants and natives, and how do the different transitions between health statuses and death contribute to these differences?

The specific objectives were (1) to compute LE and HLE separately for migrants and natives and (2) to decompose the differences in HLE into the contributions of the transitions between the different health states. First, I described the four existing explanations for the MMA and previous research to answer these questions.
The first explanation argues that unhealthy people emigrate to their country of origin because of a wish to die at home or of social networks in the home country. The second possible explanation is that migrants are healthily selected from their home population. Thirdly, it is argued that migrants from less developed countries share healthier behaviours and benefit from social support networks decreasing their risk for certain diseases. The last explanation is errors in population and death coverage, and age misreporting.
For my analysis, I then used the SHARE data to predict transition probabilities between a healthy, an unhealthy state and death based on self-rated health and mortality. These probabilities were then used as input for a multistate model. Finally, I decomposed the gaps in the different expectancies between migrants and natives.

5.2 Summary and Discussion of the Results

Plots of the transition probabilities over age mainly looked as expected. The probability of transitions to good health increases over age while the probability of transitions to bad health decreases. Both death probabilities remain constant between the ages of 50 and 60 before their exponential increase. However, the probability of bad health drops in ages above 90. This might be related to the increasing death probabilities and the small number of observations at the oldest ages. The plots also showed that previous health predicts future health as the probability of self-transitions is high. Furthermore, people in bad health are likelier to end up unhealthy than healthy people. The reverse applies to transitions into good health.

Another important finding appeared for the death probabilities. The death probability for unhealthy people is higher than for healthy people. This shows that health predicts mortality, as shown by other studies (DeSalvo, Fisher, et al., 2006). It is plausible since certain diseases and chronic conditions such as coronary heart disease or cancer increase the death probability. However, it shows that the Sullivan Method assuming the same mortality rates for healthy and unhealthy people is biased and might yield invalid results.

Then, I reported healthy, unhealthy and total life expectancies for all groups. Generally, the findings connect to research by Reus-Pons et al. (2017) that the longer life expectancy is related to longer unhealthy time which is found for both sexes. Migrants from both sexes can expect to live longer in bad health. These findings connect to previous research reporting migrants' worse self-rated health, chronic conditions and mental health (Aichberger et al., 2010; Lanari & Bussini, 2012; Solé-Auró & Crimmins, 2008).

My findings do not allow me to conclude which of the MMA explanations holds since I do not directly test them. However, there are some points to consider. If unhealthy people returned to their home country as argued by the Salmon bias, then it remains unclear why migrants in Germany are expected to live much longer unhealthy than natives. It might explain a part of the MMA, but it is implausible regarding the HLE and ULE differences. If unhealthy people, particularly at older ages, return to their home country, the stayers in the destination country should be comparatively healthier. However, they suffer from longer time in bad health than natives although they live longer which appears as a contradiction. Is it only the very oldest people in the worst health state going back to their home country driving the mortality advantage? This seems unrealistic as people in the worst health cannot always move anymore although it depends on the specific disease.

Secondly, the healthy immigrant thesis led to the argument that the health advantage of migrants should decrease over age as migrants' frailty compositions converge to that of natives. However, my results of longer unhealthy expectancy for migrants do not fit into this thesis. This hypothesis cannot explain why migrants should live longer in bad health than natives.

Thirdly, it is difficult to assess the applicability of the cultural explanation. On the one hand, the effects of healthier behaviours might be more relevant at older ages due to the longer latency behaviours which would lead to more pronounced advantages of migrants in chronic diseases and mortality at older ages. On the other hand, assimilation into the host society might compensate leading to non-existent or small health gaps. Due to the longer latency, the life expectancy might be higher for migrants. The higher LE connects to this explanation, but the cultural explanation cannot explain the much longer time in bad health.

Fourth, it is difficult to assess how my findings relate to the last thesis as data artefacts can work in different directions. If people overstate their age, time lived in poor health is assigned to an older age than their actual one. This might decrease the time spent in poor health up to this age increasing the computed time people can expect to live healthy. However, this rather contradicts my findings of lower healthy life expectancy.

The last step of my analysis was the decomposition of healthy, unhealthy and total life expectancies into the contributions of the different transition probabilities.

First, I found a higher incidence and a lower recovery probability mainly contributing to migrants' lower healthy life expectancy. Both lower death probabilities can only partially compensate for that. A part of the recovery disadvantage might be explained by the salmon bias, for instance, if migrants lose support networks of returnees. It might be that migrants' higher incidence probabilities are linked to their socioeconomic position relative to natives. This would imply that migrants' lower socioeconomic status (SES) and employment situation are contributors to this disadvantage. To test this idea, future studies could include the SES in the VGAM model and predict transitions for the same SES for migrants and natives.

Second, my findings showed that all transition probabilities contribute to migrants' higher unhealthy life expectancy. Only the contribution of the death probability for healthy people is very small and it remains unclear whether it is significantly different from zero.

Finally, I added up the decompositions of the HLE and ULE to derive the decomposition of total life expectancy. Both, migrants' lower death probabilities, particularly for unhealthy migrants, contribute to migrants' slightly higher total life expectancy.

On the other hand, migrants' higher incidence and lower recovery probabilities decrease the gap in life expectancy.

The finding of mainly the lower death probability of unhealthy migrants contributing to the mortality advantage fits well with research by Zufferey (2016) showing migrants and foreigners gaining their mortality advantage among the most vulnerable people such as labour market inactive people. Thus, future research should investigate the interrelations between social and migrant health inequalities.

5.3 Contributions of my Study

My study makes several contributions to previous research.

First, my thesis is the first study investigating differences in healthy life expectancy based on multistate models instead of the potentially biased Sullivan method. My study confirms the results of Reus-Pons et al. (2017) who found that older migrants in Belgium, the Netherlands, England and Wales suffer from a shorter HLE.

Second, this methodology allowed me to investigate the causes of the differences in the different expectancies using the Horiuchi decomposition method. My findings have shown that migrants' disadvantage in healthy life expectancy is mainly explained by contributions of a higher incidence and a lower recovery probability. In contrast, both death probabilities, particularly for unhealthy people, contribute to the migrants' advantage in total life expectancy despite their disadvantage in recovery and incidence.

Third, my study contributes to the development of the methods used to compute healthy life expectancy. My plots of the transition probabilities above clarify that the death probability is higher for unhealthy than healthy people. Their contribution to the MMA is also different with a higher contribution of the death probability for unhealthy people. This questions the assumption of the Sullivan method that the death probability is the same for unhealthy and healthy people. My research contradicts the Sullivan method's second assumption of no recovery. The recovery probability is relatively high in the younger ages before it decreases. This probability partially contributes to the difference in HLE showing its importance.

5.4 Limitations of my Study and Directions for Future Research

The first limitation of this study is that I only used German SHARE data. The reason for choosing only one country was that it is easier to work on a homogeneous population rather than having to deal with different social, economic, political and welfare contexts. However, as the results based on the mortality-adjusted probabilities looked implausible, I only reported findings based on the unadjusted probabilities. Future studies should combine findings from

several countries. It is at least possible to compute mortality-unadjusted probabilities for several countries using the SHARE data.

Second, my study relied on self-rated health (SRH). The reason was its high degree of validity and reliability in previous studies and because it measures the general health status instead of only one specific facet. However, other measures could be used for new studies such as health limitations. As previous studies have shown that migrants do not only perform poorly on SRH, it is not expected that this would fundamentally change the picture. A reason for migrants' higher ULE could be sick migrants close to death leaving Germany and dying abroad quickly, but their deaths would not be recorded in Germany. Then, it would not only increase migrants' LE but also extend their ULE because unhealthy migrants who are not close to death would be over-represented in Germany. Further research could distinguish between chronic and acute diseases. If migrants' longer ULE is only caused by chronic diseases, this would be a strong indication that their longer ULE is at least partially generated by the salmon bias.

Third, mortality is badly reported and potentially underestimated in survey data. I excluded people with unknown vital status as there is no reliable way to estimate whether they are dead or whether they only disappeared from the study. However, even among people with known vital status, the degree of certainty of death remains unsure. That is because mortality is reported by proxy people in interviews. Therefore, I used a mortality correction. The results seem implausible and are included in the Appendix. The reason for the questionable results remains an open question. Future research should look for better mortality correction methods and methods to predict mortality from unknown cases. SHARE started looking for possibilities to link SHARE data to death registers although that remains difficult due to data sensitivity (Bergmann et al., 2020, p. 13). More steps in these directions are necessary.

Fourth, although using the country with the largest migrant sample, my sample of people classified as migrants is small. This is problematic as I divided the full analysis sample into age groups. Particularly, there are only a few observations in the oldest age groups. However, I used modelled probabilities to diminish the influence of this problem. Due to the small migrant sample, it was not possible to distinguish between different countries of origin. Furthermore, it hindered the investigation of the role of the type of migration. Migrants in Germany are a diverse group from different countries consisting of labour migrants, repatriates and refugees. Therefore, the results might look different for different groups. For instance, it is known that labour migrants from Turkey or the former Yugoslavia live in more disadvantaged areas which might be related to longer time in bad health (Steinbach, 2018, p. 295). Future research should incorporate these differences although this remains challenging regarding data and sample size.

One solution to increase the migrant sample would be to oversample them. For analyses, it would then be necessary to re-weighting them.

Sixth, I rely on foreign-born status to define migrants. SHARE data does not include a better health measure. That is problematic for the following reasons: First, people born abroad are not necessarily migrants. They could be children of tourists or of people who only worked abroad for a short time. Furthermore, this definition excludes second- and third-generation migrants. Future survey studies investigating health inequalities should include more fine-grained migration status measures.

5.5 Recommendations for the non-academic Audience

It is impossible to derive strong policy recommendations because my analysis only generally investigates morbidity and mortality differences between migrants and natives. I did not analyse the specific causes explaining these differences such as specific diseases. However, my study helps to understand the demand for healthcare facilities. The knowledge about migrants living longer in bad health than natives shows it is important to offer enough healthcare facilities the more migrants come. This might be particularly important for refugees who might be traumatized.

The strong contribution to the HLE from migrants' higher incidence and lower recovery probabilities clarifies the importance that healthcare providers must investigate the reasons for migrant patients' disease onset and worse recovery. It might be necessary to consider migrants' and refugees' backgrounds in healthcare adequately, for instance, refugees' traumas history. Migrants are less likely to use specialist care, medication use, therapist consultations and counselling, rehabilitation and disease prevention as shown by Klein and Knesebeck (2018). Therefore, migrants should be particularly addressed and informed about health care and prevention as they are less familiar with it than natives.

Another preventing factor might be to decrease migrants' exposure to unfavourable working conditions decreasing their risk of future health problems (Hargreaves et al., 2019; Ronda Pérez et al., 2012).

More detailed research is needed for more specific recommendations.

6 References

Abraído-Lanza, A. F., Dohrenwend, B. P., Ng-Mak, D. S., & Turner, J. B. (1999). The Latino mortality paradox: A test of the "salmon bias" and healthy migrant hypotheses. *American Journal of Public Health*, *89*(10), 1543–1548. https://doi.org/10.2105/AJPH.89.10.1543

Aichberger, M. C., Schouler-Ocak, M., Mundt, A., Busch, M. A., Nickels, E., Heimann, H. M., Ströhle, A., Reischies, F. M., Heinz, A., & Rapp, M. A. (2010). Depression in middle-aged and older first generation migrants in Europe: Results from the Survey of Health, Ageing and Retirement in Europe (SHARE). *European Psychiatry : The Journal of the Association of European Psychiatrists*, *25*(8), 468–475. https://doi.org/10.1016/j.eurpsy.2009.11.009

Arenas, E., Goldman, N., Pebley, A. R., & Teruel, G. (2015). Return Migration to Mexico: Does Health Matter? *Demography*, *52*(6), 1853–1868. https://doi.org/10.1007/s13524-015-0429-7

Auguie, B. (2017). *gridExtra: Miscellaneous Functions for "Grid" Graphics.* https://doi.org/10.32614/CRAN.package.gridExtra

Baćak, V., & Ólafsdóttir, S. (2017). Gender and validity of self-rated health in nineteen European countries. *Scandinavian Journal of Public Health*, *45*(6), 647–653. https://doi.org/10.1177/1403494817717405

Bergmann, M., Birkenbach, T., & Groh, R. (2020). *Possibilities to Deal with Unknown Vital Status in the Survey of Health, Ageing and Retirement in Europe (SHARE)* (MEA DISCUSSION PAPERS 22-2020). Munich Center for the Economics of Aging (mea). https://doi.org/10.2139/ssrn.3768571

Bergmann, Michael: Kneip, Thorsten, Luca, G. de, & Scherpenzeel, A. (2019). *Survey participation in the Survey of Health, Ageing and Retirement in Europe (SHARE), Wave 1-7. Based on Release 7.0.0.* (SHARE Working Paper Series 41-2019).

Boardman, J. D. (2006). Self-rated health among U.S. Adolescents. *The Journal of Adolescent Health : Official Publication of the Society for Adolescent Medicine*, *38*(4), 401–408. https://doi.org/10.1016/j.jadohealth.2005.01.006

Bos, V., Kunst, A. E., Keij-Deerenberg, I. M., Garssen, J., & Mackenbach, J. P. (2004). Ethnic inequalities in age- and cause-specific mortality in The Netherlands. *International Journal of Epidemiology*, *33*(5), 1112–1119. https://doi.org/10.1093/ije/dyh189

Bostean, G. (2013). Does selective migration explain the Hispanic paradox? A comparative analysis of Mexicans in the U.S. And Mexico. *Journal of Immigrant and Minority Health*, *15*(3), 624–635. https://doi.org/10.1007/s10903-012-9646-y

Boulogne, R., Jougla, E., Breem, Y., Kunst, A. E., & Rey, G. (2012). Mortality differences between the foreign-born and locally-born population in France (2004-2007). *Social Science & Medicine (1982)*, *74*(8), 1213–1223. https://doi.org/10.1016/j.socscimed.2012.01.002

Bourbeau, R. (2002). L'effet de la « sélection d'immigrants en bonne santé » sur la mortalité canadienne aux grands âges [The "healthy immigrant effect" on Canadian mortality in advanced old age]. *Cahiers québécois de démographie*, *31*(2), 249–274. https://doi.org/10.7202/000667ar

Carmada, C., & Riffe, T [Tim]. (2015). *MortalitySmooth — Smoothing and Forecasting Poisson Counts with P-Splines*. https://github.com/timriffe/MortalitySmooth

Carnein, M., Milewski, N., Doblhammer, G., & Nussfelder, W. J. (2015). Health inequalities of immigrants: Patterns and determinants of health expectancies of Turkish migrants living in Germany. In G. Doblhammer (Ed.), *Beiträge zur Bevölkerungswissenschaft:*

Vol. 46. Health among the Elderly in Germany: New evidence on disease, disability and care need (1. Aufl., pp. 157–190). Budrich, Barbara.
Caswell, H. (2000). *Matrix population models* (Vol. 1). Sinauer.
Chandola, T., & Jenkinson, C. (2000). Validating self-rated health in different ethnic groups. *Ethnicity & Health*, *5*(2), 151–159. https://doi.org/10.1080/713667451
Chiswick, B. R., Lee, Y. L. W., & Miller, P. W. (2008). Immigrant Selection Systems and Immigrant Health. *Contemporary Economic Policy*, *26*(4), 555–578. https://doi.org/10.1111/j.1465-7287.2008.00099.x
Cox, B., van Oyen, H., Cambois, E., Jagger, C., Le Roy, S., Robine, J.-M., & Romieu, I. (2009). The reliability of the Minimum European Health Module. *International Journal of Public Health*, *54*(2), 55–60. https://doi.org/10.1007/s00038-009-7104-y
Cullati, S., Mukhopadhyay, S., Sieber, S., Chakraborty, A., & Burton-Jeangros, C. (2018). Is the single self-rated health item reliable in India? A construct validity study. *BMJ Global Health*, *3*(6), e000856. https://doi.org/10.1136/bmjgh-2018-000856
Darmon, N., & Khlat, M [M.] (2001). An overview of the health status of migrants in France, in relation to their dietary practices. *Public Health Nutrition*, *4*(2), 163–172. https://doi.org/10.1079/PHN200064
De Maio, F. G., & Kemp, E. (2010). The deterioration of health status among immigrants to Canada. *Global Public Health*, *5*(5), 462–478. https://doi.org/10.1080/17441690902942480
DeSalvo, K. B., Bloser, N., Reynolds, K., He, J., & Muntner, P. (2006). Mortality prediction with a single general self-rated health question. A meta-analysis. *Journal of General Internal Medicine*, *21*(3), 267–275. https://doi.org/10.1111/j.1525-1497.2005.00291.x
DeSalvo, K. B., Fisher, W. P., Tran, K., Bloser, N., Merrill, W., & Peabody, J. (2006). Assessing measurement properties of two single-item general health measures. *Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *15*(2), 191–201. https://doi.org/10.1007/s11136-005-0887-2
DeSalvo, K. B., Jones, T. M., Peabody, J., McDonald, J., Fihn, S., Fan, V., He, J., & Muntner, P. (2009). Health care expenditure prediction with a single item, self-rated health measure. *Medical Care*, *47*(4), 440–447. https://doi.org/10.1097/MLR.0b013e318190b716
Diaz, C. J., Koning, S. M., & Martinez-Donate, A. P. (2016). Moving Beyond Salmon Bias: Mexican Return Migration and Health Selection. *Demography*, *53*(6), 2005–2030. https://doi.org/10.1007/s13524-016-0526-2
Dudel, C. (2021). Expanding the Markov Chain Toolbox: Distributions of Occupation Times and Waiting Times. *Sociological Methods & Research*, *50*(1), 401–428. https://doi.org/10.1177/0049124118782541
Dudel, C., & Li, P. (2024). *dtms - An R package for discrete-time multistate models*. https://github.com/christiandudel/dtms
Garcia, M. A., & Chiu, C.-T. (2016). Age at migration and disability-free life expectancy among the elder Mexican-origin population. *Demographic Research*, *35*(51), 1523–1536. https://doi.org/10.4054/DemRes.2016.35.51
Gohel, D., & Skintzos, P. (2024). *flextable: Functions for Tabular Reporting.* https://doi.org/10.32614/CRAN.package.flextable
Gubernskaya, Z. (2015). Age at migration and self-rated health trajectories after age 50: Understanding the older immigrant health paradox. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *70*(2), 279–290. https://doi.org/10.1093/geronb/gbu049
Guillot, M., Khlat, M [Myriam], Elo, I., Solignac, M., & Wallace, M. (2018). Understanding age variations in the migrant mortality advantage: An international comparative perspective. *PloS One*, *13*(6), 1-17. https://doi.org/10.1371/journal.pone.0199669

Guillot, M., Khlat, M [Myriam], Gansey, R., Solignac, M., & Elo, I. (2023). Return Migration Selection and Its Impact on the Migrant Mortality Advantage: New Evidence Using French Pension Data. *Demography*, *60*(5), 1335–1357. https://doi.org/10.1215/00703370-10938784

Halford, C., Wallman, T., Welin, L., Rosengren, A., Bardel, A., Johansson, S., Eriksson, H., Palmer, E., Wilhelmsen, L., & Svärdsudd, K. (2012). Effects of self-rated health on sick leave, disability pension, hospital admissions and mortality. A population-based longitudinal study of nearly 15,000 observations among Swedish women and men. *BMC Public Health*, *12*, 1103. https://doi.org/10.1186/1471-2458-12-1103

Hargreaves, S., Rustage, K., Nellums, L. B., McAlpine, A., Pocock, N., Devakumar, D., Aldridge, R. W., Abubakar, I., Kristensen, K. L., Himmels, J. W., Friedland, J. S., & Zimmerman, C. (2019). Occupational health outcomes among international migrant workers: A systematic review and meta-analysis. *The Lancet. Global Health*, *7*(7), e872-e882. https://doi.org/10.1016/S2214-109X(19)30204-9

Horiuchi, S., Wilmoth, J. R., & Pletcher, S. D. (2008). A decomposition method based on a model of continuous change. *Demography*, *45*(4), 785–801. https://doi.org/10.1353/dem.0.0033

Huang, G., Guo, F., Taksa, L., Cheng, Z., Tani, M., Liu, L., Zimmermann, K. F., & Franklin, M. (2024). Decomposing the differences in healthy life expectancy between migrants and natives: the 'healthy migrant effect' and its age variations in Australia. *Journal of Population Research*, *41*(1). https://doi.org/10.1007/s12546-023-09325-8

Ichou, M., & Wallace, M. (2019). The Healthy Immigrant Effect: The role of educational selectivity in the good health of migrants. *Demographic Research*, *40*, 61–94. https://www.jstor.org/stable/26726993

Imai, K., & Soneji, S. (2007). On the Estimation of Disability-Free Life Expectancy: Sullivan' Method and Its Extension. *Journal of the American Statistical Association*, *102*(480), 1199–1211. https://doi.org/10.1198/016214507000000040

International Organization for Migration. (2009). *Migration health: better health for all in Europe.* International Organization for Migration.

Kennedy, S., Kidd, M. P., McDonald, J. T., & Biddle, N. (2015). The Healthy Immigrant Effect: Patterns and Evidence from Four Countries. *Journal of International Migration and Integration*, *16*(2), 317–332. https://doi.org/10.1007/s12134-014-0340-x

Khlat, M [Myriam], & Courbage, Y. (1996). Mortality and causes of death of Moroccans in France, 1979-91. *Population: An English Selection*, *8*, 59–94. https://www.jstor.org/stable/2949157

Kim, I.-H., Carrasco, C., Muntaner, C., McKenzie, K., & Noh, S. (2013). Ethnicity and postmigration health trajectory in new immigrants to Canada. *American Journal of Public Health*, *103*(4), e96-104. https://doi.org/10.2105/AJPH.2012.301185

Klein, J., & Knesebeck, O. von dem (2018). Inequalities in health care utilization among migrants and non-migrants in Germany: A systematic review. *International Journal for Equity in Health*, *17*(1), 160. https://doi.org/10.1186/s12939-018-0876-z

Kouris-Blazos, A. (2002). Morbidity mortality paradox of 1st generation Greek Australians. *Asia Pacific Journal of Clinical Nutrition*, *11 Suppl 3*, S569-75. https://doi.org/10.1046/j.1440-6047.11.supp3.2.x

Lanari, D., & Bussini, O. (2012). International migration and health inequalities in later life. *Ageing and Society*, *32*(6), 935–962. https://doi.org/10.1017/S0144686X11000730

Law, M., & Wald, N. (1999). Why heart disease mortality is low in France: The time lag explanation. *BMJ (Clinical Research Ed.)*, *318*(7196), 1471–1476. https://doi.org/10.1136/bmj.318.7196.1471

Markides, K. S., & Eschbach, K. (2011). Hispanic Paradox in Adult Mortality in the United States. In R. G. Rogers & E. M. Crimmins (Eds.), *International handbooks of*

population: vol. 2. International handbook of adult mortality (pp. 227–240). Springer Science+Business Media B.V. https://doi.org/10.1007/978-90-481-9996-9_11
McGee, D. L., Liao, Y., Cao, G., & Cooper, R. S. (1999). Self-reported health status and mortality in a multiethnic US cohort. *American Journal of Epidemiology*, *149*(1), 41–46. https://doi.org/10.1093/oxfordjournals.aje.a009725
Moretti, M., Riffe, T [Timothy], & Lorenti, A. (2023a). *Multistate analysis and decomposition of disability-free life expectancy trends in Italy 2004-2019.* https://doi.org/10.4054/MPIDR-WP-2023-030
Moretti, M., Riffe, T [Timothy], & Lorenti, A. (2023b). *R code for Multistate Analysis and Decomposition of Disability Life Expectancy Trends in Italy, 2004-2019*. https://osf.io/fm7zw/?view_only=44b4dd51398f4e86831d276e021f06a2 https://doi.org/None
Namer, Y., & Razum, O. (2016). Convergence Theory and the Salmon Effect in Migrant Health. In Y. Namer & O. Razum (Eds.), *Oxford Research Encyclopedia of Global Public Health.* Oxford University Press. https://doi.org/10.1093/acrefore/9780190632366.013.17
Newbold, K. B. (2005). Self-rated health within the Canadian immigrant population: risk and the healthy immigrant effect. *Social Science & Medicine*, *60*(6), 1359–1370. https://doi.org/10.1016/j.socscimed.2004.06.048
Palloni, A., & Arias, E. (2004). Paradox lost: Explaining the Hispanic adult mortality advantage. *Demography*, *41*(3), 385–415. https://doi.org/10.1353/dem.2004.0024
Patrick Deboosere, & Sylvie Gadeyne (2005). Adult Migrant Mortality Advantage in Belgium: Evidence Using Census and Register Data. *Population*, *60*(5), 655–698. https://www.cairn.info/revue-population-english-2005-5-page-655.htm
Perneger, T. V., Gayet-Ageron, A., Courvoisier, D. S., Agoritsas, T., & Cullati, S. (2013). Self-rated health: Analysis of distances and transitions between response options. *Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *22*(10), 2761–2768. https://doi.org/10.1007/s11136-013-0418-5
Razum, O. (2008). Migrant Mortality, Healthy Migrant Effect. In W. Kirch (Ed.), *Encyclopedia of public health: A-H* (pp. 932–935). Springer. https://doi.org/10.1007/978-1-4020-5614-7_2188
Razum, O., Zeeb, H., Akgün, H. S., & Yilmaz, S. (1998). Low overall mortality of Turkish residents in Germany persists and extends into a second generation: Merely a healthy migrant effect? *Tropical Medicine & International Health : TM & IH*, *3*(4), 297–303. https://doi.org/10.1046/j.1365-3156.1998.00233.x
Rechel, B., Mladovsky, P., Ingleby, D., Mackenbach, J. P., & McKee, M. (2013). Migration and health in an increasingly diverse Europe. *The Lancet*, *381*(9873), 1235–1245. https://doi.org/10.1016/S0140-6736(12)62086-8
Reus-Pons, M., Kibele, E. U. B., & Janssen, F. (2017). Differences in healthy life expectancy between older migrants and non-migrants in three European countries over time. *International Journal of Public Health*, *62*(5), 531–540. https://doi.org/10.1007/s00038-017-0949-6
Reus-Pons, M., Mulder, C. H., Kibele, E. U. B., & Janssen, F. (2018). Differences in the health transition patterns of migrants and non-migrants aged 50 and older in southern and western Europe (2004-2015). *BMC Medicine*, *16*(1), 1–15. https://doi.org/10.1186/s12916-018-1044-4
Riosmena, F., Wong, R., & Palloni, A. (2013). Migration selection, protection, and acculturation in health: A binational perspective on older adults. *Demography*, *50*(3), 1039–1064. https://doi.org/10.1007/s13524-012-0178-9

Ronda Pérez, E., Benavides, F. G., Levecque, K., Love, J. G., Felt, E., & van Rossem, R. (2012). Differences in working conditions and employment arrangements among migrant and non-migrant workers in Europe. *Ethnicity & Health*, *17*(6), 563–577. https://doi.org/10.1080/13557858.2012.730606

Rubalcava, L. N., Teruel, G. M., Thomas, D., & Goldman, N. (2008). The healthy migrant effect: New findings from the Mexican Family Life Survey. *American Journal of Public Health*, *98*(1), 78–84. https://doi.org/10.2105/AJPH.2006.098418

Ruiz, J. M., Steffen, P., & Smith, T. B. (2013). Hispanic mortality paradox: A systematic review and meta-analysis of the longitudinal literature. *American Journal of Public Health*, *103*(3), e52-60. https://doi.org/10.2105/AJPH.2012.301103

Schimany, P., Rühl, S., & Kohls, M. (2012). *Ältere Migrantinnen und Migranten. Entwicklungen, Lebenslagen, Perspektiven: Forschungsbericht 18.* Bundesamt für Migration und Flüchtlinge. http://www.naschdom.de/files/_bilder/Fremndspr.%20Material/fb18-aeltere-migranten-1.pdf

Seo, S., Chung, S., & Shumway, M. (2014). How good is "very good"? Translation effect in the racial/ethnic variation in self-rated health status. *Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *23*(2), 593–600. https://doi.org/10.1007/s11136-013-0522-6

Solé-Auró, A., & Crimmins, E. M. (2008). Health of Immigrants in European countries. *The International Migration Review*, *42*(4), 861–876. https://doi.org/10.1111/j.1747-7379.2008.00150.x

Statistisches Bundesamt. (2024a, April 2). *Population in private households by immigration history by sex*. https://www.destatis.de/EN/Themes/Society-Environment/Population/Migration-Integration/Tables/immigration-by-sex.html#644566

Statistisches Bundesamt. (2024b, April 10). *Migration between Germany and foreign countries, 1950 to 2023*. https://www.destatis.de/EN/Themes/Society-Environment/Population/Migration/Tables/migration-year-01.html

Statistisches Bundesamt. (2024c, June 14). *Average life expectancy (period life table): Germany, period of years, sex, completed age*. https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=12621-0002&zeitscheiben=15&sachmerkmal=ALT577&sachschluessel=ALTVOLL000,ALTVOLL020,ALTVOLL040,ALTVOLL060,ALTVOLL065,ALTVOLL080&language=en#abreadcrumb

Steinbach, A. (2018). Older Migrants in Germany. *Journal of Population Ageing*, *11*(3), 285–306. https://doi.org/10.1007/s12062-017-9183-5

Survey of Health, Ageing and Retirement in Europe. (2020, June 26). *Conditions of Use.* SHARE-ERIC. https://share-eric.eu/data/data-access/conditions-of-use

Survey of Health, Ageing and Retirement in Europe. (2024a). *SHARE Research Data Center*. https://releases.sharedataportal.eu/login?redirect=%2Fusers%2Flogin

Survey of Health, Ageing and Retirement in Europe. (2024b, June 13). *SHARE wave 1 documentation: dn004_*. https://www.share-datadocutool.org/data-variables/view/20206

Survey of Health, Ageing and Retirement in Europe. (2024c, June 13). *SHARE wave 2 documentation: dn004_*. https://www.share-datadocutool.org/data-variables/view/22288

Survey of Health, Ageing and Retirement in Europe. (2024d, June 13). *SHARE wave 5 documentation: dn004_*. https://www.share-datadocutool.org/data-variables/view/17829

Tarnutzer, S., & Bopp, M. (2012). Healthy migrants but unhealthy offspring? A retrospective cohort study among Italians in Switzerland. *BMC Public Health*, *12*, 1104. https://doi.org/10.1186/1471-2458-12-1104

Tezcan, T. (2019). What initiates, what postpones return migration intention? The case of Turkish immigrants residing in Germany. *Population, Space and Place*, *25*(3), Article e2175. https://doi.org/10.1002/psp.2175

van Mol, C., & Valk, H. de. (2016). Migration and Immigrants in Europe: A Historical and Demographic Perspective. In B. Garcés-Mascareñas & R. Penninx (Eds.), *IMISCOE Research Series. Integration Processes and Policies in Europe* (pp. 31–55). Springer international publishing. https://doi.org/10.1007/978-3-319-21674-4_3

Wallace, M., & Kulu, H. (2014). Low immigrant mortality in England and Wales: A data artefact? *Social Science & Medicine (1982)*, *120*, 100–109. https://doi.org/10.1016/j.socscimed.2014.08.032

Wallace, M., & Kulu, H. (2018). Can the salmon bias effect explain the migrant mortality advantage in England and Wales? *Population, Space and Place*, *24*(8), Article e2146. https://doi.org/10.1002/psp.2146

Wallace, M., & Wilson, B. (2019). Migrant Mortality Advantage Versus Origin and the Selection Hypothesis. *Population and Development Review*, *45*(4), 767–794. http://www.jstor.org/stable/45285989

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., . . . Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

World Health Organization. (2024, February 7). *Healthy Life Expectancy (HALE) at birth*. https://www.who.int/data/gho/indicator-metadata-registry/imr-details/66

Yee, T. W. (2015). *Vector Generalized Linear and Additive Models: With an Implementation in R* (1st ed. 2015). *Springer Series in Statistics*. Springer New York. https://doi.org/10.1007/978-1-4939-2818-7

Zheng, H., & Yu, W.-H. (2022). Diminished Advantage or Persistent Protection? A New Approach to Assess Immigrants' Mortality Advantages Over Time. *Demography*, *59*(5), 1655–1681. https://doi.org/10.1215/00703370-10175388

Zufferey, J. (2014). *Pourquoi les migrants vivent-ils plus longtemps? Analyse des différentiels de mortalité des populations suisse et étrangères 1990-2008* [Thèse de doctorat, Université de Genève, Genève]. DataCite.

Zufferey, J. (2016). Investigating the migrant mortality advantage at the intersections of social stratification in Switzerland. *Demographic Research*, *34*, Article 32, 899–926. https://doi.org/10.4054/DemRes.2016.34.32

zur Nieden, F., & Sommer, B. (2016). Life Expectancy in Germany Based on the 2011 Census: Was the Healthy Migrant Effect Merely an Artefact? *Comparative Population Studies - Zeitschrift Für Bevölkerungswissenschaft*, *41*(2), 145–174. https://doi.org/10.12765/CPoS-2016-05en (314 kB / Comparative Population Studies, Vol 41, No 2 (2016)).

Appendix

List of Figures and Tables of the Appendix

A Sample Selection

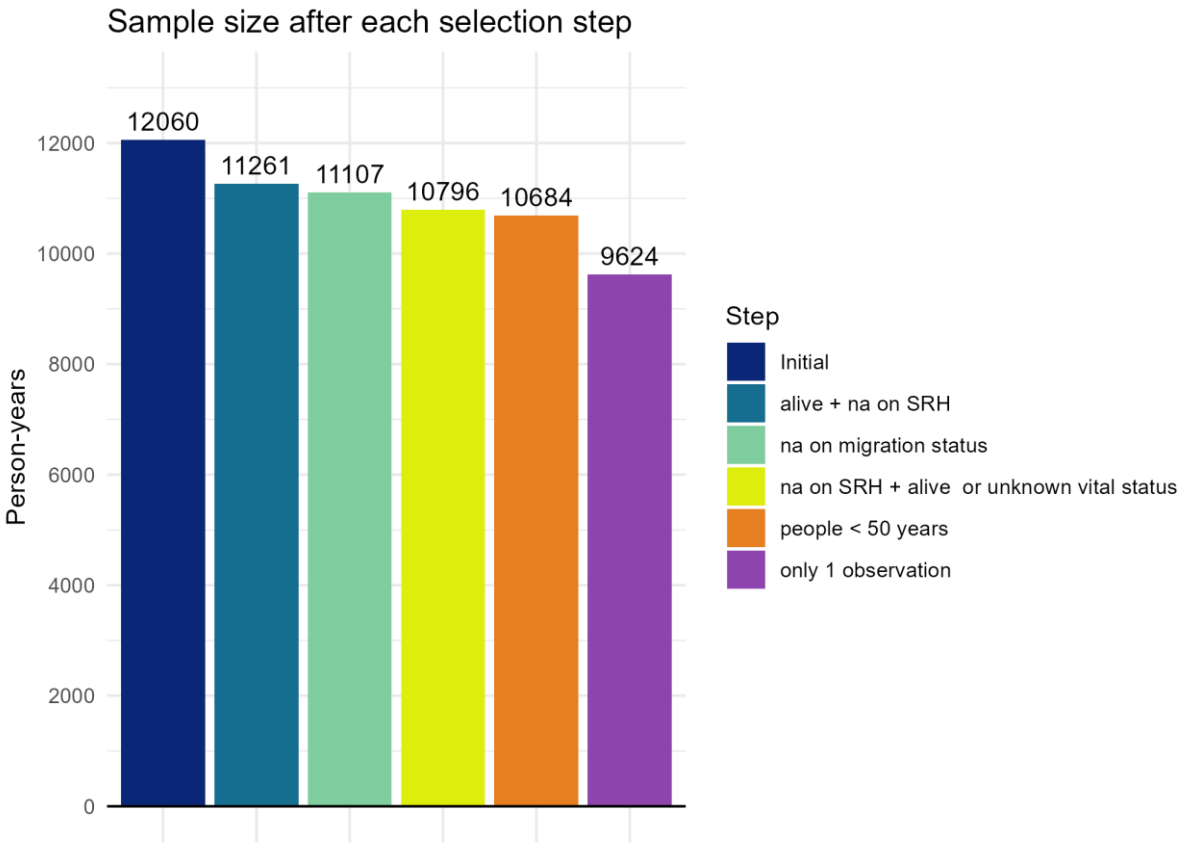

Figure A 1: Development of the sample size after each selection step
Source: own computations, data from SHARE

B Appendix for the unadjusted Probabilities

B.1 Death Probabilities from the unadjusted Probabilities (log scale)

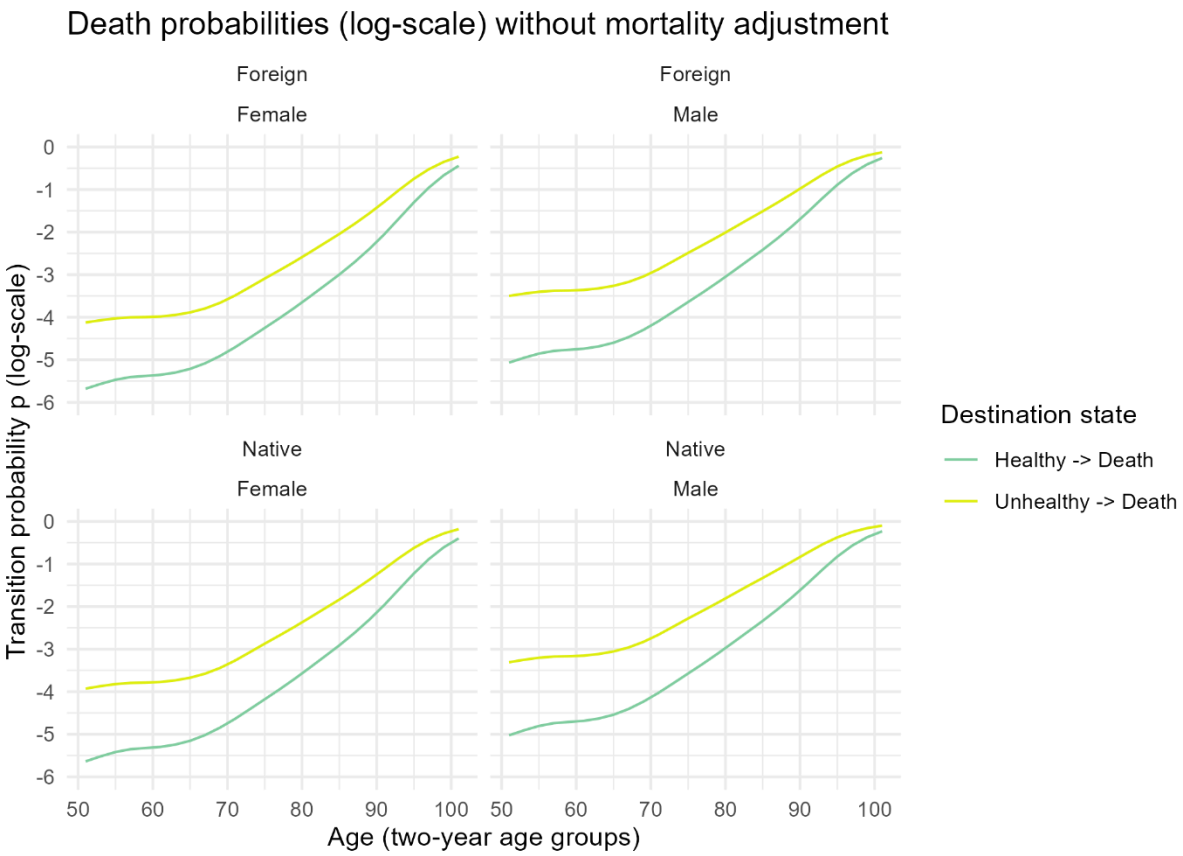

Figure A 2: mortality-unadjusted death probabilities (log scale)
Source: own computations, data from SHARE

B.2 Decomposition of Unhealthy Life Expectancy (mortality-unadjusted probabilities)

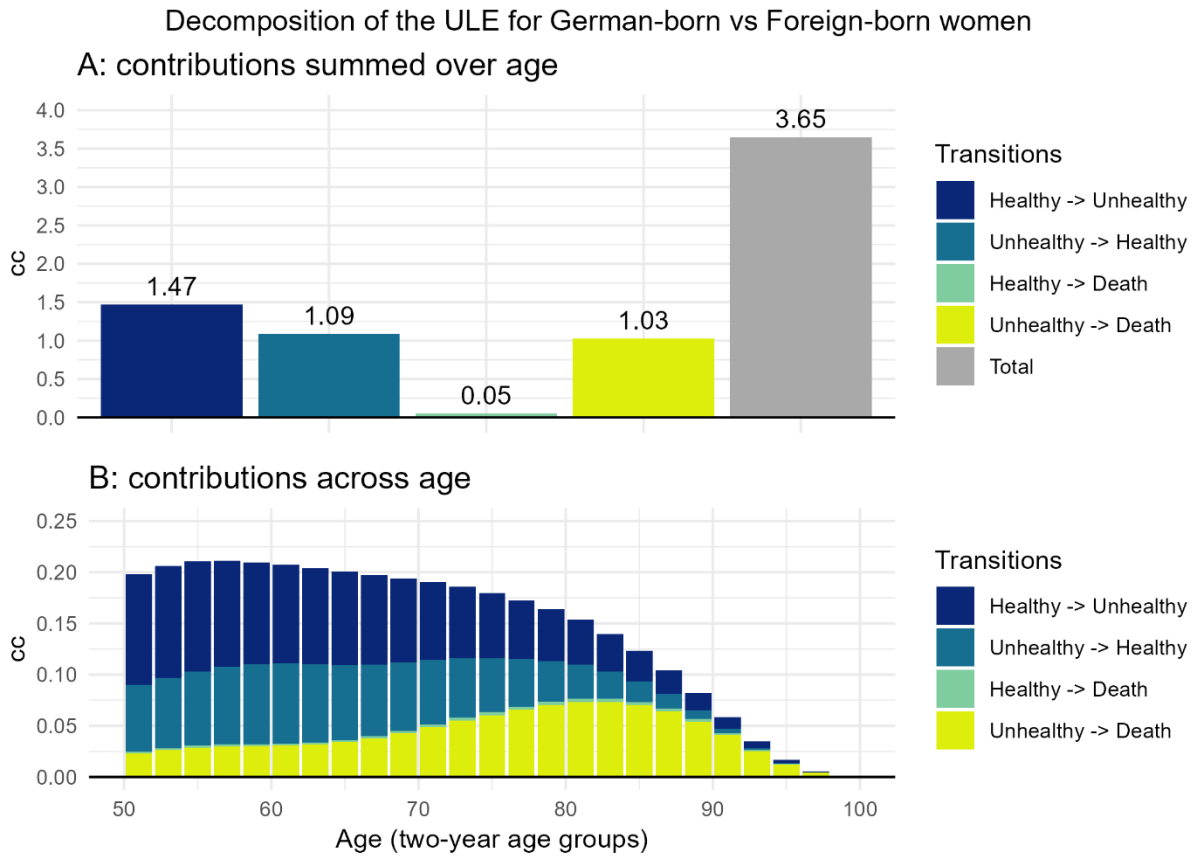

Figure A 3: Decomposition of Unhealthy Life Expectancy for women based on unadjusted probabilities
Source: own computations, data from SHARE

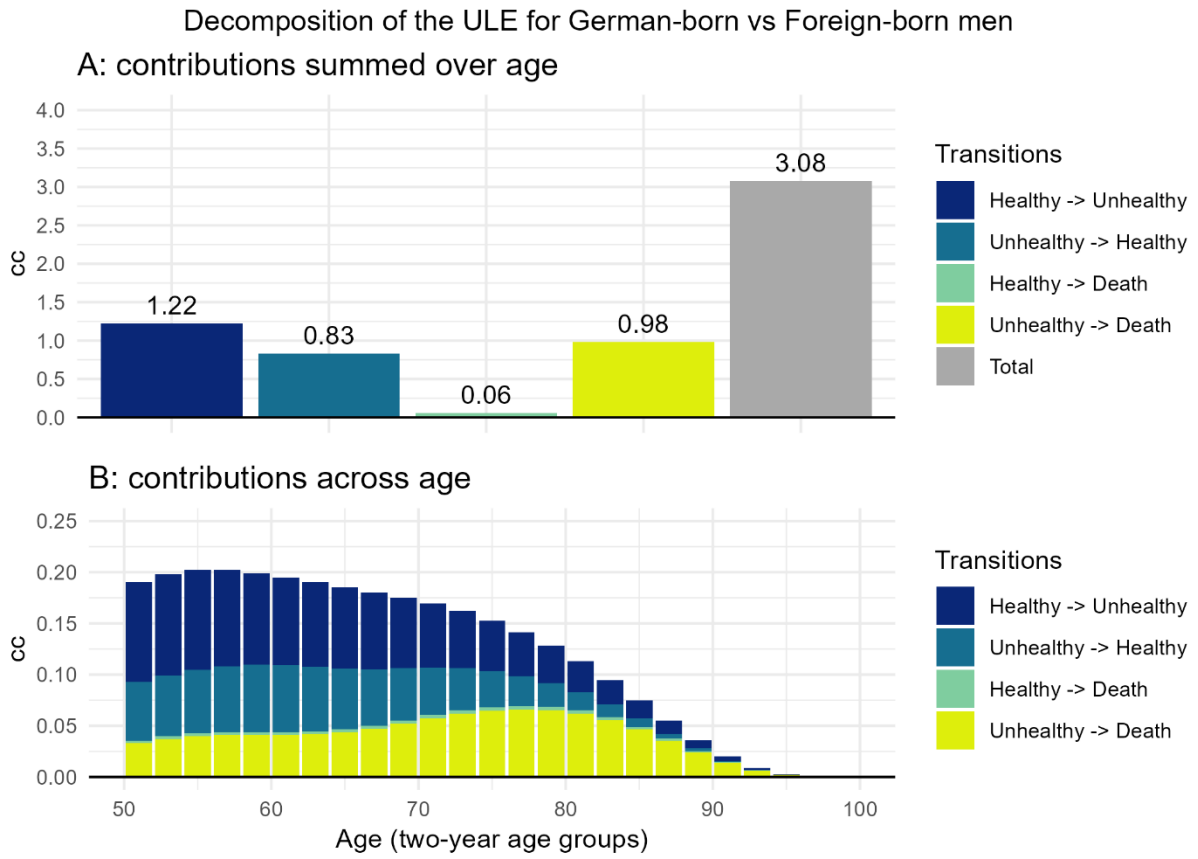

Figure A 4: Decomposition of Unhealthy Life Expectancy for men based on unadjusted probabilities
Source: own computations, data from SHARE

C All results based on the mortality-adjusted Probabilities

Although I do not describe the full results from the analyses based on the mortality-adjusted probabilities in the results section, I will briefly discuss them now.

Some findings need to be pointed out: After an exponential curve of all mortality probabilities, the probabilities decrease in the oldest age groups for natives, particularly men. This results in the mortality of migrants, particularly for men which is slightly higher and in a lower mortality for native men than women. Therefore, the life expectancies look accordingly distorted. Men are expected to live longer than women and natives longer than migrants. Particularly, native men have an implausible high life expectancy of around 36 years at age 50-51. Due to these discrepancies, I will refrain from discussing the decomposition results.

C.1 Adjusted Transition Probabilities

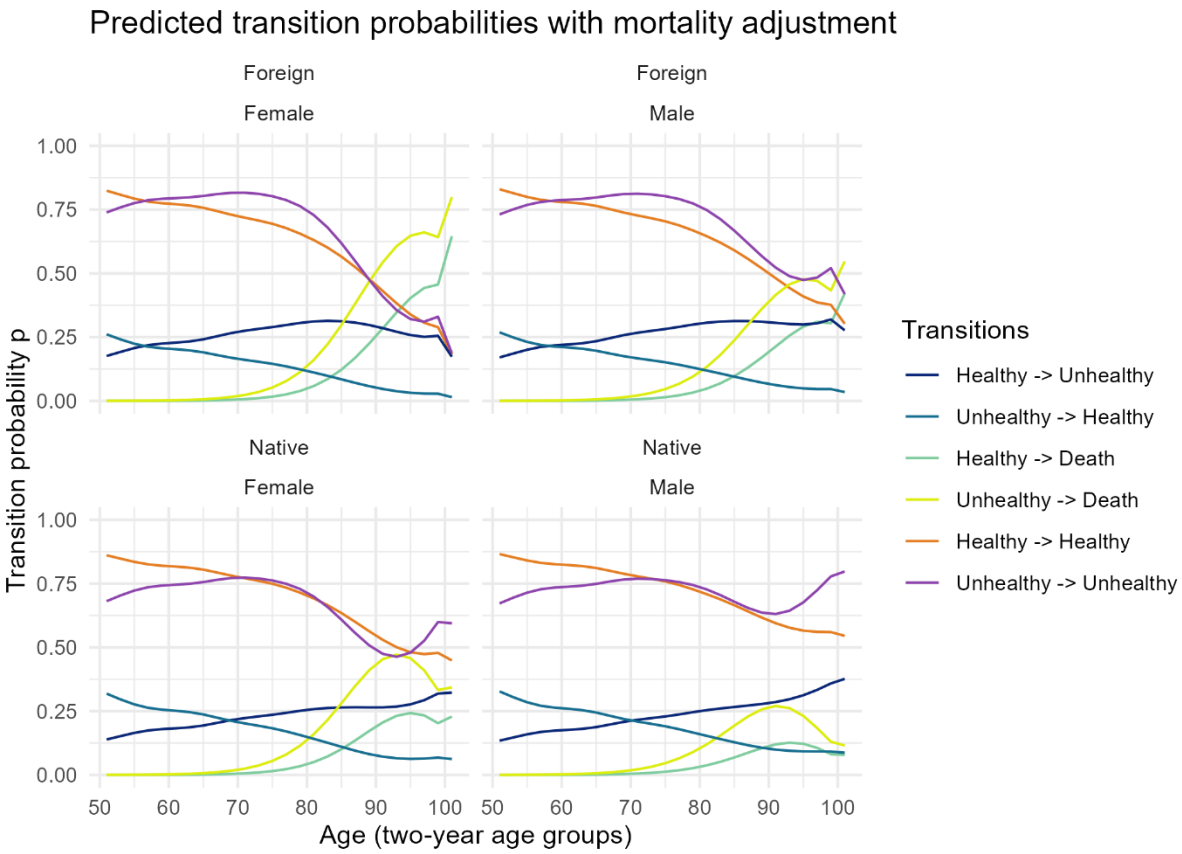

Figure A 5: mortality-adjusted transition probabilities
Source: own computations, data from SHARE

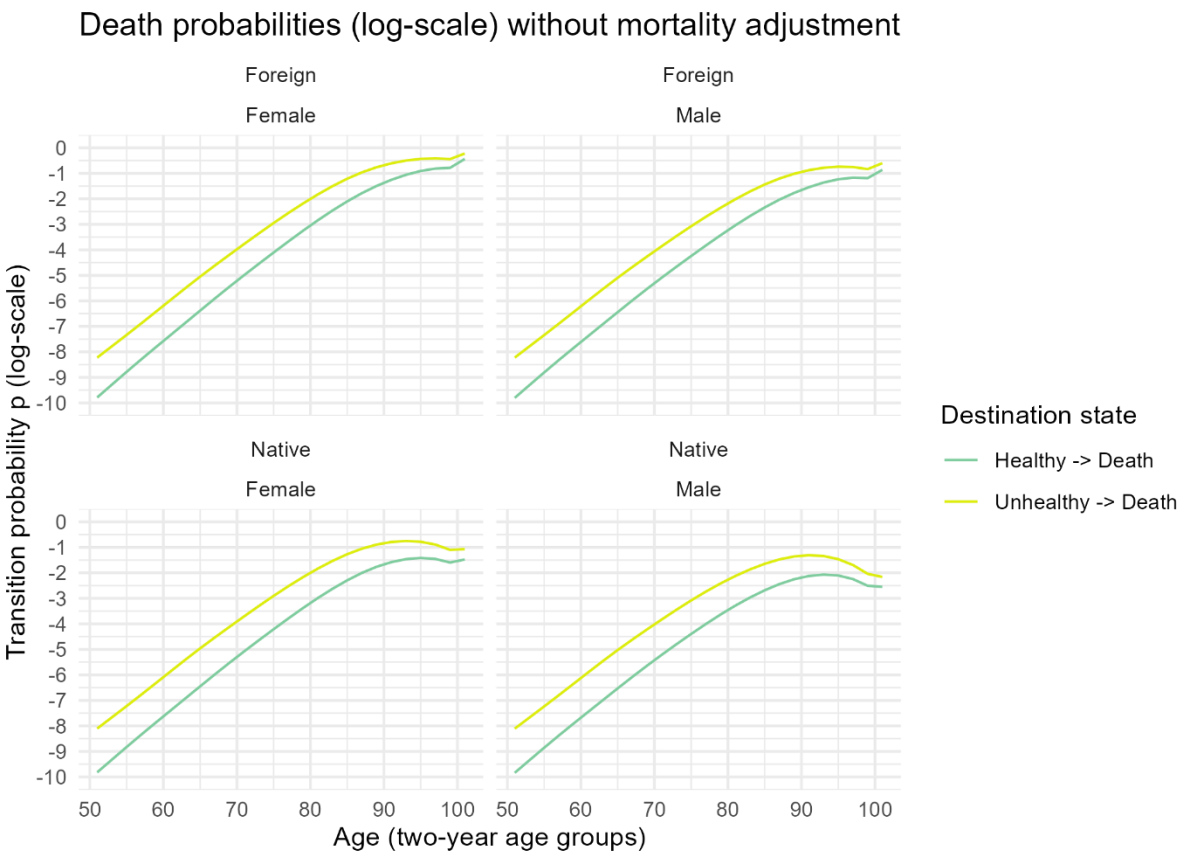

Figure A 6: adjusted death probabilities (log scale)
Source: own computations, data from SHARE

C.2 Life Expectancies from the adjusted Probabilities

Table A 1: Life expectancies for German-born women based on adjusted probabilities

(Healthy) Life expectancy for Female Natives

Starting States	HLE	ULE	LE
Healthy	19.17	14.18	33.35
Unhealthy	15.63	17.67	33.29
Average	16.65	16.66	33.31

Source: own computations; data from SHARE

Table A 2: Life expectancies for foreign-born women based on adjusted probabilities

(Healthy) Life expectancy for Female Foreign-born

Starting States	HLE	ULE	LE
Healthy	15.70	16.57	32.27
Unhealthy	12.05	20.17	32.22
Average	13.26	18.98	32.24

Source: own computations; data from SHARE

Table A 3: Life expectancies for German-born men based on adjusted probabilities

(Healthy) Life expectancy for Male Natives

Starting States	HLE	ULE	LE
Healthy	20.48	15.66	36.14
Unhealthy	16.97	19.11	36.08
Average	17.96	18.14	36.10

Source: own computations; data from SHARE

Table A 4: Life expectancies for foreign-born men based on adjusted probabilities

(Healthy) Life expectancy for Male Foreign-born			
Starting States	HLE	ULE	LE
Healthy	16.47	17.10	33.56
Unhealthy	12.83	20.68	33.51
Average	14.00	19.53	33.53

Source: own computations; data from SHARE

C.3 Decomposition based on the unadjusted Probabilities

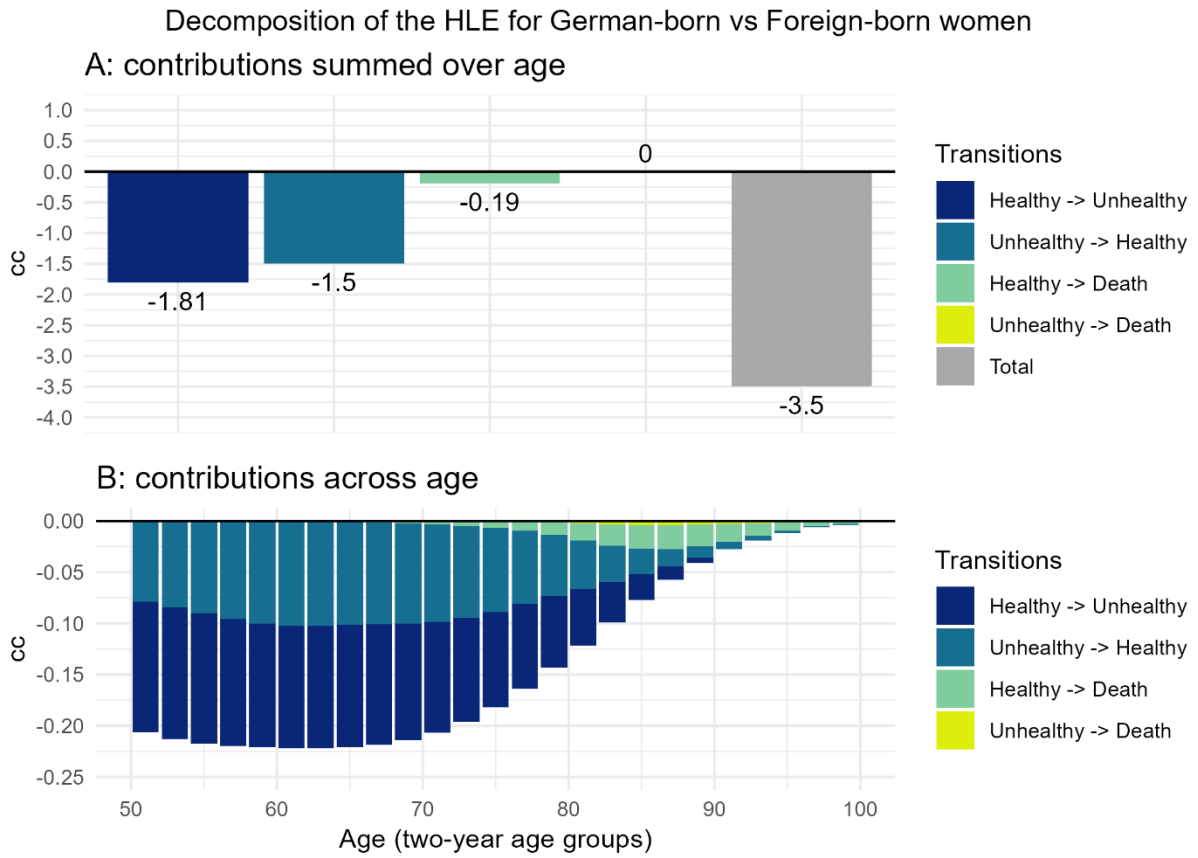

Figure A 7: Decomposition of the Healthy Life Expectancy for women (adjusted probabilities)
Source: own computations, data from SHARE

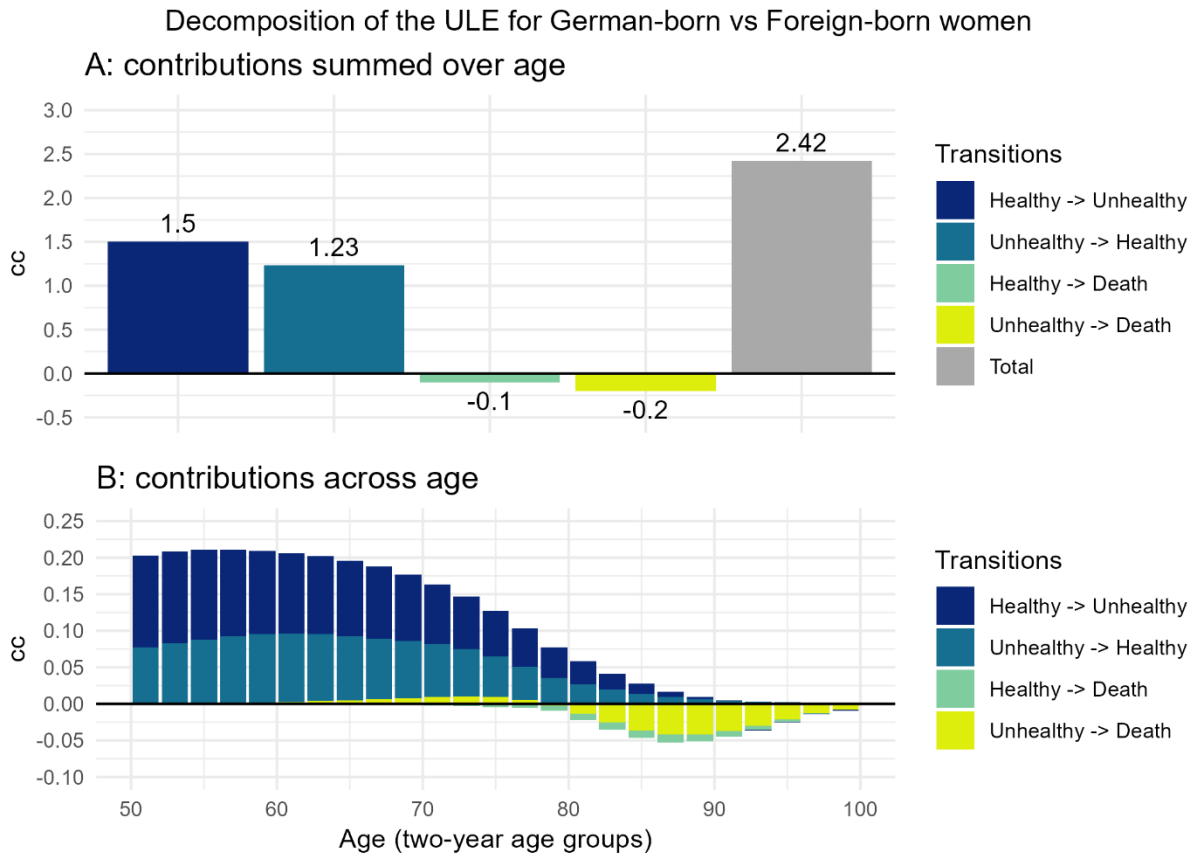

Figure A 8: Decomposition of the unhealthy Life Expectancy for women (adjusted probabilities)
Source: own computations, data from SHARE

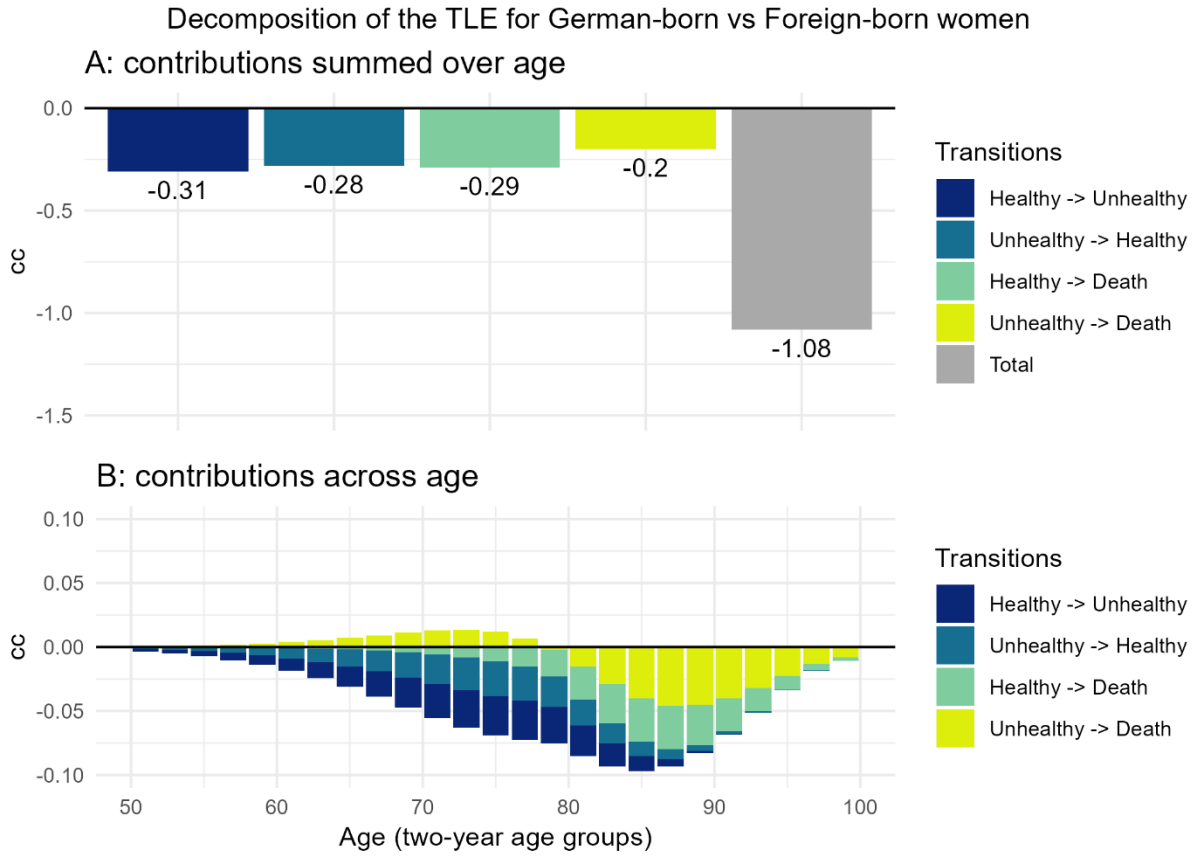

Figure A 9: Decomposition of the Total Life Expectancy for women (adjusted probabilities)
Source: own computations, data from SHARE

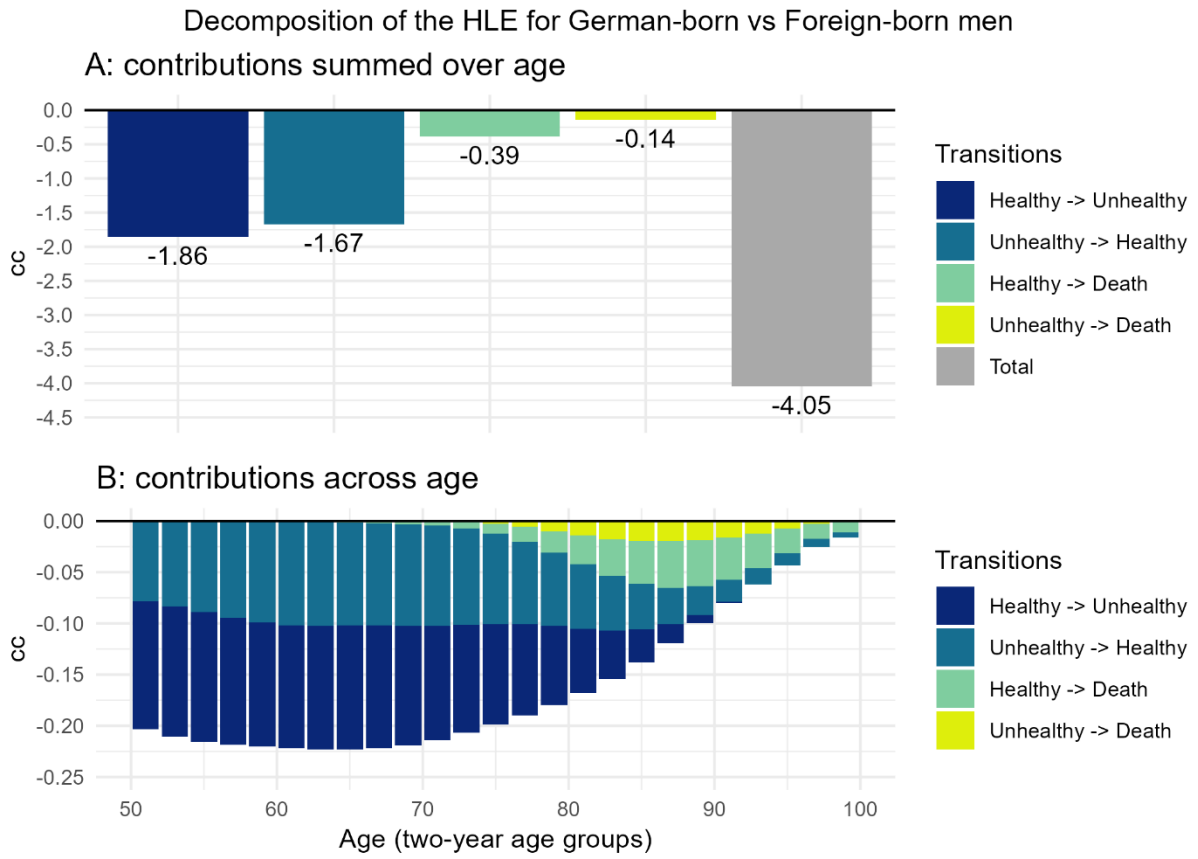

Figure A 10: Decomposition of the Healthy Life Expectancy for men (adjusted probabilities)
Source: own computations, data from SHARE

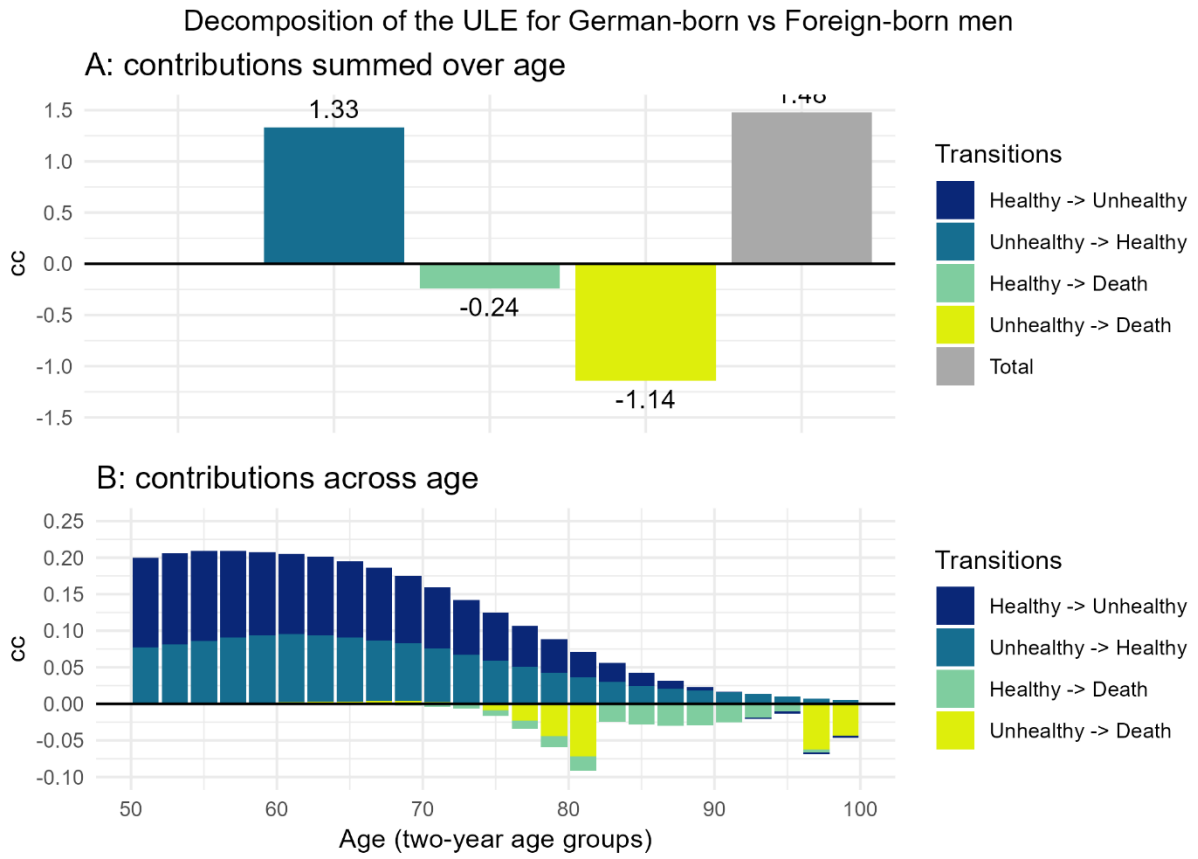

Figure A 11: Decomposition of the Unhealthy Life Expectancy for men (adjusted probabilities)
Source: own computations, data from SHARE

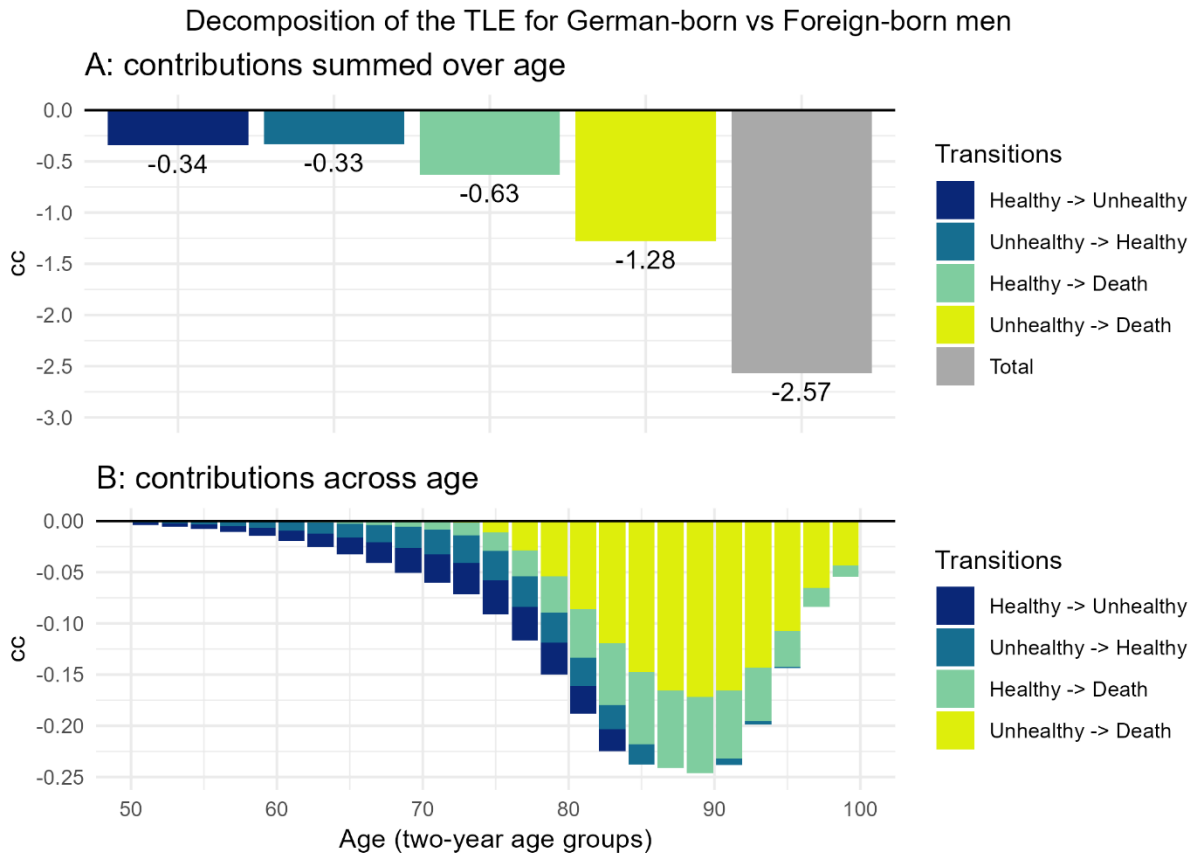

Figure A 12: Decomposition of the Total Life Expectancy for men (adjusted probabilities)
Source: own computations, data from SHARE