Carlotta Bochert

Student number S5352282

Frailty of Older Migrants in Europe – Gendered Dynamics?

Master's thesis

Primary supervisor: Max Frentz with Dr. Adrien Remund

Secondary supervisor: Prof. Dr. Karsten Hank

Study programme "Population Studies"

at the Department of Demography,

Faculty of Spatial Sciences,

Rijksuniversiteit Groningen

17th of November 2023

Abstract

European societies are ageing. It has therefore become an urgent question what can be done to enable ageing in health – both for personal well-being and fiscal sustainability. Ageing migrants are particularly vulnerable in terms of health. This study explores frailty differences between native and immigrant populations in Europe, aged 60+, with special attention paid to the intersection of migration status and gender. It uses 2015 data of almost 50.000 respondents from the SHARE project.

In logistic regression, immigrants are shown to be slightly more likely to be frail than natives in their country of residence. In separate models for men and women, the effect of migration is even more pronounced for women. For both men and women, economic factors play a large role in mediating immigrant-native frailty differences. Health behaviours and social capital measures show smaller mediation effects. After controlling for all of these mediators, immigrant-native differences in frailty are no longer statistically significant. The relation of immigrants' likelihood of frailty to that of their native counterparts varies across different regions of Europe, appearing most disadvantageous for migrants in Northern Europe.

These results contribute to the growing empirical literature that finds there to be no lasting "healthy migrant effect" for older migrants in Europe. This study's findings suggest that older migrants are disproportionately negatively affected by a variety of social determinants of health and thus need to be considered as an important target group for health interventions in the older population.

Keywords: Migration, gender, frailty, population ageing, health inequality, Europe, SHARE.

Contents

List of tables and figures

List of abbreviations

ADL	Activities of Daily Living
BMI	Body-Mass-Index
HDI	Human Development Index
IADL	Instrumental Activities of Daily Living
ICC	Intraclass Correlation Coefficient
KHB	Karlson-Holm-Breen
MEA	Munich Center for the Economics of Aging
MIPEX	Migrant Integration Policy Index
OADR	Old-Age Dependency Ratio
OR	Odds ratio
SHARE	Survey of Health, Ageing and Retirement in Europe
SHARE-ERIC	Survey of Health, Ageing and Retirement in Europe – European Research Infrastructure Consortium
SRH	Self-rated health
UNDP	United Nations Development Programme
UNSD	United Nations Statistical Division
VIF	Variance inflation factor

1. Introduction

European populations are ageing. It has therefore become an urgent question what can be done to enable ageing in health – both for personal well-being and fiscal sustainability (Ramos-Herrera & Sosvilla-Rivero, 2020; Rouzet et al., 2019). The older cohorts increasingly include those who migrated internationally to or within Europe over the second half of the 20th century (White, 2007). While migrants' health has been the focus of much academic work, with inconclusive and context-specific results (Vonneilich et al., 2021), how it develops in older ages is studied less (Cela & Di Barbiano Belgiojoso, 2021). Furthermore, migration studies have honed in on the different migration experiences of men and women in the last decades, citing differential reasons for migration and differences in the process of mobility and settling down (Fresnoza-Flot, 2022; Llácer et al., 2007). These differences presumably also show up in health in later life. Consequently, I study the gendered long-term effects of migration on old-age health.

1.1 Background and objective

Population ageing is a global phenomenon that is currently most advanced in 'more developed' regions, which includes Europe. The process can be seen as the end stage of the larger demographic transition from high fertility and mortality levels to low levels of both. In Europe, it is currently accelerated by the post-World War 2 "baby boom" generation reaching retirement age (Bloom & Luca, 2016). Along with low levels of fertility, often below the levels required for replacement of the population stock, increases in life expectancy drive population ageing. Older age is associated with health deterioration, and it is not yet clear whether increases in total life expectancy will always go hand in hand with improved old-age health, or increases in healthy (working) life expectancy (Loichinger & Weber, 2016; Parker et al., 2020). All of this leads to increased old-age dependency ratios (OADR), or the proportion of people over 65 to those of working age.

Increasing OADRs pose serious challenges for societies. For one, there are fiscal implications: Smaller proportions of the population at working age have to support larger proportions in retirement, consumption and saving patterns change, and asset values and economic growth develop accordingly (Bloom & Luca, 2016; Ramos-Herrera & Sosvilla-Rivero, 2020; Rouzet et al., 2019). Additionally, more people might be in worse health for longer, an unfortunate prospect for societies and their economies as well as individuals. This has spurred many investigations into "healthy" or "active ageing" (see, e.g., Malva et al., 2023).

For a long time, migration to Europe has been seen as a way to combat the effects of low fertility and mortality. With migrants being disproportionally of working age at the time of their migration, they especially increase the size of the labour force and thus change the overall population age structure. Furthermore, many immigrants to European countries show higher fertility levels than the non-migrated population. This leads to a secondary effect of increased overall birth rates compared to a hypothetical society without any immigration (Bloom & Luca, 2016). Of course, immigrants do not only count towards the population at the time of their migration. Even among those that migrated within a 'guest worker' scheme, which explicitly or implicitly expected them to return at some point in their lives when they were not economically useful to the host countries anymore, many stayed on (White, 2007). The large migrant cohorts of post-World War 2 Europe are now increasingly reaching retirement age in many Western European countries (Ciobanu et al., 2017; White, 2007). As such, migrant-specific policy cannot assume all migrants to be young and healthy; instead, the established programmes and institutions need to take into consideration this growing group of older migrants, which faces ageing-related health challenges like their non-migrated counterparts. Naturally, how immigrants in Europe age might be different from the non-migrated older population, requiring more or less intervention and specialised programmes, as are established for other sub-groups of immigrants (White, 2007).

Migrants are not, and have never been, a homogeneous group. Even though most of the migrants originally recruited through 'guest worker' schemes were men, they were often later joined by their spouses, leading to a mixed gender composition in the ageing migrant population today (Liversage & Mizrahi Mirdal, 2017), same as in the non-migrated population. Gender has been shown to be an important influence on health, including at older ages (Aner & Dosch, 2023; Gómez-Costilla et al., 2022; Hankivsky, 2012). While the results are not quite conclusive, they suggest that women's lower socio-economic position leads to a disadvantage in health in later life compared to men despite their higher life expectancy (Gómez-Costilla et al., 2022). However, gender's intersection with migrant status has not been studied extensively (Krobisch et al., 2021; Wandschneider et al., 2020).

Similarly, health differentials between migrants and non-migrants at older ages are understudied. Previous findings for younger migrants indicate health advantages, thought to be mainly due to selectivity in who decides to migrate and protective health behaviours and community norms (Jasso et al., 2004). There is also an indication that migration might have differential effects on health by gender (Fresnoza-Flot, 2022; Llácer et al., 2007; Trappolini & Giudici, 2021), potentially affecting both selectivity in choosing migration and adverse (post-)migration experiences. Over time spent in the destination country, health appears to converge between non-migrants and migrants (Jasso et al., 2004; Viruell-Fuentes, 2007). Generally, studies following up on immigrants' health long after their migration are rarer than those looking at the time of migration and its immediate aftermath or at intergenerational transmission of (dis)advantage. Since a big proportion of migrants migrate at (young) working ages, this means that studies on health of older migrants are scarce. Additionally, destination country contexts shape selectivity and (post-)migration experience as well, meaning that the large body of work on migration and health in a US immigration context may not be easily applicable to Europe, reducing the available information for this region further. This work is intended to contribute to this sparse field of literature.

Health itself can refer to a broad scope of dimensions of physical and mental well-being. In a complex societal system like migration, exploring effects on different dimensions is important. In quantitative research, respondents are often simply asked to rate their own health. However, especially in the context of migration, this might come with problems in comparability across cultural understandings of health (Lazarevič, 2023). For this work, I use the concept of frailty, which refers to "a biologic syndrome of decreased reserve and resistance to stressors, resulting from cumulative declines across multiple physiologic systems and causing vulnerability to adverse outcomes" (Sirven et al., 2020, p. 716). Within Verbrugge and Jette's (1994) conceptual scheme of the disablement process, frailty can be located at the level of impairment, before the onset of functional limitations or disability. Frailty has been shown to have good predictive abilities for mortality (Chang & Lin, 2015) and hospitalization (Chang et al., 2018). At the same time, it is a reversible state, not the negative health outcome itself, thus indicating opportunities for intervention in health trajectories (Etman et al., 2012). It is especially salient in older people (Santos-Eggimann et al., 2009) and therefore well-suited to the population of interest.

The objective of this work is thus to explore frailty in older men and women living in Europe, comparing those with and without international migration experience. The aim is to answer four research questions:

1) Are there differences in frailty levels between immigrant and native older adults in Europe?

2) Do differences in frailty levels between immigrant and native older adults differ by gender?

3) What pathways shape differences in frailty between immigrant and native older adults, and how does this vary by gender?

4) What role does the destination country context play in immigrant-native differences in frailty?

1.2 Structure of the thesis

To investigate the questions posed above, I will first outline existing theories on health and migration as well as review empirical research on older migrants' health in Europe in section 2. Using this knowledge base, I formulate hypotheses based on the research questions. I evaluate the hypotheses by estimating regression models for the frailty outcome, using data from the Survey of Health, Ageing, and Retirement in Europe (SHARE). These data and my analysis steps are presented in section 3. I present results from this analysis in section 4 and discuss them critically in section 5. In section 6, I summarize and conclude the thesis. Supplementary tables and figures are presented in the appendix.

2. Literature review

2.1 Theoretical Approaches

Health inequality within societies is widespread. Whether it is gender, class, sexuality, race/ethnicity, or other dimensions; many social factors are shown to play a role in who experiences illness, disability and death at what rates. This is also true for health at older ages (Huisman et al., 2003; Kneale et al., 2021; Matthews et al., 2006; Rose, 2021; Schofield, 2018). One part that plays into this are the resources people do or do not have, e.g., in the form of income, education, skills, and social connections (Hornby-Turner et al., 2017). Another are their life experiences, both in older age itself and earlier, including as children (van der Linden et al., 2020). All of these factors are affected by migration. During migration, especially when it is through unofficial channels, people are often vulnerable and more likely to experience negative events like deprivation and violence (Llácer et al., 2007), and if they migrate towards a wealthier destination, their pre-migration life may have also posed more (health) challenges than for the average citizen of their destination. At the same time, skills and resources that were region-specific lose some of their protective power, and others, like local language skills, become relevant. As such, migration can be seen as an important factor affecting health at older ages.

Here, I will present previous theoretical work on the pathways in which migration is thought to affect health at older ages, including differences in these pathways by gender. I first cover experiences before, during, and after migration. Subsequently, I introduce selection effects that might affect which individuals decide to migrate in the first place, and which of those choose to return from their destination to their origin country before or during old age.

2.1.1 Before and during migration

Much of the theory on migration and health relates to migration from low-income countries in the global south to high-income countries in the global north. In this context, the prevalence of adverse conditions is likely higher and the quality of health care lower in the origin country than in the destination (Llácer et al., 2007). This puts migrants at a disadvantage compared to people born in the destination country, with increased disadvantage the later in the life course migration takes place. Llácer et al. (2007) also argue that in the average origin country, women's social position in relation to men is much lower than in destination countries, disproportionately increasing health risks for migrant women.

Potential health-affecting experiences during migration are especially common for refugees and irregular migrants, up to and including situations of trafficking. Even when there is no outright violence, this type of migration is associated with prolonged stress, which is detrimental to health (Hossin, 2020). Again, women are often even more vulnerable to these negative experiences than men (Llácer et al., 2007).

2.1.2 Post-migration

Even after arrival in their destination country, migrants' circumstances are not the same as those of non-migrants. One often stressed aspect of this are differences in health behaviours. Norms around health behaviours (like smoking, alcohol consumption, eating and exercising habits) are thought to be more favourable in many origin countries, especially in the global South, than in the more industrialised destinations. If migrants take these norms with them, they can protect their health (Guillot et al., 2018). At the same time, as migrants adjust to their destination countries, their health behaviours might also converge with those of the non-migrated population of the destination country. These changes commonly include increased smoking, higher alcohol consumption, and a sedentary lifestyle, as well as a loss of community values and cultural heritage (Finch et al., 2004). Especially in the context of the United States, this has been considered an important reason for declining health trajectories of immigrants after migration (Abraído-Lanza et al., 2016; Alidu & Grunfeld, 2018; Cela & Di Barbiano Belgiojoso, 2021; Jasso et al., 2004; Loi & Hale, 2019).

Women are thought to be more likely to develop worse health behaviours than men after migrating to the US. This is attributed both to more positive patterns pre-migration and disadvantageous social networks encouraging "risky and nontraditional behaviours" post-migration (Gorman et al., 2010, p. 442). At the same time, even with this steeper decline in health behaviours, migrant women retain some advantages over their male counterparts. This is the case for US smoking patterns, where the gap between men's higher and women's lower rates of smoking is much larger in immigrant groups than in the non-Latino white population, and smoking behaviour is positively associated with orientation towards US culture (Gorman et al., 2014).

In addition to cultural norms, the post-migration experience is shaped by many structural factors: Being positioned as a migrant, and often in the same process as racially 'other', shifts structural positions and opens the floor for discrimination and disadvantages in societal power imbalances (Finch et al., 2004; Viruell-Fuentes, 2007). Therefore, Viruell-Fuentes (2007) criticises acculturation approaches' focus on individual-level cultural differences as lacking in structural explanations. The criticism remains anchored in the context of immigration into the US and demonstrates how Mexican immigrants' lives there are shaped by the local understanding of race and their racialized position therein. Especially in neighbourhoods and institutions with only a small Mexican population, experiences of "othering" are common: Being subjected to discriminatory remarks and treatments that mark a divide between Mexican immigrants and both Black and white Americans. This othering is understood as a function of structural racism, shaping the ascribed racial status of individuals and, through that, their access to resources and institutions. This, in turn, is thought to influence health outcomes. In addition to this indirect pathway, experiences of othering and racism may in themselves be thought of as psychosocial stressors that negatively impact health (Viruell-Fuentes, 2007).

In order to get an accurate picture of how social structure shapes immigrants' health post-migration, Viruell-Fuentes et al. (2012) employ an intersectionality framework. In it, immigration does not just intersect with race, but also with other demographic and social categories such as gender and class to create many-dimensional experiences of (dis)advantage. At the same time as race is constructed interdependently with immigrant status and place as described above, gender and class are also shaped by and shape (post)migration experiences and access to resources. In Viruell-Fuentes et al.'s view, all of these factors work in conjunction to influence health outcomes.

Still, gender's intersection with migrant status in the country of destination has not been studied extensively (Krobisch et al., 2021; Wandschneider et al., 2020). On the one hand, attitudes and norms around gender might include protective factors for migrant women's health, potentially making them less likely to work in physically strenuous occupations. If that is the case, female migrants might have a smaller health gap from their native counterparts compared to men. On the other hand, women's generally marginalized position could instead compound health disadvantages associated with migration. Even when their jobs are not as physically dangerous, their precariousness also exposes women to health risks. In addition, if women who migrated primarily for their husbands do not seek (or find) employment, they have fewer opportunities to build social networks than men and risk isolation. This is especially true for social support beyond family members, and networks only made up of family networks have also been shown to be detrimental to migrant women's health (Llácer et al., 2007). This compounding of disadvantage seems to be predominant in Europe, as reviewed by Krobisch et al. (2021), making migrant women an especially vulnerable group.

2.1.3 Selectivity

Before all the mechanisms described above lies the decision to migrate. Migration often involves high investments before potential returns can be seen, and not everybody who might be able to migrate can expect their returns to outweigh the investments. As such, migrants are a highly selective group, and this selectivity is thought to include factors related to health. Many choose to migrate for reasons related to attractive labour markets abroad. In that case, the choice to migrate is more attractive if a person is in good enough health to apply their skills and generate economic returns. Therefore, healthy people are more likely to decide to migrate. This becomes more relevant the higher the costs of migration are. For other people, labour market opportunities might not be central in their migration decision. Especially among older people, migration might actually be more attractive for those (expecting to be) in poor health if health care is better or less costly in the destination country than in the country of origin (Jasso et al., 2004). This would lead to negative selection on health but is considered less impactful than the positive selection due to labour market concerns. In addition to self-selection processes, many high-income countries also require immigrants to be in good health or at least to fulfil requirements of wealth and formal education that correlate with good health. These procedures could increase positive health selectivity further (Constant et al., 2018).

Selection effects could also persist into the post-migration period since most migratory movements include the option of return. If those in ill health choose to return more frequently than others (because they can no longer work as they planned to, because they want to make use of care networks in the country of origin, or even because they want to die in their place of birth), the remaining migrants are those in comparatively good health (Abraído-Lanza et al., 1999; Guillot et al., 2018).

These selection effects are likely to be shaped by gender dynamics. When women's migration follows their husbands' under a family reunification scheme, they are less likely to be positively selected on health (Garcia et al., 2019; Krobisch et al., 2021). Their husbands' selectivity on some health-influencing factors might, however, also hold for them to some degree (e.g., in terms of wealth or lifestyle). For the US context, Lu and Li (2020) also suggest that women rely more on social networks than men, especially in the risky circumstances of undocumented migration, reducing their health selectivity. Additionally, Garcia et al. (2019) suggest an opposite "salmon bias" effect: Older women might want to join their migrant children in their host countries as their health deteriorates, resulting in negative health selectivity. For return migration selectivity, the opposite would apply for those women who left previous children in the origin country. However, it appears more common that women have children after migration, who then keep them in the destination country even if they have an interest in returning (see, e.g., Liversage & Mizrahi Mirdal, 2017).

In summary, while the experiences and resources relevant for old-age health appear to be disadvantageous for migrants, selectivity might counteract this to some degree. Women, however, are theorised to be less positively selected and even more negatively affected by migration-related experiences, which would make them more vulnerable to frailty.

2.2 Empirical literature

To go beyond the theoretical mechanisms outlined above, I now introduce the empirical literature relevant to my research questions. While migrants' health and its relation to that of natives has been the object of study for decades (see e.g. Jasso et al., 2004; Razum et al., 1998), only more recently has this interest expanded to the growing population of older migrants in Europe (e.g., Bousmah et al., 2019; Cela & Di Barbiano Belgiojoso, 2021; Ciobanu et al., 2017; Guillot et al., 2018; Jacobsen et al., 2023; Kristiansen et al., 2016; Lanari & Bussini, 2012; Reus-Pons et al., 2017; Reus-Pons et al., 2018; Vonneilich et al., 2021; Walkden et al., 2018). For a wide array of health indicators, these studies indicate that while there are sometimes health advantages at the time of migration (Bousmah et al., 2019; Gkiouleka & Huijts, 2020; Guillot et al., 2018; Malmusi et al., 2010; Trappolini & Giudici, 2021; Wallace et al., 2019), in the studies where age trajectories are allowed to vary between non-/migrants, health advantages do not usually persist into older age (Guillot et al., 2018; Trappolini & Giudici, 2021; Wallace et al., 2019).

In order to review the literature that is most applicable to my research questions, I focus on European destination contexts only. First, I review previous findings on differences between older migrants and natives for different health indicators, including frailty. Secondly, I review literature on the specific impact of gender on the relationship between migration status and health. Lastly, I summarize those findings from the previously reviewed papers that pertain to the relevance of destination country contexts.

2.2.1 Comparing older immigrants and natives

Frailty

As discussed in section 1.1, the outcome of interest in this work is frailty, which has previously been explored in the population of interest by Brothers et al. (2014), Franse et al. (2018) and Walkden et al. (2018).

Walkden et al. (2018) use data from SHARE waves 1 to 5 (years 2004-2013), which sample the 50+ population from 19 European countries. Using a frailty index, they show that migrants tend to be frailer than native Europeans. Even after controlling for age, gender, height (as a proxy for adverse early life experiences), education (as a proxy for pre-migration SES), home ownership, household income and host country citizenship, frailty likelihood was 16% higher for migrants than natives. Disparities between migrants and natives are stronger the lower the performance of the host country on the health care part of the Migrant Integration Policy Index MIPEX, indicating the relevance of host country policy. When differentiating by region, they find a small negative effect of migration status on frailty for Eastern Europe along with the bigger impacts in Western and Northern Europe, but not for Southern Europe (Walkden et al., 2018).

This relevance of the host country is also supported by Brothers et al.'s (2014). Since they base their analysis on only the two first waves of the SHARE, which included fewer countries, they pool a) Northern/Western Europe and b) Eastern/Southern Europe. They find that migrants are only statistically significantly frailer than natives in Northern/Western Europe (on a frailty index scale of 0 to 1, migrants' means are 0.01 to 0.03 points higher than non-migrants' mean of 0.15). The differences are not statistically significant in Eastern/Southern Europe (Brothers et al., 2014).

Franse et al. also use a frailty index, but different data, gathered in the Netherlands: The Older Persons and Informal Caregivers Survey Minimum DataSet, which includes persons aged 55 and up of Dutch, Indonesian, Surinamese, Moroccan or Turkish ethnic backgrounds. This includes second-generation immigrants, but those make up less than 10% of all immigrant respondents. The main finding of the study is that persons with Surinamese, Moroccan, and Turkish backgrounds tended to be frailer than those with a Dutch background, after adjustment for age, sex, living alone and education: Their odds of crossing the frailty threshold were between 2 and 6 times larger than for those with a Dutch background. For older people with an Indonesian background, the direction of effects was also positive, but they were small and non-significant (Franse et al., 2018).

In summary, existing research on frailty in older migrants in Europe suggests that they tend to be frailer than their native counterparts, but that origin and host country play a role in whether and how that effect manifests.

Frailty is a construct meant to integrate several dimensions of health simultaneously (see e.g. the six components for Franse et al.'s (2018) operationalization of the frailty index, p.7). Other research focuses instead on separate dimensions. Since they also play a role in frailty, reviewing studies using self-rated health (SRH), functional limitations, and chronic conditions as outcome variables enables me to get a fuller picture.

Self-rated health

SRH appears to be the most studied outcome variable in the population of interest. To my knowledge, three studies recently investigated older migrants' SRH in Europe, all using SHARE data: Bousmah et al. (2019), Lanari and Bussini (2012), and Solé-Auró and Crimmins (2008). All three studies dichotomized the corresponding interview item for analysis (Bousmah et al., 2019; Lanari & Bussini, 2012; Solé-Auró & Crimmins, 2008). The latter two studies define immigrants as foreign-born (Lanari & Bussini, 2012; Solé-Auró & Crimmins, 2008) and work with just the first wave of SHARE data, collected in 2004-05 in 11 European countries. Solé-Auró and Crimmins analyse all countries separately and observe stronger SRH disadvantages for immigrants compared to natives in Northern than Southern European countries, which they attribute to the higher health levels of natives in the Northern countries. Controlling for age and sex, they find that most statistically significant differences in SRH are to the disadvantage of immigrants (Solé-Auró & Crimmins, 2008).

Lanari and Bussini (2012) reduce the scope to just Northern European countries with long-standing immigration histories. They come to generally similar conclusions of SRH more likely being poor for immigrants than for natives. However, compared to Solé-Auró and Crimmins, they include a broader set of migration-specific (nativity, length of stay and citizenship status) and socio-economic (household type, employment status, economic resources) variables as well as age, sex, and education. This reveals that SRH disparities differ by origin and host country, are more pronounced for non-citizens (after living in the host country for more than ten years) than for naturalized immigrants and are explained by socioeconomic factors only for some of the immigrant groups (Lanari & Bussini, 2012).

Bousmah et al. (2019) make use of SHARE waves 1 to 5 (collected between 2004 and 2013), which expands the number of participating European countries to 19. Instead of looking at migration history, they compare citizens to non-citizens[1]. Additionally, they work with the panel structure of the data to investigate the effects of duration of stay in the host country on health longitudinally. Controlling for a wide range of socio-economic, behavioural and health-care specific variables, Bousmah et al. find a healthy immigrant (or non-citizen) effect for SRH that is reversed with increasing share of lifetime spent in the host country: Non-citizens' probability to be in poor SRH converges with that of citizens at around 35% of life spent in the host country, corresponding to an average of about 20 years, and then continues to grow beyond that of citizens. They also study health trajectories over share of lifetime spent in the host country by wealth of origin country (grouped into low, medium and high values on the Human Development Index (HDI)). For SRH, the trajectories are close to parallel, with the lowest likelihood of poor SRH estimated for immigrants from low-HDI backgrounds and the highest for those with medium-HDI backgrounds. However, confidence intervals overlap substantially (Bousmah et al., 2019).

Furthermore, Constant and Milewski (2021) find evidence for positive self-selection into migration based on SRH for older intra-European migrants, based on SHARE waves 1 to 6 (2004 to 2015). Stawarz et al. (2021) echo this for German emigrants, adding that the health advantage is bigger for those who migrate later in life.

In summary, previous studies show the SRH of migrants to be similar to or worse than that of non-migrants at older ages, even when there is a positive effect earlier in life. There are, however, positive selection effects when comparing migrants to their non-migrated peers in their (European) origin countries.

Functional limitations and chronic conditions

Bousmah et al. (2019) and Solé-Auró and Crimmins (2008) also investigated functional limitations and chronic conditions in the same studies reviewed above. Bousmah et al. look only at limitations in Activities of Daily Living (ADL). They find a similar trajectory as for SRH – originally better outcomes for non-citizens that are lost over time spent in the destination

[1] This leads to a much smaller subsample of immigrants than in the other studies: 3% of the total sample, compared to 8% Solé-Auró and Crimmins (2008) and 11% Lanari and Bussini (2012) when defined as those born outside the country of residence.

country – but with the non-citizen advantage being lost much later. Non-citizens' likelihood of having ADL limitations only converges with citizens' after spending 80% of their life in the host country, equivalent to about 50 years. However, the trajectory is reversed for those migrants from low-HDI countries.

The trajectory for chronic conditions forms a middle point between SRH and ADL difficulties. Among non-citizens, those with low-HDI backgrounds were estimated to have the lowest number of chronic conditions originally, and those with high-HDI backgrounds the highest, all of which converge with citizens' levels after about 50% of non-citizens' lives (or about 30 years) spent in the host country (Bousmah et al., 2019).

Solé-Auró and Crimmins (2008) look at Nagi's (1976) functioning indicator and difficulty in Instrumental Activities of Daily Living (IADL) in addition to ADL difficulties. Controlling for age, sex, and education, all statistically significant effects point towards worse functioning for immigrants compared to natives, but effect sizes are very small for most countries. For chronic diseases, the picture is similar: After controlling age, sex, and education, immigrants are statistically significantly more likely to have two or more chronic diseases than natives in Germany, Spain, Sweden, and Switzerland. However, they are statistically significantly less likely to in Austria, which the authors speculate might have to do with less frequent doctor's visits, which is a prerequisite for diagnosis of chronic issues (Solé-Auró & Crimmins, 2008).

Other studies on SRH, functional limitations and chronic diseases of older European migrants lack a comparison to native populations (e.g., Cela & Di Barbiano Belgiojoso, 2021; Krobisch et al., 2021; Vaillant & Wolff, 2010) or exclude those at increasingly older ages to minimize selection effects instead of trying to understand them (Gkiouleka & Huijts, 2020; Malmusi et al., 2010) and are thus not as applicable to my research questions.

In conclusion, the existing research suggests that older migrants are more likely to be in bad health than their native counterparts, especially after spending significant portions of their life in the host country. This applies to different measures of health, including frailty and related concepts. The size and significance of effects vary by origin and host countries, and only some of it can be explained by socio-economic factors.

2.2.2 Gender dynamics

Unfortunately, despite several calls to integrate a gender perspective into migration and health research (Curran et al., 2006; Fresnoza-Flot, 2022; Llácer et al., 2007; Wandschneider et al., 2020), there is little empirical literature on health of ageing migrants in Europe that explicitly elaborates gender dynamics (Krobisch et al., 2021). When they reported gender effects at all, the papers above only reveal that regardless of migration status, frailty is more common in women than in men (Brothers et al., 2014; Walkden et al., 2018). Below, I present current research on gender effects on migrants' health that looks at broader age ranges and allows inferences to older age, especially through presenting health trajectories.

Gkiouleka and Huijts (2020) studied health differentials between first- and second-generation migrants and natives across 27 European countries, focusing on the interactions of occupational status and gender with migrant generation. They pooled data from six waves of the European Social Survey from 2004 to 2014, looking at the population between 25 and 75 years of age. As outcomes, they chose two measures of subjective health: General SRH and the perception of being hampered in everyday life by a (physical or mental) medical condition. They find that in most constellations, women are more likely to report poor health and hampering conditions, though the differences between men and women are not statistically significant for all migration-occupation categories. Their multi-level models also show that there is a substantial part of variance in health that is due to country context, about a quarter to a third of the variance in each model (Gkiouleka & Huijts, 2020).

Four additional studies look at the intersection of gender, migration, and health in specific country contexts: Krobisch et al. (2021) for Germany, Malmusi et al. (2010) for Spain, Trappolini and Giudici (2021) for Italy, and Wallace et al. (2019) for France.

Krobisch et al. (2021) investigate SRH differences between men and women from Turkey aged 55 and above in Berlin, Germany, in a mixed-methods design with self-collected data. In their quantitative mediation analysis, they include age, marital status, education, ethnic identity, smoking behaviour, physical activity, functional limitations (combining ADL and IADL), emotional loneliness and frequency of medical check-ups. They find that women tend to report worse SRH, which is mediated by higher functional limitations and loneliness. Marital status also affected women's health insofar that they were more likely to be in the "widowed or divorced" category than men and this was counterintuitively linked to better SRH. Qualitatively, interviews suggested that migrant women's worse health could be linked to the pressures of multiple social roles (as employees and housemakers) and unsupportive partnerships (Krobisch et al., 2021). Unfortunately, the study does not include a non-migrated German comparison group. Therefore, it cannot be estimated whether the gap between men and women is especially big among this migrant group.

Malmusi et al. (2010) studied SRH inequalities between international migrants, internal migrants, and natives aged 25 to 64 in Catalonia, Spain, looking specifically at the effect of gender, class, and HDI of the origin country. They use data from two official government surveys collected in 2006. First, only controlling for age, they find that women report poor health more often than men across the board, and that while immigrant men from low-HDI origin countries had a health advantage over native men, their female counterparts had worse health than native women. Controlling for economic variables changes the picture: When social class and material assets are included in the model in addition to age, female immigrants from low-HDI countries who had been in Catalonia for at least six years have statistically significantly lower odds of having poor health than their native counterparts. The originally negative health effect for more recently migrated women from low-HDI countries also changes direction and becomes non-significant when controlling for economic factors. The odds of poor health for migrant women's from higher-HDI backgrounds are not statistically significantly different from native women's. For men from low-HDI countries (both recent and less recent migrants), odds of poor health are significantly lower than natives' after controlling for economic variables (Malmusi et al., 2010). Thus, it appears that while there is to some extent a healthy migrant effect for men from low-HDI origin countries, the same is less clear for women, suggesting a female health disadvantage.

Trappolini and Giudici (2021) looked at gender specificities in the migrant-native health convergence over duration of stay in Italy, using the Italian Health Survey 2013. Their sample consists of native Italians and immigrants aged 20 to 64, for whom they assess functional limitations and chronic conditions in addition to SRH. They also find health advantages for recent immigrants as well as worse health and a slightly bigger negative effect of length of stay for women than for men (Trappolini & Giudici, 2021). However, since the data are cross-sectional, length of stay cannot be disentangled from the timing of migration, which might confound the findings.

Wallace et al. (2019) studied gender differences in migrant mortality patterns over the duration of stay in France. They use 2004-2014 data from France's largest socio-demographic panel, including immigrants and natives aged 20 and above. In their survival models, they show that the pattern of an initial mortality advantage at the time of migration and later convergence to native mortality levels holds for both men and women (Wallace et al., 2019).

Summarizing these findings, the interaction of migration and gender is not quite clear, but most often appears slightly disadvantageous for migrated women. Again, length of stay and ageing are associated with worse health outcomes, whereas migration itself is sometimes positively associated with health, especially for men.

2.2.3 Country-level impacts

In the studies reviewed above that looked at different host country contexts in a comparative perspective, marked differences in migrant-native health gaps were found across countries or country groups (Brothers et al., 2014; Gkiouleka & Huijts, 2020; Lanari & Bussini, 2012; Solé-Auró & Crimmins, 2008; Walkden et al., 2018). Explanations include the health levels of the general population and the composition of migrant flows.

Brothers et al. (2014) and Walkden et al. (2018) find bigger migrant disadvantages in Northern and Western European countries, smaller ones in Eastern Europe and none in Southern Europe, which correlates with regions' different migration histories (also suggested for younger migrant cohorts by Moullan & Jusot, 2014). For policy measures, Walkden et al. (2018) indicate that more inclusive health care systems reduce migrants' potential health disadvantage. This is corroborated for younger refugees in Germany and Austria by Georges et al. (2021); Giannoni et al. (2016) make this case for a wider array of migrant integration measures. However, in Walkden et al.'s study, health care inclusion has the strongest effect compared to other areas of migrant integration.

2.3 Conceptual model

The factors shaping immigrants' health and their connections, as outlined above, are visualized in the conceptual model of this study in Fig. 1. Experiences before and during migration, including selectivity effects, are the starting point. These go on to influence acculturation processes and experiences of disadvantage in the destination country, which in turn impact the likelihood of poor health and specifically frailty. All three stages are themselves influenced by gender dynamics: Decision making for migration has been shown to be different for men and women, and acculturation processes might work in gender-informed ways, e.g. in terms of who has access to which parts of the host country society (e.g. Llácer et al., 2007). Similarly, structural disadvantages include sexism as well as class, race, and citizenship issues, all of which intersect and result in gender-specific experiences (Viruell-Fuentes et al., 2012). The way subsequent stress and health behaviours lead to frailty might also differ for men and women, since frailty rates in general do (Santos-Eggimann et al., 2009; Walkden et al., 2018). Finally, these processes take place in country contexts with different migration histories and policies. These provide the backdrop against which selection, acculturation, disadvantage and gender intersections play out to influence frailty risk (see, e.g., Moullan & Jusot, 2014).

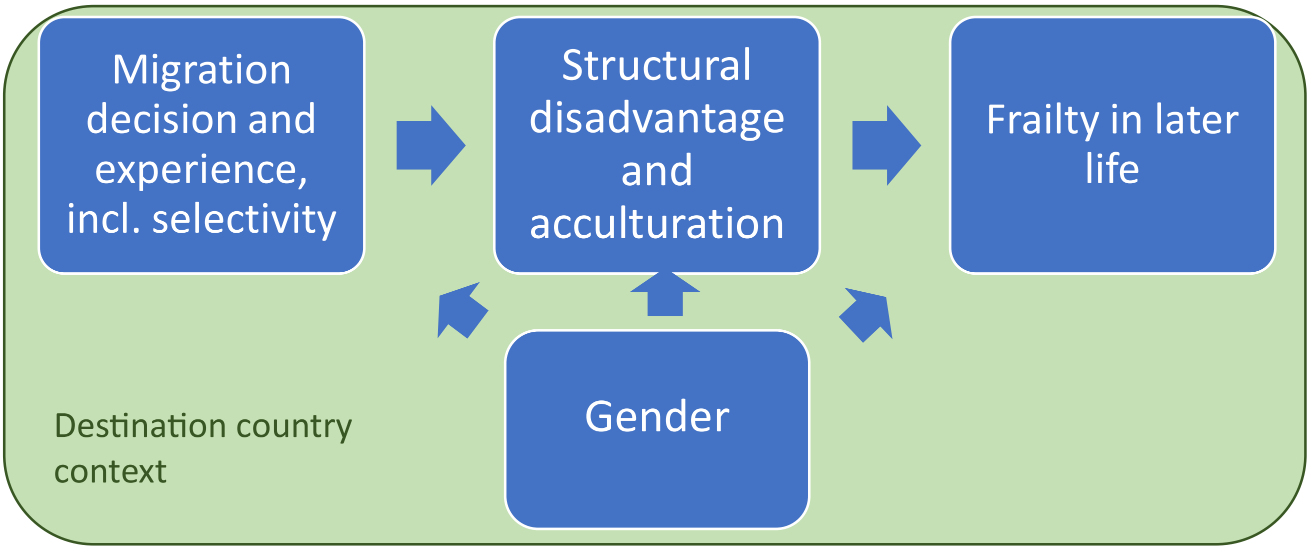

Figure 1: Conceptual model

2.4 Hypotheses

Based on theories and empirical literature reviewed above, I now formulate hypotheses to answer my four research questions:

1) Are there differences in frailty levels between immigrant and native older adults in Europe?

I hypothesize that migration to and within Europe results in structural disadvantages for immigrants, such as discrimination, lack of access to health care, and disadvantageous working and living conditions (Viruell-Fuentes et al., 2012). It appears that at older ages, there is no lasting healthy migrant effect in a European context (Walkden et al., 2018). Thus, Hypothesis 1 is: Immigrated older adults will be more likely to be pre-frail and frail than native older adults.

2) Do differences in frailty levels between immigrant and native older adults differ by gender?

The structural disadvantage model posits that gender interacts with migrant status to shape the structural conditions of people's lives. It has been shown that women's health in later life is more vulnerable compared to men, and that this is likely to be even more pronounced among migrated populations (Gkiouleka & Huijts, 2020; Llácer et al., 2007). Thus, Hypothesis 2 is: The difference in frailty between men and women will be bigger for migrants than for natives.

3) What pathways shape differences in frailty between immigrant and native older adults, and how does this vary by gender?

Based on acculturation and structural disadvantage theories (Jasso et al., 2004; Viruell-Fuentes et al., 2012) and their operationalisation in previous studies, Hypothesis 3a is: Differences in health behaviours, socio-economic status, and social networks mediate immigrant-native differences in frailty. To take into account gender dynamics as elaborated e.g. by Llácer et al. (2007), Hypothesis 3b is: While selectivity effects and economic pathways affecting immigrant-native differences in frailty will be more important among men, social pathways and acculturation measures will be more important among women.

4) What role does the destination country context play in immigrant-native differences in frailty?

Northern and Western European countries have seen cohorts of older migrants for a longer time than other countries in the sample. Countries' immigration policies also vary across Europe and inform the accessibility of health care to migrants. Given that both of these have been shown to play a role in immigrant-native health differences (Walkden et al., 2018), Hypothesis 4 is: Immigrant-native differences in frailty will be moderated by countries' region within Europe and the inclusivity of policies for migrants' access to health care.

3. Data and Methods

Above, I derived hypotheses from previous research and the theoretical literature. I will now elaborate on how I evaluate them and answer my research questions empirically. To this end, I conduct descriptive analysis and multinomial logistic regressions using data from the SHARE project. This thesis uses data from SHARE wave 6 (DOI: 10.6103/SHARE.w6.800); see Börsch-Supan et al. (2013) for methodological details. The SHARE data collection has been funded by the European Commission, DG RTD through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982, DASISH: GA N°283646) and Horizon 2020 (SHARE-DEV3: GA N°676536, SHARE-COHESION: GA N°870628, SERISS: GA N°654221, SSHOC: GA N°823782, SHARE-COVID19: GA N°101015924) and by DG Employment, Social Affairs & Inclusion through VS 2015/0195, VS 2016/0135, VS 2018/0285, VS 2019/0332, and VS 2020/0313. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2,

P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C, RAG052527A) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

As could be seen above in the literature review, SHARE is commonly used for analyses of older European's health and frailty in particular. Most of the studies reviewed above that used other data sources focused on a single country or just a small number of countries, using national data. Of course, there are also other European data collections besides SHARE; however, they tend not to focus on older people, who consequently make up only a small proportion of the samples (e.g., the European Social Survey used by Gkiouleka & Huijts, 2020). The main drawback of the SHARE for my research question is that it is not specifically focused on migration studies. Thus, there are no special efforts to reach undocumented populations or those with limited knowledge of the countries' official languages. Furthermore, the migration-specific information asked of respondents is limited and does not correspond to all the demands raised for studies of migration and health e.g. by Hossin (2020). It is, however, still sufficient for my interests. Overall, the SHARE's focus on the older population, its large international probability sample and the accordance with data standards make it a good fit for my analyses. In the following, I further elaborate information on the data, its advantages and shortcomings, followed by operationalisation of key concepts and the analysis plan.

3.1 Data

The Survey of Health, Ageing and Retirement in Europe (SHARE) is a European research infrastructure conducting a panel study since 2004 (Börsch-Supan et al., 2013; SHARE-ERIC, 2023c). It has collected information from the population aged 50+ and their cohabiting partners in 27 European countries and Israel. So far, a total of 140,000 respondents have participated in 530,000 interviews (SHARE-ERIC, 2023c) across eight regular waves and in two pandemic-adjusted "Corona Surveys" (Munich Research Institute for the Economics of Aging and SHARE Analyses [MEA], 2022, p. 8). The specific participating countries change from wave to wave. Thematically, SHARE is concerned with "the effects of health, social, economic and environmental policies" for Europeans (SHARE-ERIC, 2023c) and connected to several similar European and global research initiatives. It collects data on a broad array of topics in older Europeans' circumstances: Health and health care usage, finances and housing, activities and social networks (MEA, 2022, p. 15). In additional life history interviews (waves 3 and 7), further information on respondents' life course was collected (MEA, 2022, pp. 12–13). Except for the Corona Surveys, SHARE interviews are computer-assisted personal interviews, which allows the inclusion of physical tests (MEA, 2022, p. 12).

At the moment, the most recent regular SHARE data are from wave 8, collected in 2019 and 2020. Due to the onset of the COVID-19 pandemic, interviews had to be suspended in all countries over the course of March 2020. As a result, only about 70% of expected longitudinal and 50% of refreshment sample responses were collected, with large differences in coverage across countries (Bergmann & Börsch-Supan, 2021; Scherpenzeel, Annette et al., 2020). This means that any inference to the complete population would have to be undertaken with great caution and, of course, that there are fewer responses overall. SHARE reacted to the pandemic outbreak by establishing two "Corona Surveys" in 2020 and 2021 that were conducted via telephone. For them, the regular questionnaire was changed, shortened and supplemented by pandemic-specific items (MEA, 2022, pp. 6–7). This unfortunately makes them unsuitable for investigating my research questions. Wave 7 was conducted in 2019, but due to the inclusion of life history interviews, about 80% of respondents were given only an abridged version of the usual questionnaire (Bergmann, Börsch-Supan, & Scherpenzeel, Annette, 2019). This makes the wave unsuitable for frailty analyses as well. As a result, I will work with the next most recent data collected in 2015 for wave 6 (Börsch-Supan, 2022e). Here, the 18 participating countries were Austria, Belgium, Croatia, the Czech Republic, Denmark, Estonia, France, Germany, Greece, Israel, Italy, Luxembourg, Poland, Portugal, Slovenia, Spain, Sweden, and Switzerland (Malter & Sand, 2017, p. 100). Data from respondents' first interviews in previous waves 1, 2, 4, and 5 are added as needed (Börsch-Supan, 2022a, 2022b, 2022c, 2022d).

3.1.1 Sampling

Each participating country independently draws its sample of respondents. The target population consists of persons 1) at least 50 years of age and 2) their spouse/partner living in the same household, who a) have regular domicile in the country the survey is being conducted in, b) are not incarcerated or hospitalized over the fieldwork period, and c) speak (one of) the country's official language(s). Those living in nursing homes and care institutions besides hospitals are generally included, but it varies how well they are taken into account depending on each country's sampling frame (Bergmann et al., 2017, p. 77). For the study of an elderly population, this is an important limitation. The language criterion is especially limiting in the study of migrant populations, in particular for those who migrated recently or live without much contact to the majority population. Thus, people with limited local language skills form a vulnerable group in terms of health (Schouten et al., 2020) which is not well represented in SHARE samples[2].

The ideal for SHARE countries is to draw probability samples from a sampling frame spanning the whole population and containing reliable information on age, like population registers. Due to limited availability and access to those, however, some country teams use other sampling frames: Census and election registers, insurance registers, and even random walk procedures with screening for age eligibility (Bergmann et al., 2017, pp. 78–79).

Sampling strategies within those sampling frames are also allowed to differ between countries. Sampling probabilities and corresponding design weights resulting from each country's procedures are computed by the central SHARE coordination to facilitate harmonization over countries and waves. They also take into account the issues of 1) inclusion probabilities for couples where both members are age-eligible and 2) cross-selection probabilities occurring when the longitudinal sample overlaps with the sampling frame for refreshment samples (Bergmann et al., 2017, pp. 81–82).

3.1.2 Non-response and attrition, weights and imputations

Of course, not all respondents selected in the sampling procedures actually respond, and not all of those who participated in an earlier wave (the longitudinal sample) respond when contacted again. In the longitudinal sample in wave 6, most countries achieved a household retention rate (at least one completed interview per household) between 68% (Austria) and 82% (Portugal). The only exception was France with a very low household retention rate of 53%, attributed to a switch in survey agency (Malter & Sand, 2017, p. 104).

Refreshment samples for wave 6 were recruited in Belgium, Croatia, Denmark, Estonia, France, Greece, Italy, Luxembourg, Poland, and Slovenia. Cooperation of contacted households was much lower for the refreshment sample than for the longitudinal sample, resulting in household response rates between 30% (Luxembourg) and 64% (Greece) (Malter & Sand, 2017, pp. 113–114). For more information on response and retention rates in waves 1 to 7, see Bergmann, Kneip, et al. (2019).

Non-response and attrition do not just reduce the sample size, consequently increasing standard errors; there are also likely systematic differences between respondents and non-respondents, which can bias both estimates and standard errors (Bergmann et al., 2017, p. 84).

[2] In some countries, the survey is also administered in a significantly spoken non-official language (Russian in Israel and Estonia; MEA, 2022), potentially providing access to significant migrant populations. For wave 7 and 8, this is documented as also offering the survey in non-official languages spoken by more than 10% of the population (Bergmann, Börsch-Supan, and Scherpenzeel, Annette (2019)).

SHARE deals with this issue by calculating calibrated weights that take into account gender, region, and age of respondents and non-respondents in addition to their design weights. (Bergmann et al., 2017, p. 88). The resulting weights can then be used to adjust analyses so that inferences to the whole population are more valid. For this study, cross-sectional calibrated weights at the individual level are most appropriate.

In addition to non-response of whole respondents (unit non-response), there is also item non-response: Respondents refusing an answer to specific questions. For SHARE wave 5, item non-response over 5% occurred only for monetary variables regarding, for example, income, assets, and expenditures. Response rates for these items vary considerably between countries. The documentation for wave 6 does not include updated information; however, some non-monetary items such as current smoking behaviour and social network data are missing for over 5% here. To offset some of the initial item non-response, refusal of a point estimate for monetary items results in being asked a series of "unfolding bracket" questions. In this procedure, respondents are sequentially given up to three country-specific pre-determined threshold values and are asked to say whether their amount was larger, smaller, or about the same as those. This procedure results in an approximate point estimate (amount roughly equal to one of the thresholds) or an interval estimate (between two thresholds or above/below the outer ones). Of course, respondents can also choose "Refuse" or "Don't know" as answers to the unfolding brackets (Luca et al., 2015, pp. 86–87). In wave 5, unfolding bracket answers provided information on over half of the initially missing data points; again, there is substantial variance between countries (Luca et al., 2015, pp. 89–90).

Aside from unfolding brackets, missing data points were also imputed from the information respondents gave on other aspects of their situation. To this end, hot-deck imputations and fully conditional specification imputation procedures were used, depending on how often the specific variable to be imputed was missing in the total sample. In any imputation, there is uncertainty. To account for this and make it assessable, SHARE provides five different imputations per missing value. For more information on imputation procedures, see MEA (2022, pp. 48–52).

Summing up the data description, it has become clear that SHARE goes to considerable lengths to provide good-quality data. All country samples are intended to be nationally representative and are based on the best available sampling frames, and there are important steps in coordination and harmonization across countries. Of course, like any survey study, SHARE has issues with imperfect sampling frames, potential nonresponse bias and specific methodological challenges. Especially in refreshment samples, response rates are low in many countries, and monetary interview items also show substantial non-response. However, using calibrated weights and estimates derived from unfolding brackets and imputations can offset these issues to some extent, even as interpretation of the data needs to remain conscious of limits to inference to the population.

3.1.3 Ethics

SHARE data are collected in accordance with international standards of data protection and research ethics. No identifying information is included in any of the data (SHARE-ERIC, 2023a). I accessed the data through SHARE's user license process. All my analyses are in line with SHARE's terms of use; in particular, research is of a scientific nature only, and I make no efforts towards re-identification of participants.

3.2 Analysis plan

3.2.1 Operationalisation

Dependent variable

The dependent variable is frailty. I use the frailty phenotype operationalisation established for SHARE by Santos-Eggimann et al. (2009) and used e.g. by Sirven et al. (2020). It is both less demanding of data and better theoretically founded than the alternative frailty index (Sirven et al., 2020). It includes criterions for the five dimensions exhaustion, shrinking, weakness, slowness, and low activity.

1) Exhaustion was fulfilled if respondents answered affirmatively to the question "In the last month, have you had too little energy to do things you wanted to do? (yes/no)." 2) Shrinking was fulfilled if respondents reported a "diminution in desire for food" when asked about their appetite. If their response to that question was not codable, they were asked whether they had been "eating more or less than usual" and coded as shrinking if they reported eating less. 3) Weakness was fulfilled if the highest of four grip strength measures was below a threshold defined for the applicable combination of gender and body mass index (BMI) as set by Fried et al. (2001). 4) Slowness would ideally have been assessed by measuring walking speed; however, SHARE only includes that test for respondents aged 75 and over. Walking speed measures were shown to correspond well to mobility questions by Santos-Eggimann et al. (2009), so they were used instead. Slowness was thus fulfilled if respondents answered affirmatively to either of the two questions "Because of a health problem, do you have difficulty [expected to last more than 3 months] walking 100 meters" or "… climbing one flight of stairs without resting". 5) Finally, low activity was fulfilled if respondents answered "one to three times a month" or "hardly ever or never" when asked "How often do

you engage in activities that require a low or moderate level of energy such as gardening, cleaning the car, or going for a walk?" (MEA, 2015; Santos-Eggimann et al., 2009). If none of the criterions are fulfilled, the respondent is coded as robust; if one or two of them are fulfilled, as pre-frail; and if three to five of them are fulfilled, as frail.

Main independent variables

The main independent variables are migration status and gender. Migration status is derived from the variable asking whether respondents were born in the country of interview (which is by definition their country of regular residence, Bergmann et al., 2017, p. 77). If the answer is no, they are recorded as immigrants; if their answer is yes, they are recorded as native. I do not distinguish immigrants' children from natives. Respondents who are selected for a country's sample and move abroad before being interviewed (before their first interview or after already participating in an earlier wave) are classed as ineligible and not followed up on (Malter & Sand, 2017, pp. 94–96).

Gender is measured in SHARE by interviewer observation, with instructions to "[n]ote sex of respondent from observation (ask if unsure)" as a binary measure (male/female) (MEA, 2017).

Mediators

I analyse three groups of potential mediators of the relationship between migration and frailty: Health behaviours, as related to acculturation theory; economic indicators; and social capital indicators. Health behaviours under study are smoking behaviour, with a binary variable indicating whether the respondent is a current smoker, and alcohol consumption, with three categories of three or more drinks per week, less than 3 drinks per week, and non-drinkers, both in line with Sirven et al. (2020).

Economic indicators are formal education (coded as none or primary, secondary, and post-secondary according to ISCED 1997 as done by Walkden et al., 2018), occupational status (retired, employed or self-employed, unemployed, homemaker or permanently sick, Bousmah et al., 2019), household wealth and household income (Sirven et al., 2020). The latter two are equivalized by dividing by the square root of household size (as done by e.g. the OECD Income Distribution Database; OECD, 2020). Household income is transformed into its natural logarithm, wealth is transformed into its cubic root. Both transformations are intended to account for the highly right-skewed distributions and the diminished utility of increasing monetary values at higher levels (Greene & Baron, 2001).

SHARE offers two specifications of household income: One aggregating different streams of income, and one asking respondents directly for their estimate of their total income (SHARE-ERIC, 2023b). I use the latter because it appears plausible that respondents have an easier time estimating the sum of all their income accurately rather than each stream by itself. Additionally, this variable contains fewer missing values, making it less reliant on imputations or prone to bias when excluding missing cases.

Social capital is assessed via marital status (married or in a registered partnership, widowed, divorced, never married as by Bousmah et al., 2019), participation in social clubs in the past year (Sirven et al., 2020), and not mentioning anyone when asked about people in one's social network. These variables are intended to assess the level of social integration and support, which has often been shown to be a difficulty for migrants (Makwarimba et al., 2010) and an important factor for health (Brandt & Hagge, 2020).

Moderators

Migration- and country-specific variables potentially moderate the impact migration status has on frailty for each gender.

To take into account differential migration experiences, I assess time since migration in years (using the year of migration), wealth of the country of origin, and citizenship in the destination country. Untangling age effects, age at migration, and duration of stay is a statistical challenge. Since the three measures are interdependent, they cannot all be included in full in the same model (for elaboration of this point see also Wallace et al., 2019). For my theoretical background and research interest, duration of stay appears more central than age at migration, as it is directly related to acculturation and disadvantage processes in the destination country. It therefore takes precedence over including age at migration.

Origin country wealth is operationalized into a categorical variable, sorting origin countries by their value on the HDI. It would be most informative to use information on origin countries specific to the year of migration to assess respondents' situation before migration and the circumstances that informed their migration decision more accurately. However, HDI has only been computed since 1990, and the global average has been on the rise since then (United Nations [UN], 2023). So not only is no HDI information available for many of respondents' times of migration, but it is also not very well comparable across time. Instead, I opt for comparability with Bousmah et al. (2019) and use 2015 HDI values. Following their procedure, I classify origin countries as low-HDI if they had an HDI below 0.700, as medium-HDI between 0.700 and 0.836 and as high-HDI from 0.827 onwards. The first threshold is an often used value to separate low- and medium-HDI countries; the second is derived from the lowest value of an EU member state in that year, in this case Croatia (Bousmah et al., 2019; also in line with Malmusi et al., 2010)[3].

Lastly, citizenship in the destination country can be an important factor in the accessibility of health care and other institutions and services (Lanari & Bussini, 2012). It is self-reported.

To assess the impact of country attributes, I investigate both regional groupings and a more political division by health care inclusivity for migrants. Israel is not included (see section 3.2.2 below). For the first categorisation, countries are divided into the four regions Northern, Eastern, Southern, and Western Europe. Following Walkden et al. (2018), I use United Nations definitions of these regions (United Nations Statistics Division [UNSD], 2023). However, for the purposes of migration studies, Slovenia fits in better with the Eastern countries, even though the UN classifies it as Southern (Walkden et al., 2018). The same is the case for Croatia[4], whose migration history and present has been shaped in similar ways to the other post-Yugoslavian countries in the sample (Božić & Kuti, 2016). I thus reclassify them into the Eastern category. Thus, Denmark, Estonia, and Sweden were classified as Northern; Croatia, the Czech Republic, Poland, and Slovenia as Eastern; Greece, Italy, Portugal, and Spain as Southern; and Austria, Belgium, France, Germany, Luxembourg, and Switzerland as Western European.

Additionally, I want to assess the relevance of policy regarding health care accessibility for migrants. To make this comparable to a previous study by Walkden et al. (2018), I again follow their example and sort the countries into terciles according to their performance on the health care strand of the Migrant Integration Policy Index (MIPEX) IV (2014) (Huddleston et al., 2015). They also perform sensitivity checks, using the labour market strand and the complete eight-dimensional MIPEX instead of the health care strand and find that for both alternative measures, their explanatory power is not as strong as that of the health care strand (Walkden et al., 2018). This suggests that the health care strand is the most relevant when looking at frailty as an outcome. SHARE countries in the first tercile of MIPEX healthcare inclusiveness (most inclusive) were Switzerland, Italy, Austria, and Sweden; those in the middle tercile were Spain, Belgium, Denmark, France, the Czech Republic,

[3] Bousmah et al. (2019) refer to Hungary as being the EU member state with lowest HDI (0.836); this is not correct based on United Nations Development Programme (UNDP) (2016), which lists the 2015 values. In my sample, the correction affected 185 persons, which is unlikely to skew results substantially.

[4] Walkden et al. (2018) use SHARE wave 5, where Croatia did not participate.

Germany, Luxembourg, and Portugal[5]; and in the third tercile (least inclusive) were Estonia, Greece, Poland, Croatia, and Slovenia (Huddleston et al., 2015).

Control variables

Age is a central variable that can confound the relationship between migration status and frailty, since health usually deteriorates with age and migrant populations tend to be younger than the majority (e.g., Solé-Auró & Crimmins, 2008). As is usual in studies of frailty, to distinguish it from advanced disability, I also control for the number of limitations in (I)ADL and chronic diseases (see, e.g., Sirven et al., 2020). For parsimony, I combine these into one index that specifies in how many of the three dimensions respondents are having health issues.

Item nonresponse

For item non-response on some items, there are available imputations. Since the frailty dimension of weakness was assessed relative to BMI, some of the missing units regarding weakness were respondents who did participate in the grip strength measure but did not provide height and/or weight information. However, BMI is included in SHARE's imputation protocol, allowing tentative recovery of these cases. I also use imputations for missing income and wealth information. For all three instances, I compute the mean of SHARE's five offered solutions to the imputation procedures. For all other missing variables, I generate separate categories of missing values so that a) the sample size is not reduced unnecessarily and b) estimations are not unduly biased by missing information, as would be the case when only analysing complete cases (Pedersen et al., 2017). Using multiple imputations in parallel would be even more robust (Sterne et al., 2009), but is not feasible to combine with all steps of the analysis in Stata.

3.2.2 Analytical sample

Beyond SHARE's own sampling (frame) constraints, I will further restrict the analytical sample. Since the outcome measure is frailty status and frailty is much less common in middle age (Santos-Eggimann et al., 2009), I only use information from respondents aged 60 and up. Additionally, I include only European countries and exclude Israel, because its migration history and present is a very special case which should not be conflated with the other countries (see e.g. Constant et al., 2018; Brothers et al., 2014). As is standard in studies of frailty, I also exclude people living in institutions from the analytical sample (Chang &

[5] Spain, Belgium, and Denmark all achieve the same score on the cusp between the first and second tercile. In the interest of variation between terciles, they are all grouped into the middle tercile.

Lin, 2015; Etman et al., 2012; Franse et al., 2018; Santos-Eggimann et al., 2009). This also means less uncertainty and better comparability across countries regarding how well people in institutions were represented in sampling frame and sample. The remaining analytical sample consists of n=49.333 respondents, 21,641 of which had any of the relevant variables imputed (mainly wealth, making up for 20,370 of the imputed cases due to its value being composed of the answers to many individual questions). Descriptive characteristics of the analytical sample are shown in section 4.1.

3.2.3 Statistical analysis

Descriptive and bivariate statistics

First, I compute descriptive statistics (weighted, using calibrated cross-sectional individual weights as described in section 3.1.2). These include means and standard deviations for metric variables as well as percentages and numbers of observations for categorical variables. I compute descriptives 1) for the complete sample, as well as 2) separately for migrants and non-migrants and 3) separately for men and women. Additionally, I compute descriptive statistics by frailty type (robust, pre-frail, frail). These bivariate associations facilitate interpretation of later mediation analysis.

Regression models

I compute a number of multinomial logistic regression models for frailty phenotypes to answer each research question. Using multinomial models lets me compare the three established phenotypes instead of conflating pre-frailty with either robustness or frailty. Even though they are more parsimonious, I do not use ordinal logistic regression models since these usually assume that the predictors have the same impact on all steps between values of the dependent variable (DeMaris, 2004, pp. 304–305). Here, this would mean the same impact on 1) being robust vs. pre-frail as on 2) being pre-frail vs. frail. This assumption does not seem well-founded: Frailty is much less common than pre-frailty or robustness (Santos-Eggimann et al., 2009). There also appears to be a non-linear relationship between the different frailty states and adverse health outcomes like hospitalisation and mortality (Chang & Lin, 2015; Chang et al., 2018). This suggests the possibility of different underlying processes leading to pre-frailty and frailty, which in turn might be differentially affected by migration experience and gender. Therefore, I use multinomial models. The approach is cross-sectional.

Since each country participating in SHARE draws its own, independent sample, respondents can be thought of as clustered in countries. Ignoring the resulting multi-level structure might result in conflating differences between countries (e.g., in number of immigrants and prevalence of frailty) with effects within countries. This suggests the use of multilevel regressions. However, the number of country clusters in the sample is rather small to accurately estimate random effects models (Sommet & Morselli, 2017). Furthermore, in logistic regression, coefficients (and derived odds ratios) of nested same-sample models cannot be compared directly. To account for this, I will use the Karlson-Holm-Breen (KHB) method for comparisons between same-sample models (Karlson et al., 2012). Unfortunately, Stata does not support a combination of multinomial multilevel logistic regression with KHB. Consequently, my main models are estimated with regular multinomial logistic regression, including fixed effects for countries (McNeish & Kelley, 2019; Moehring, 2012). Their coefficients are then compared to both multilevel random effects models and the KHB estimation of mediation effects.

In regards to weights, I follow Winship and Radbill (1994): After determining that the weights do not appear to be a function of the dependent variable, weights are omitted in order to obtain unbiased, consistent and efficient estimates.

The main models build on each other as follows:

- Model 1 includes only migrant status and gender as the independent variables as well as controls to provide first insights into research question 1.
- Model 2 adds an interaction term between gender and migration status to answer research question 2.

Models 1 and 2 include only the controls so that they can give full estimates for non-/migrant differences. Explaining some of this difference is the intent of the models below, but models 1 and 2 aim to show potential frailty discrepancies as they exist in society, only adjusted for factors that are not in themselves the result of societal mechanisms[6]. They also provide a baseline for comparison with multilevel models and different specifications of missing values (see below).

- Models 3a1 and 3b1 are the same models computed separately by gender, functionally replacing interaction terms. In a mediation analysis, Models 3a2-3a5 and 3b2-3b5 add sets of potential mediators for the relationship between migrant status and frailty to the separate gender models. The variable sets are first added separately to see which of them account for how much of a possible immigrant-

[6] The control for other health problems is intended to specify the outcome measure correctly, not to adjust for differential distributions.

native gap in frailty: Health behaviours in models 2, economic variables in models 3, and social variables in models 4. Finally, all variable groups are included in the full models 5 to account for potential dependencies between them. This set of regression models will elucidate research question 3 and provide more information on research question 1.

- Models 4a and 4b are migrant-only variants of models 3a5 and 3b5. Here, the migration-specific variables are added to further answer research question 3.
- Models 5a and 5b exclude the fixed effects controls for countries. Instead, they include dummy variables for the geographical region and MIPEX rank and interactions of these with migration status to the full mediation models 3a5 and 3b5. This will give insight into research question 4.

Model diagnostics, robustness checks and sensitivity analyses

As is the case for linear regression, logistic regression estimates and standard errors are only reliable when a number of assumptions are met. First, the independent variables cannot be linearly dependent upon each other (DeMaris, 2004, p. 170). To assess whether this condition is met, I compute variance inflation factors (VIF) for all independent variables and compare them to the VIF threshold of 10 (DeMaris, 2004, p. 110). Secondly, there should be no strong outliers potentially distorting estimates (Stoltzfus, 2011). In linear regression, the standard way to check for outliers would be an investigation into residuals. However, residuals do not make sense in the same way for logistic regression due to the categorical outcome. I therefore investigate this assumption by checking the distribution of metric predictors.

Lastly, while the relationship between independent and dependent variables is modelled as non-linear, it is still modelled as linear for the log odds (DeMaris, 2004, p. 287). This means that I keep in mind the possibility of non-linear association between the metric predictors and the likelihood of frailty and test different specifications of these variables to best fit the model.

To check whether random effects multilevel models would be (more) appropriate, I also compute an empty model 0 and variants of models 1 and 2 as random effects models. Because no intraclass correlation coefficient (ICC) can be computed for multinomial models, I also estimate empty binarized models comparing pre-frailty to robustness, frailty to robustness, and frailty to pre-frailty. ICC values and loglikelihood ratio tests for the other models will help assess the relevance and appropriateness of specifying a random effects

model. In addition, I compare the estimated coefficients in the random effects specification of models 1 and 2 to their fixed effects counterparts as a robustness check.
As further robustness checks, I also provide models 1 and 2 a) without imputed values (i.e., treating imputed values as missing) and b) with listwise deletion of missing cases.

4. Findings

In this section, I present the results of the analysis outlined above. I start with descriptive statistics of the sample and sub-groups, then present the key findings from regression models and robustness checks.

4.1 Descriptive statistics

The analytical sample consists of n=49333 respondents with a mean age of 72 years. 55% of respondents are female; 9% of respondents are migrants, and for an additional 1%, migrant status could not be assessed. There is substantial variation in the percentage of migrant respondents per country; it ranges from only 1% in Italy and Portugal to about a quarter of all respondents in Luxembourg and Estonia. In most countries, information on country of birth (and thus, of migrant status) is missing for very few respondents (less than 2%). However, in Poland (8%) and Spain (3%), the percentage of people with missing information is elevated. In Poland, it is even higher than the percentage of known migrant respondents (see appendix, Table A1).
Table A2 (appendix) shows the weighted means and standard deviations for continuous variables as well as the weighted percentages and number of observations for categorical variables first by migrant status (excl. missing), then by gender, and for the whole sample. Here, weighting allows for better inferences to the population, instead of just describing the sample. For many of the variables, differences between the groups are insubstantial, while for others, they are noteworthy: According to the weighted sample, women are on average slightly older than men. They are also more likely to report health problems and be pre-frail or frail than men. In terms of health behaviours, while men are far more likely to drink alcohol than women, women are more likely to smoke[7]. Economically, women's households have lower average wealth and income. This might be connected to the lower percentage of women being retired and the higher percentage of them being homemakers or permanently sick compared to men. Women also tend to have lower levels of formal education than men.

[7] It is important to note, however, that a substantial portion of respondents has missing information on smoking behaviour. Therefore, discrepancies in the non-missing results might also be contorted by skews in the nonresponse.

Regarding social factors, women are more likely to be widowed and less likely to be married or partnered than men. This might in turn influence household finances.

For migrants, the economic picture is less clear. Despite not being employed more often than non-migrants, they have a higher average income. They are more commonly retired and less commonly in the homemaker or sick category than non-migrants. Their average household wealth is lower than that of non-migrants. However, they tend to have higher levels of formal education than non-migrants. Regarding health, non-migrants are more likely to report no health problems; for frailty, the picture is complicated, with migrants more often being in the pre-frail state and non-migrants having higher percentages in both robust and frail states. Migrants appear a bit more likely to be divorced than non-migrants, and less than a quarter of migrants in the weighted sample had citizenship in their country of residence, compared to 99.8% of non-migrants. Despite a slightly higher percentage of migrants being women than men, a higher percentage of men had citizenship compared to women. Big differences between migrants and non-migrants emerge in their area of residence. Compared to the full sample, migrants are strongly overrepresented in West European countries and underrepresented in East and especially in South European countries. Migrants are also more likely than natives to live in countries with a medium-inclusive health care policy according to MIPEX, and less likely than natives to reside in countries with low or high inclusivity.

For variables not mentioned here, differences between men and women or migrants and non-migrants are small or non-existent.

4.2 Regression models

As outlined in the methods section, the main models are multinomial logistic regression models utilizing fixed effects by controlling for country. These are presented first. For the mediation analysis, I also present KHB regression results that allow the comparison of (exponentiated) coefficients. Additionally, I show two sets of models that allow insight into specific types of variables: One without fixed country effects that looks at country characteristics instead, and one including only migrants to look at migration-specific influences on frailty. The main models are supplemented by a multilevel model to assess the relevance of the way in which country clusters are integrated into the analysis. Lastly, I check model assumptions and the robustness of results against models that a) use weighted data, b) do not use imputations for metric variables or c) where respondents with relevant item nonresponse are removed from the sample instead of including "missing" as a category for categorical variables.

4.2.1 Base and mediation models

Model 1 includes only migrant status and gender as the independent variables as well as controls; model 2 adds an interaction term between gender and migration status (Table 1; see appendix, Table A3 for the full table of coefficients). Both models show a statistically significant effect of being female on the likelihood of being pre-frail or frail as opposed to robust, with women having an odds ratio (OR) of 1.5 for pre-frailty and 1.9 for frailty. In other words, women's odds of being frail as opposed to pre-frail are almost twice as large as men's odds. In model 1, migrants are also shown to be statistically significantly more likely than non-migrants to be pre-frail or frail (ORs 1.1 and 1.2, respectively), though the effects are much smaller than for gender. In model 2, migrants' OR for pre-frailty is slightly smaller and no longer statistically significant (OR 1.1). Migrants' OR for frailty remains at a similar size. The ORs of the interaction terms are slightly above 1 for pre-frailty, slightly below 1 for frailty, and neither of them is statistically significant. The category for missing migrant information and its interaction with gender also do not yield any statistically significant results.

Of the control variables, the ORs for age are all positive and highly statistically significant, as are those for health problems. The fixed effect country controls, using Germany as a reference category, also mostly yield ORs above 1 and many of them are (sometimes highly) statistically significant. This is the case for the control variables across all models, including those presented below.

Table 1. Regression models with and without interaction of gender and migrant status

	Model 1		Model 2	
	Pre-frail	Frail	Pre-frail	Frail
Migrant	1.125**	1.237**	1.094	1.276*
	(0.002)	(0.001)	(0.102)	(0.018)
Female	1.512***	1.904***	1.501***	1.911***
	(0.000)	(0.000)	(0.000)	(0.000)
Female*migrant			1.052	0.962
			(0.494)	(0.760)
Observations	49333		49333	

Exponentiated coefficients; *p*-values in parentheses
All models controlling for missing migrant information, age, health problems, and countries (coefficients not shown).
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Mediation models are computed separately for men and for women. Migrants' ORs for (pre-)frailty in both sets of mediation models are shown in Table 2 (see appendix, Tables A4 and A5 for the full list of coefficients). In all men's models, migrants are slightly more likely to be pre-frail than non-migrants, but none of the coefficients are statistically significant. For frailty, ORs are slightly larger, but also only statistically significant in the non-mediated base model (OR 1.2, p-value 0.041). In the women's models, ORs are all slightly bigger, especially for pre-frailty. Here, the coefficient for migrants remains statistically significant in the three separate mediation models, though not in the full model including all mediators. For frailty, migrant women's OR is not statistically significant in the model with economic factors as well as in the full model. In all mediated models, migrants' (pre-)frailty ORs are smaller than in the base model, indicating mediation taking place.

Table 2. Odds Ratios of (pre-)frailty for migrants (ref. non-migrants) in models with different groups of mediators

	Basic model	Health be-haviours	Economic factors	Social capital	All mediators
Pre-frailty					
Women	1.16^{**}	1.15^{**}	1.12^{*}	1.13^{*}	1.10
	(0.005)	(0.009)	(0.034)	(0.017)	(0.069)
Men	1.09	1.08	1.04	1.07	1.03
	(0.121)	(0.189)	(0.482)	(0.204)	(0.639)
Frailty					
Women	1.25^{**}	1.22^{*}	1.17	1.22^{*}	1.15
	(0.007)	(0.015)	(0.057)	(0.019)	(0.109)
Men	1.24^{*}	1.19	1.15	1.23	1.10
	(0.041)	(0.098)	(0.201)	(0.056)	(0.392)

Exponentiated coefficients; *p*-values in parentheses
All models controlling for missing migrant information, age, health problems, and countries (coefficients not shown).
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In addition to mathematically adjusting for skew due to the number of variables included, KHB analysis decomposes the total effect of migration status on (pre-)frailty into an indirect component explained by the added mediators and a direct component that is not explained by the mediators. The direct component as estimated by KHB agrees with the estimates provided by the separate multinomial logistic regression models (for full KHB results, see appendix, Table A6). The total effect is slightly bigger than the direct effect in all models. Indirect effects themselves are only statistically significant in the models with economic factors and most of the full models. The percentage of the total migrant effect that is mediated varies from 6% mediated by health behaviours for pre-frailty in women to 69% mediated by all mediators for pre-frailty in men.

The regression models also estimate the effect of the mediators themselves (see appendix, Tables A4 and A5). Regarding health behaviours, smoking is positively associated with (pre-)frailty, and drinking alcohol is negatively associated with the two outcomes for both men and women. These effects are highly statistically significant. Many of the missing categories also have statistically significant results.

Economic factors showed the strongest mediation according to KHB analysis. Education levels above primary education are negatively associated with both pre-frailty and frailty, though for pre-frailty the effect of secondary education is only statistically significant for women. Employment status shows a heterogeneous picture. For both men and women, I find that being (self-)employed is positively associated with pre-frailty, but negatively with frailty compared to the "Retired" base category. The effect of unemployment is only statistically significant for pre-frailty in the women's model, where it is positive. Being in the "homemaker or permanently sick" category is positively associated with both outcomes, as is the "other" category; coefficients are mostly statistically significant. Missing information on occupational status has a strong negative association with both outcomes that is statistically significant in the women's model.

Household income and wealth are negatively associated with both outcomes for men and women, and are statistically significant, except for the income term for frailty in the men's model.

The third set of mediators represents social capital. For men and women, being widowed, divorced, or never married are all positively associated with both outcomes compared to being married or in a registered partnership. These coefficients are all statistically significant except for the effect of never having been married on frailty in the women's model. Having attended a social club in the past year is strongly negatively associated with the two outcomes; these effects are highly statistically significant. Not naming anybody as part of one's social network is positively associated with both outcomes, though the coefficient is not statistically significant for frailty in the men's model. Again, some of the missing categories also show statistically significant effects.

In the full models that include all mediators, the KHB-estimated total effect of migrant status is largest, as well as the percentages that are mediated. The specific variables' coefficients are largely very similar to those in the individual models and most retain the same direction as previously reported. Only the OR for missing employment status changes its direction in the women's model and is now strongly and positively associated with frailty (OR 3.95). Some of the effects of marital status lose their statistical significance, as does secondary education's effect on pre-frailty for women.

Looking at the percentage explained per mediator in the KHB analysis (see appendix, Table A7) reveals that for both outcomes, wealth plays the biggest explanatory role with about a fifth of the total migrant effect mediated for men, and about a third for women. It is followed by social club attendance for both men and women. However, missing information on club attendance confounds migrants' OR for frailty for women by 13%, whereas it is negligible for men. Another difference is that for women, whether they drink three or more alcoholic drinks per week mediates migrants' OR for pre-frailty by 10% and for frailty by 12% (2% and 5%, respectively, for men). The only sizeable confounding effect that is not for a "Missing" category is of higher levels of formal education. It is only slightly stronger for women than for men (6% vs. 5% and 8% vs. 7% for (pre-)frailty, respectively). It also stands out that being a homemaker or permanently sick has a mediating effect for women that it does not have for men (5% for pre-frailty and 2% for frailty among women; both 0% for men).

In Figures 2 and 3, below, men's and women's predicted probabilities for the three frailty states based on the full mediation models are shown. The graphs show how impactful age is for frailty, with the probability of being robust reducing consistently the older the respondents. There are also clear gender discrepancies: While for men, being robust is the most likely state up into their seventies, for women, pre-frailty is already about as likely as robustness at age 60, the youngest age in the sample. Within each gender, the trajectories for migrants and non-migrants are very close to identical. Only for women do the point estimates between non-/migrants differ slightly, with migrated women having slightly higher probabilities of being pre-frail and slightly lower probability of being robust than their non-migrated counterparts. However, confidence intervals overlap substantially at all ages.

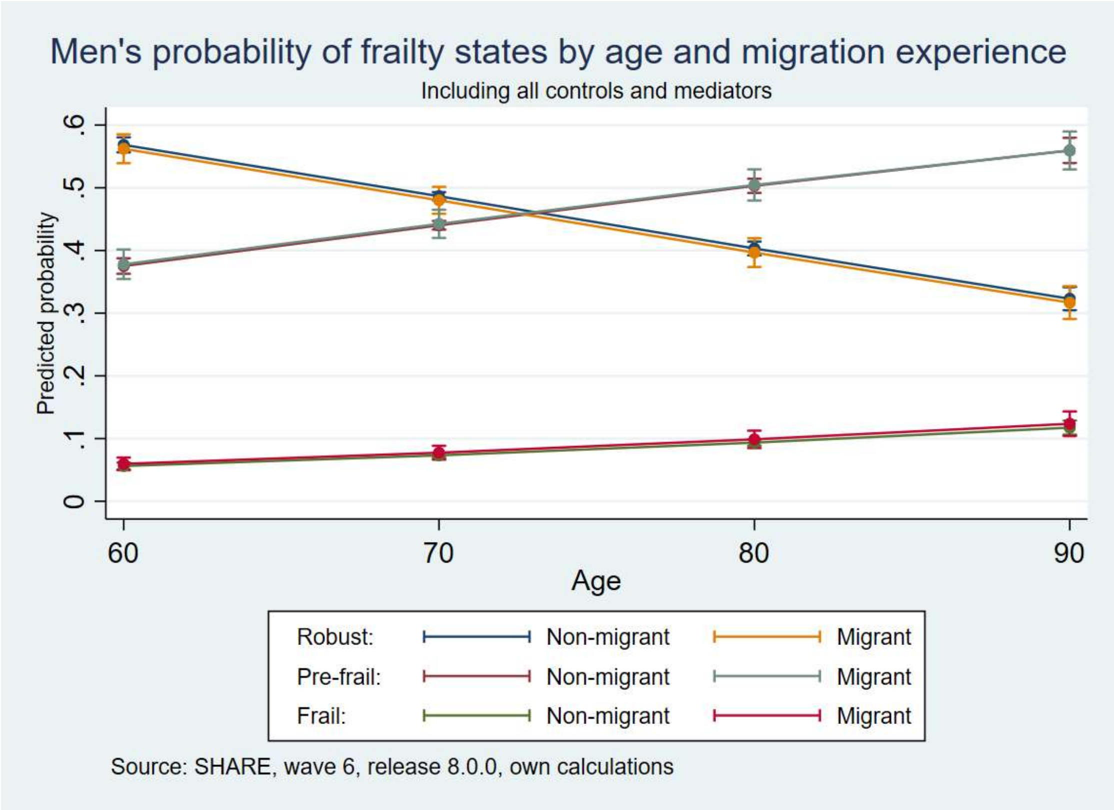

Figure 2. Men's probability of frailty states (all mediators included)

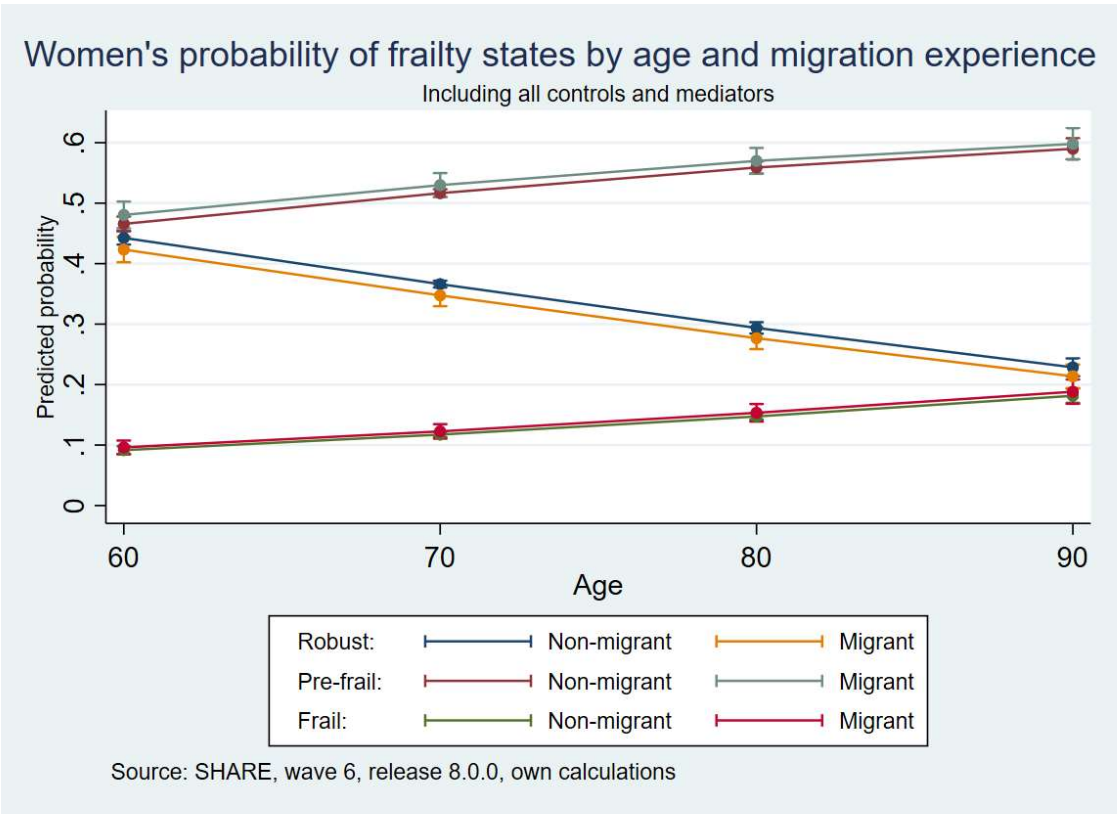

Figure 3. Women's probability of frailty states (all mediators included)

4.2.2 Moderation models

Migrant-only variants of the full mediation models allow for the addition of migration-specific variables: Years since migration, HDI level of the origin country, and whether the respondent has citizenship in their country of residence (Table 3). For the full table of coefficients, see appendix, Table A8.

For years since migration, the ORs for men and women and both outcomes are very close to 1 and not statistically significant. For citizenship, ORs are mostly slightly below 1, though again, none of them are statistically significant. Compared to migrants from high-HDI origin countries, those from medium-HDI countries have higher odds of being (pre-)frail, though the effects are not statistically significant for women being frail. Male migrants from the lowest HDI group have statistically significantly higher odds of being pre-frail (OR 1.68) but lower odds of being frail (OR 0.76, not statistically significant) than those from high-HDI countries. For low-HDI origin women, both ORs are slightly above 1 but not statistically significant.

Table 3. Migrant-only fixed effects regression models including migration-specific variables

	Men		Women	
	Pre-frail	Frail	Pre-frail	Frail
Years since migration to interview country	1.00	1.00	1.00	1.01
	(0.463)	(0.926)	(0.286)	(0.357)
HDI level of origin country (ref. highest)				
Medium	1.41*	1.91*	1.42*	1.42
	(0.030)	(0.032)	(0.021)	(0.145)
Least wealthy	1.64*	0.76	1.19	1.53
	(0.036)	(0.593)	(0.442)	(0.290)
Citizenship in country of interview (ref. no)				
Yes	1.01	0.90	0.92	0.79
	(0.935)	(0.724)	(0.523)	(0.231)
Observations	1882		2464	

Exponentiated coefficients; p-values in parentheses
Source: SHARE, wave 6, release 8.0.0, own calculations.
All models controlling for age, health problems, and countries as well as all mediators and missing categories for HDI level and citizenship (coefficients not shown).
Respondents aged 60+, unweighted, imputed values included.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Next, I assess the impact of health care inclusivity for migrants and of geographical region in the full mediation models, leaving out the fixed effects controls for countries (Table 4; see appendix, Table A9 for the full list of coefficients). In these models, which also contain interactions of migrant status with country groups, the ORs for (pre-)frailty due to being a migrant are mostly below 1 and none of them are statistically significant. In reference to Western Europe, (pre-)frailty odds are lower in Northern and Eastern Europe and higher in Southern Europe for both men and women. Almost all of these coefficients are statistically significant. The interaction of region with migration status shows that migrated men do not have the same lower risk of (pre-)frailty in Northern Europe or of frailty in Eastern Europe as their non-migrated counterparts. On the other hand, the higher (pre-)frailty risk in Southern Europe is (partially) compensated for by being a female migrant.

Compared to countries with high health care inclusivity, frailty odds are lower in countries with medium inclusivity for non-migrants, but less so for migrants. Pre-frailty odds are higher in countries with low inclusivity, an effect that is estimated to be even stronger for migrated women and less strong for migrated men. Most of the interaction terms are not statistically significant.

Table 4. Regression models with country group interactions (no fixed effects)

	Men		Women	
	Pre-frail	Frail	Pre-frail	Frail
Migrant	0.86	0.85	1.00	0.71
	(0.285)	(0.582)	(0.972)	(0.115)
Geographical region (ref. West)				
North	0.97	0.72**	0.87**	0.56***
	(0.590)	(0.003)	(0.004)	(0.000)
East	0.78***	0.71**	0.86**	0.74***
	(0.000)	(0.002)	(0.004)	(0.001)
South	1.24***	1.29**	1.25***	1.32***
	(0.000)	(0.008)	(0.000)	(0.000)
North*migrant	1.46*	2.29*	1.06	1.38
	(0.044)	(0.024)	(0.709)	(0.272)
East*migrant	1.01	1.86	0.77	0.76
	(0.944)	(0.093)	(0.147)	(0.359)
South*migrant	0.94	1.10	0.84	0.43*
	(0.766)	(0.840)	(0.380)	(0.030)
Health care inclusivity (ref. high inclusivity)				
Medium inclusivity	0.95	0.78**	0.94	0.74***
	(0.257)	(0.003)	(0.097)	(0.000)
Low inclusivity	1.28***	1.10	1.20***	0.98
	(0.000)	(0.360)	(0.000)	(0.777)
Med. inclusive*migrant	1.26	1.08	1.20	1.77*
	(0.133)	(0.821)	(0.176)	(0.020)
Low inclusive*migrant	0.90	0.60	1.23	1.67
	(0.589)	(0.165)	(0.219)	(0.086)
Observations	22322		27011	

Exponentiated coefficients; *p*-values in parentheses
Source: SHARE, wave 6, release 8.0.0, own calculations.
All models controlling for missing migrant information, age, health problems, and all mediators (coefficients not shown).
Respondents aged 60+, unweighted, imputed values included.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.2.3 Model diagnostics and robustness checks

To assess whether the assumption of no multicollinearity is met, I compute VIFs for all three sets of variables used in models: 1) the full mediation models, 2) the models containing country-level variables, and 3) the migration-specific models. VIF estimation is not supported as part of Stata's logistic regression commands used in the main analysis. However, multicollinearity is a feature of the variables and their values themselves, not of the model: Whether variables are colinear is not affected by what models they are included in. Therefore, I compute VIFs as postestimation of a non-meaningful linear model; the results are shown in the appendix, Table A10. For the migrant-only model, there is no indication of multicollinearity (all VIFs well below the customary threshold of 10). In the other two models, there is some correlation between missing information on migrant status and missing information on the level of education (each with a VIF of about 14). No other VIFs are concerning.

Other model assumptions could only be addressed through robustness checks. This includes the assumption of a linear relationship between the log of the dependent variables and the (potentially transformed) metric independent variables. For this, I also estimated models using a quadratic specification of age as well as linear specifications of wealth and income. None of these differences in specification made statistically significant differences to the model. For wealth and income, I used the transformations that best adjusted their distribution towards a normal distribution, which also helps reduce the impact of outliers – another condition that should be fulfilled for a reliable model.

In KHB analysis, weights did not appear to be a function of frailty. I also checked for robustness to weighting the data by computing weighted versions of the full mediation models (see appendix, Table A11). This did lead to some differences: In the weighted models, many coefficients are closer to 1 and have bigger p-values, though for some it is also the other way around, and a few even change direction. Differences vary between men's and women's models and the two outcomes; one difference that is persistent across the weighted models is that the effect of income is no longer statistically significant. In sum, not all previously estimated effects are robust to weighting the data.

Instead of using country fixed effects by including country variables in the model, the multilevel structure of the sample could also be represented through a random intercept model, where respondents are explicitly clustered in countries (see section 3.2.3). Log likelihood ratio for the empty multilevel model is highly significant, suggesting the cluster structure to be relevant to the data. Since ICC cannot be calculated for multinomial models, I can only report ICC for binarized models. For all three combinations of the three frailty states, ICC is small, but statistically significantly different from zero. For being pre-frail rather than robust, country ICC is 0.03; for being frail rather than robust, it is 0.10; and for being frail rather than pre-frail, it is 0.03. Comparing models 1 and 2 of the fixed effects regression with their random intercept counterparts shows that the results are robust to this difference in specification, with only very small reductions in effect size in the random intercept models (see appendix, Table A12).

Regarding imputed values, I compared models 1 and 2 to the same models estimated based on a sample without imputations (and, thus, with those respondents deleted who had missing values on metric variables, notably income and wealth). Some of the effects change slightly in size, though none change direction. In the non-imputed model 2 (with gender*migrant interaction), migrants' OR for frailty is no longer statistically significant, even though the effect is slightly bigger than in the imputed model. Otherwise, significance levels are robust (see appendix, Table A13).

Lastly, I checked for differences in results when missing categories are omitted and respondents with missing values on categorical variables are removed from the sample. In models 1 and 2, this concerns migrant status and health problem information. Here, effect sizes and significance levels of migrant status are increased slightly compared to the models including missing categories. At the same time, coefficients for being female are slightly smaller. The main terms for being a migrant are statistically significant in the model including the gender*migrant interaction, unlike in the original model 2. Still, the interaction term ORs themselves are not statistically significantly different from 1 (see appendix, Table A14).

5. Discussion

5.1 Interpretation

I will now interpret the results reported above to evaluate the four hypotheses formulated in section 2.4 and relate the findings to the literature.

> Hypothesis 1: Immigrated older adults will be more likely to be pre-frail and frail than native older adults.

The main take-away of model 1 is that migrants have slightly elevated odds for both frailty and pre-frailty compared to their non-migrated counterparts. This is also found in the total effects estimated in the KHB mediation models and supports hypothesis 1. The only exception is pre-frailty in men, where being a migrant did not yield statistically significant effects. In comparison to the literature reviewed in section 2, my results show a smaller health disadvantage for older migrants. However, they do lend support to the general impression that any health advantages observed in migrants at younger ages are not likely to be present at older ages.

> Hypothesis 2: The difference in frailty between men and women will be bigger for migrants than for natives.

In model 1, there is a clear difference in frailty between men and women, with men being more likely to be robust. Introducing an interaction of migration status and gender in model 2 does not show substantial effects: Migrants and non-migrants appear to be subject to the same impact of gender on frailty. In the models separated by gender, the estimated effects of migrant status are larger for women. Since the models are based on different samples, comparing their estimations directly is fraught. Therefore, there is no clear evidence of the effect of gender differing by migrant status, and hypothesis 2 is not supported.

The role of gender in the processes of shaping migrants' health was opaque in the literature and remains so in my findings. Some of the opposing theoretical and empirical arguments for gender differences in migrants' health could apply at the same time, resulting in impacts that cancel each other out, which would lead to the absence of a clear unidirectional effect. Alternatively, negative influences on female migrants' health might be mirrored by similar processes in the non-migrated population, so that the higher frailty risk for women is the same across non-/migrants. This study cannot answer whether any or all of these processes occur. The analysis results do, however, show clearly that migrated women carry the combined weight of higher frailty odds due to being a migrant and higher frailty odds due to being a woman, even if there is no additional intersecting (dis)advantage.

> Hypothesis 3a: Differences in health behaviours, socio-economic status, and social networks mediate immigrant-native differences in frailty.
>
> Hypothesis 3b: While selectivity effects and economic pathways affecting immigrant-native differences in frailty will be more important among men, social pathways and acculturation measures will be more important among women.

Previous research has focused on reporting coefficients of potential mediators, sometimes also providing descriptive statistics, but not on comparing nested models (e.g., Bousmah et al., 2019; Brothers et al., 2014; Lanari & Bussini, 2012). Consequently, it is hard to tell whether what they document are actual mediating effects or whether the additional predictors are associated with health outcome(s) independently of migrant status. However, the assumption of more advantageous health behaviours among migrants that is found in the theoretical literature appears unfounded in the SHARE sample: Here, based on descriptive statistics, migrants' behaviour around alcohol and smoking is not substantially more restrictive than that of non-migrants, contrary to the healthy migrant argument. Consequently, smoking appears to play little mediating role, even though its independent effect on (pre-)frailty is strong.

Despite the lack of differences in alcohol intake in descriptive statistics, it does mediate some of migrants' frailty odds in the multivariate models, especially for women. This indicates behavioural differences when gender and migration status intersect, which were not shown in the bivariate tables. Drinking alcohol, while usually considered an unhealthy habit, is in my analysis negatively associated with (pre-)frailty. This might suggest reverse causality if people in worse health consciously limit their alcohol intake, e.g., to avoid interfering with medications. The mediating effect drinking levels have on (female) migrants' frailty odds could also be understood like that: Migrated women in worse health choose to drink less alcohol.

Economic factors played the clearest role in mediating migrants' frailty odds in my analysis, with household wealth standing out as especially relevant. Descriptive statistics showed that migrants tended to have higher income than their non-migrated counterparts, but substantially lower wealth. Lower wealth and income were also associated with (pre-)frailty in descriptives. Interestingly, income had little to no mediating (or confounding) impact on migrants' frailty odds, perhaps because of the simultaneous inclusion of employment status. Wealth, however, mediated up to 37% of the total effect of migration (on pre-frailty for women). This is a sign that what might be generational wealth discrepancies generate health inequalities independently from someone's personal income level. The size of the effect also encourages the inclusion of wealth measures in research of this sort, despite its difficulties in measurement. The commonly used measure of "being able to make ends meet" (e.g., Bousmah et al., 2019; Lanari & Bussini, 2012), while interesting, does not give the same insight into the objective monetary situation. However, interpretation of the wealth measure in this study needs to be cautious, given that about half of the wealth estimates were imputed in some way. Further analysis of how large the portion of total wealth that had to be imputed was on average and of how close the multiple imputation estimates were for each respondent would be necessary to say whether this result is likely to be biased.

Education levels were also included among the economic variables and were the most important confounders out of all variables (excluding missing categories). Descriptively, migrants had on average higher levels of education, and those were in turn associated with lower levels of (pre-)frailty. While some of this may have been explained by differences in the age profiles, a confounding effect remained even in the age-controlled multivariate models. This indicates that migrants do not receive the same health protection from higher levels of education that non-migrants do. While not explicitly discussed by the studies reviewed above, this matches up with prior research on diminished returns to education for international migrants (see e.g. Brandt & Hagge, 2020; Roura, 2017; Walton et al., 2009). The last component of the economic variable set, employment status, provided ambiguous coefficients and no strong mediation (or confounding) effects.

Within the social variable set, "having attended a social club" showed the strongest mediation effect. While intended as a measure of social integration that might influence health, the causal relationship could also be reversed, as for alcohol consumption above: When people get frailer, they become less able to attend social events. Of course, both processes might reinforce each other in a feedback loop. Either way, it appears that (pre-)frail migrants are especially unlikely to attend social clubs and that this makes up considerable amounts of the difference between migrants and natives in (pre-)frailty. This finding fits well with the literature on migrants' lack of social support in the destination country (see, e.g., Makwarimba et al., 2010). The measure of empty social networks did not meaningfully contribute to the mediation despite substantial and statistically significant effects in most models, since differences between migrants and non-migrants were small on this variable. In future research, it might be fruitful to use a measure more concerned with how close respondents' support networks are to them spatially, which might differ by migration status, or with whether support networks are mainly familial (see Llácer et al., 2007). Marital status also does not contribute meaningfully to mediation or confounding.

Regarding the migration-specific variables, the main difference between men and women can be seen in the impact of origin countries: Even though generally, lower origin-country HDI was associated with higher odds of (pre-)frailty, male migrants from low-HDI countries seem to be exempt. This could indicate a protective selection effect (Bousmah et al., 2019) that was not applicable for women.

Contrary to the literature, years since migration do not appear to have any bearing on frailty states. This could suggest that at these advanced ages, age is more important and covers other effects. It might be interesting to investigate this and its collinearity with the age at migration further, e.g., with dummy variables for the timing of migration, either in calendar time (picking up policy changes) or over the life course. These further research questions would also benefit from longitudinal studies.

While citizenship was slightly negatively associated with frailty in accordance with the literature, the effects were not statistically significant. The relatively small size of the migrant sample and its heterogeneity given the wide range of contributing countries might obscure effects.

In sum, KHB analysis showed that including health behaviours, economic factors, and social capital in the models explained some of the total effect of migrant status. For the separate sets of variables, the indirect effect was only statistically significant for the set of economic factors, not for health behaviours and social capital by themselves. However, there were substantial percentages of the total effect mediated by social club attendance and, especially for women, by the frequency of alcohol consumption. Hypothesis 3a can therefore be deemed partially supported: Economic variables as well as two of the health behaviour and social indicators explain some of the migration effect.

This is the case for both men and women. It also appears that economic variables play at least as big of a role for women as they do for men, given the great proportion of mediation due to wealth. Other variables also show bigger confounding and mediation effects in the women's model than in the men's; social variables do not stand out specifically. For the acculturation-related variables in the migrants' models, citizenship and years since migration, there were no clear gender differences, either. However, wealth of the origin country did affect men and women differently, even if the size of the sample did not suffice for statistically significant results. This gives a small indication of positive health selection for men from low-HDI backgrounds, but not for women. Most of hypothesis 3b is not supported: Mediation effects appear generally stronger for women than for men, without there being a clear difference in the type of confounder or mediator that is most relevant. Acculturation measures did not show differential impact, either, though there is an indication for stronger selectivity effects for men.

> Hypothesis 4: Immigrant-native differences in frailty will be moderated by countries' region within Europe and the inclusivity of policies for migrants' access to health care.

The models for country-specific variables made clear that both health care policy and geographical region have an impact on frailty in both the total population and for migrants specifically. These impacts do not always match up: In Northern and Eastern Europe, migrant men were shown to be more vulnerable than native men, and in Southern Europe, migrant women were shown to be less vulnerable than their non-migrated counterparts. In terms of health care inclusivity, the differences between high- and medium-/low-inclusivity contexts were mostly less advantageous for migrants compared to natives in in the medium- and low-inclusivity groups. Thus, hypothesis 4 is supported: The country contexts under study did have moderating effects on migrants' (pre-)frailty odds. The relevance of country context is also supported by the multi-level model. There, ICCs and log likelihood ratio emphasize the need to adjust the analysis for the clustered sample structure. However, since the results are similar to fixed effects models, this is an adequate specification given the small number of clusters.

My finding of a smaller difference in frailty between migrants and non-migrants in Southern Europe fits in with Walkden et al. (2018), who attribute it to a different composition of the migrant flows to the South compared to the rest of Europe. However, I only find this effect in the women's model. This raises further questions about who is migrating to Southern Europe for what reasons, and whether that differs between men and women. Conversely, I replicate Brothers et al.'s (2014) finding of frailer migrants in North-western Europe mostly for men. They attribute regional differences to discrepancies in the overall health of the population. This does not seem to be the full story here, given that the non-migrated Northern European frailty advantage is even stronger for women in my data.

Walkden et al. (2018) find a stronger association between frailty and migrant status the lower the health care inclusivity. I did not replicate this finding, as in my analysis only one of the relevant interaction terms is statistically significant and migrated men even appear somewhat less vulnerable in low-inclusivity contexts. More research on health care policies' effect on migrants could shed further light on this issue.

5.2 Strengths and limitations

Any statistical model can only approximate real connections, and its approximations are necessarily limited. There is already potential bias in the data collection: Both unit and item nonresponse, as detailed in section 3.1.2, might be systematic and thus skew the results. Systematic unit nonresponse is offset by weighting in the descriptive statistics. Since the regression models already use the variables also employed in weighting, and weighting cannot account for the relationships between variables as well as it can for simple frequencies, it is not quite clear whether a weighted regression model is more accurate than an unweighted one. The differences in results do, however, caution against broad claims based on the unweighted models.

In terms of item nonresponse, the statistically significant results of missing categories and collinearity between some of them suggest a not completely random pattern[8]. This implies that removing all respondents with missing items from the analytical sample would have biased the results. Another option for dealing with item nonresponse would have been using multiple imputation, which is currently not supported for all steps of my analysis. Using the mean of multiple imputation estimates for metric variables is the next best option available. Given the robustness of results against different treatments of missing values, however, item nonresponse does not seem to have skewed the results much.

Missing categories' significant results are also a warning sign in another direction: With increasing numbers of variables included in the model, the likelihood of observing spurious 'connections' as statistically significant increases, especially when also estimating more and more models. As such, minor differences in significance levels and effect sizes should not be over-interpreted.

While one danger is over-specifying the model and including too many variables, another is the vagueness that comes with categories that are too large. In this analysis, this is especially apparent for the categories of employment status. The statistically significant results of the "Other" category beget the question who exactly is represented in them. Similarly, combining homemakers and those who are permanently sick into one group, while also done in the literature and a useful way of dealing with low numbers in either category, obscures differences between them. Also, it appears likely that women more often consider themselves homemakers than men, making the category mean different things in the women's vs. the men's models. These examples illustrate that, while all models are specified to give the most insight into processes happening in the general population, they might not be reliable for each small effect. However, many more general trends hold across different models, which increases the confidence in their interpretation.

In the same vein, interpreting all country fixed effects and their significance levels themselves would be very vulnerable to spurious effects, as explained above. However, the robustness checks for random effects and associated tests for clustering did indicate that accounting for the clustered structure of respondents was necessary, and that using fixed effects provided an agreeable way of doing so. Still, further research into country effects would be

[8] Collinearity also means not being able to accurately estimate either variable's effect size. This is not as worrying here, given that the missing categories are not interpretable either way.

very useful; within the scope of this analysis, only the broadest of trends could be covered. For example, there was no room for models that let the migration effect itself vary by country or country group, which would be an interesting research project in itself.

More generally, the outcome variable also must be interpreted with care. While the frailty phenotype used here is well-supported by evidence, other studies in this area instead used a frailty index. In that regard, my work has added to a fuller understanding of frailty. However, the differences in operationalisation complicate the comparison of effect sizes and might account for some differences in results.

Additionally, controlling for health problems while looking at frailty ensures that what is measured is not bad health itself, but frailty as an independent predictor of morbidity and mortality. Thus, the analysis is more focused on a physical and mental state that itself might not be distressing or considered in need of treatment but provides an early warning sign (see discussion in section 1.1). Therefore, it has to be kept in mind in all interpretations that potential differences in the frequency of health problems have already been controlled for and are not represented in the frailty odds. More general health inequalities might therefore be underestimated by this measure.

6. Conclusion

In this thesis, I have found a slightly elevated risk of (pre-)frailty for older migrants in Europe compared to their non-migrated peers. This effect is stronger and less likely to be a statistical artifact for women. It is partially explained by economic and behavioural differences between migrants and non-migrants. Both origin countries and host countries play a role in how frailty differences manifest. These results support theories of structural disadvantage for immigrants, especially in terms of (generational) wealth and returns on education. Some support for selectivity and acculturation arguments could be found as well, though the main measure of acculturation, duration of stay, did not have substantial effects. Most of the mediation and moderation analyses yielded similar results for men and women, but the effect of migrating from a low-HDI origin country appears to differ by gender, and wealth had an even bigger mediating effect on migrant women's frailty odds than it did on men's. In summary, migrant women's frailty odds are slightly more disadvantageous than for their male counterparts, but the processes influencing their frailty appear to be similar.

6.1 Future research

As shown in both this study and in previous research, the SHARE infrastructure is an invaluable tool for analysing older migrants' health in Europe. However, the life history interviews in wave 7 and the fieldwork restrictions due to the COVID-19 pandemic mean that since wave 6 in 2015, no data that are easily comparable to previous waves have been published, resulting in few publications with more recent data. Given that Europe's older migrant population is growing continuously, more current data are needed, especially for fine-grained analysis of migrant sub-groups. It would therefore be an important step to develop operationalisations of key health measures, including frailty, whose data requirements are met by the shorter telephone surveys conducted in the pandemic years. This would also pave the way for expanded longitudinal analyses.

The country-specific effects documented in this thesis are complex and could only be touched upon cursorily. To analyse them more precisely, studies on the specific impact of policy measures in single countries or in comparative case studies continue to be needed. Multi-level models would be able to integrate country-specific variables more precisely than the models used here, as well. However, meaningful multi-level modelling needs a larger set of clusters. To this end, continuing the expansion of SHARE coverage to more countries and its integration with other (inter)national surveys of the older population is important.

6.2 Policy implications

My results support the existing evidence that European governments cannot expect their migrant populations to age with lower levels of health problems than the non-migrated populations, despite potential "healthy migrant effects" at younger ages. Persistent inequalities in wealth as well as diminished health returns on education for migrants contribute to frailty risks for this group. The study has also made apparent that migrants are a heterogeneous group affected by general social and/or physical conditions. Migrant women show the same frailty trajectory as non-migrated women, which is substantially more disadvantageous than that of men.

However, it is also clear that the proportion of older people becoming frail is low at most ages, while high percentages of all groups can be categorized as pre-frail. This provides the option of early intervention to health care providers and social programmes: Before frailty sets in, observing its components, like weakness or a decreased appetite, gives opportunities to intervene. Given that the growing numbers of ageing migrants are likely to be at least as (pre-)frail as their non-migrated counterparts, it becomes imperative that health care interventions for older people are designed with migrants in mind. Whether that means trainings for intercultural competence in social workers and nurses or activities offered in different languages, the migrant populations of Europe cannot be left out in discussions of healthily ageing societies.

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

Abraído-Lanza, A. F., Echeverría, S. E., & Flórez, K. R. (2016). Latino Immigrants, Acculturation, and Health: Promising New Directions in Research. *Annual Review of Public Health*, *37*(1), 219–236. https://doi.org/10.1146/annurev-publhealth-032315-021545

Alidu, L., & Grunfeld, E. A. (2018). A systematic review of acculturation, obesity and health behaviours among migrants to high-income countries. *Psychology & Health*, *33*(6), 724–745. https://doi.org/10.1080/08870446.2017.1398327

Aner, K., & Dosch, E. (2023). Lebenslagen oder Intersektionalität? Gerontologie aus kritisch feministischer Perspektive. *Zeitschrift fur Gerontologie und Geriatrie*, *56*(1), 9–12. https://doi.org/10.1007/s00391-022-02149-w

Bergmann, M., & Börsch-Supan, A. (Eds.). (2021). *Share wave 8 methodology: Collecting cross-national survey data in times of COVID-19*. Munich. Munich Research Institute for the Economics of Aging and SHARE Analyses (MEA). https://share-eric.eu/fileadmin/user_upload/Methodology_Volumes/SHARE_Methodenband_WEB_Wave8_MFRB.pdf

Bergmann, M., Börsch-Supan, A., & Scherpenzeel, Annette (Eds.). (2019). *Share wave 7 methodology: Panel innovations and life histories*. Munich. Munich Research Institute for the Economics of Aging and SHARE Analyses (MEA). https://share-eric.eu/fileadmin/user_upload/Methodology_Volumes/SHARE_Methodenband_A4_WEB_Wave7MFRB.pdf

Bergmann, M., Kneip, T., Luca, G. de, & Scherpenzeel, Annette. (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). München. https://share-eric.eu/fileadmin/user_upload/SHARE_Working_Paper/WP_Series_41_2019_Bergmann_et_al.pdf

Bergmann, M., Luca, G. de, & Scherpenzeel, Annette. (2017). Sampling design and weighting strategies in SHARE Wave 6. In F. Malter & A. Börsch-Supan (Eds.), *SHARE Wave 6: Panel innovations and collecting Dried Blood Spots* (pp. 77–93). Munich: Munich Research Institute for the Economics of Aging and SHARE Analyses (MEA).

Bloom, D. E., & Luca, D. L. (2016). The Global Demography of Aging. In J. Piggott & A. Woodland (Eds.), *Handbook of the Economics of Population Aging* (Vol. 1, pp. 3–56). Elsevier. https://doi.org/10.1016/bs.hespa.2016.06.002

Börsch-Supan, A. (2022a). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 1: Release Version: 8.0.0* [Data set]. SHARE-ERIC. https://doi.org/10.6103/SHARE.W1.800

Börsch-Supan, A. (2022b). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 2: Release Version: 8.0.0* [Data set]. SHARE-ERIC. https://doi.org/10.6103/SHARE.W2.800

Börsch-Supan, A. (2022c). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 4: Release Version: 8.0.0* [Data set]. SHARE-ERIC. https://doi.org/10.6103/SHARE.W4.800

Börsch-Supan, A. (2022d). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5: Release Version: 8.0.0* [Data set]. SHARE-ERIC. https://doi.org/10.6103/SHARE.W5.800

Börsch-Supan, A. (2022e). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6: Release Version: 8.0.0* [Data set]. SHARE-ERIC. https://doi.org/10.6103/SHARE.W6.800

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S., & Zuber, S. (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*, *42*(4), 992–1001. https://doi.org/10.1093/ije/dyt088

Bousmah, M.-A.-Q., Combes, J.-B. S., & Abu-Zaineh, M. (2019). Health differentials between citizens and immigrants in Europe: A heterogeneous convergence. *Health Policy*, *123*(2), 235–243. https://doi.org/10.1016/j.healthpol.2018.12.005

Božić, S., & Kuti, S. (2016). New international borders – old social spaces: Transnational migrant networks across the boundaries of post-socialist Croatia. *Österreichische Zeitschrift Für Soziologie*, *41*(4), 409–426. https://doi.org/10.1007/s11614-016-0246-2

Brandt, J., & Hagge, K. S. (2020). Education and social support: Do migrants benefit as much as natives? *Comparative Migration Studies*, *8*(1), Article 41. https://doi.org/10.1186/s40878-020-00199-w

Brothers, T. D., Theou, O., & Rockwood, K. (2014). Frailty and migration in middle-aged and older Europeans. *Archives of Gerontology and Geriatrics*, *58*(1), 63–68. https://doi.org/10.1016/j.archger.2013.07.008

Cela, E., & Di Barbiano Belgiojoso, E. (2021). Ageing in a foreign country: Determinants of self-rated health among older migrants in Italy. *Journal of Ethnic and Migration Studies*, *47*(15), 3677–3699. https://doi.org/10.1080/1369183X.2019.1627863

Chang, S.-F., & Lin, P.-L. (2015). Frail phenotype and mortality prediction: A systematic review and meta-analysis of prospective cohort studies. *International Journal of Nursing Studies*, *52*(8), 1362–1374. https://doi.org/10.1016/j.ijnurstu.2015.04.005

Chang, S.-F., Lin, H.-C., & Cheng, C.-L. (2018). The Relationship of Frailty and Hospitalization Among Older People: Evidence From a Meta-Analysis. *Journal of Nursing Scholarship*, *50*(4), 383–391. https://doi.org/10.1111/jnu.12397

Ciobanu, R. O., Fokkema, T., & Nedelcu, M. (2017). Ageing as a migrant: Vulnerabilities, agency and policy implications. *Journal of Ethnic and Migration Studies*, *43*(2), 164–181. https://doi.org/10.1080/1369183X.2016.1238903

Constant, A. F., García-Muñoz, T., Neuman, S., & Neuman, T. (2018). A "healthy immigrant effect" or a "sick immigrant effect"? Selection and policies matter. *The European Journal of Health Economics*, *19*(1), 103–121. https://doi.org/10.1007/s10198-017-0870-1

Constant, A. F., & Milewski, N. (2021). Self-selection in physical and mental health among older intra-European migrants. *The Journal of the Economics of Ageing*, *19*, Article 100322. https://doi.org/10.1016/j.jeoa.2021.100322

Curran, S. R., Shafer, S., Donato, K. M., & Garip, F. (2006). Mapping Gender and Migration in Sociological Scholarship: Is It Segregation or Integration? *International Migration Review*, *40*(1), 199–223. https://doi.org/10.1111/j.1747-7379.2006.00008.x

DeMaris, A. (2004). *Regression with social data: Modeling continuous and limited response variables*. *Wiley series in probability and statistics*. Wiley-Interscience. https://doi.org/10.1002/0471677566

Etman, A., Burdorf, A., van der Cammen, T. J. M., Mackenbach, J. P., & van Lenthe, F. J. (2012). Socio-demographic determinants of worsening in frailty among community-

dwelling older people in 11 European countries. *Journal of Epidemiology and Community Health*, *66*(12), 1116–1121. https://doi.org/10.1136/jech-2011-200027
Finch, B. K., Frank, R., & Vega, W. A. (2004). Acculturation and Acculturation Stress: A Social-Epidemiological Approach to Mexican Migrant Farmworkers' Health. *International Migration Review*, *38*(1), 236–262. https://doi.org/10.1111/j.1747-7379.2004.tb00195.x
Franse, C. B., van Grieken, A., Qin, L., Melis, R. J. F., Rietjens, J. A. C., & Raat, H. (2018). Ethnic differences in frailty: A cross-sectional study of pooled data from community-dwelling older persons in the Netherlands. *BMJ Open*, *8*(8), Article e022241. https://doi.org/10.1136/bmjopen-2018-022241
Fresnoza-Flot, A. (2022). Gender Gaps in Migration Studies. In E. Fornalé (Ed.), *World Trade Institute advanced studies: Vol. 10. Gender equality in the mirror: Reflecting on power, participation and global justice* (pp. 110–137). Brill Nijhoff. https://doi.org/10.1163/9789004467682_007
Fried, L. P., Tangen, C. M., Walston, J., Newman, A. B., Hirsch, C., Gottdiener, J., Seeman, T., Tracy, R., Kop, W. J., Burke, G., & McBurnie, M. A. (2001). Frailty in older adults: Evidence for a phenotype. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *56*(3), 146–156. https://doi.org/10.1093/gerona/56.3.m146
Garcia, M. A., Reyes, A. M., & Rote, S. (2019). Disability and the Immigrant Health Paradox: Gender and Timing of Migration. In W. A. Vega, J. L. Angel, L. M. F. Gutiérrez Robledo, & K. S. Markides (Eds.), *Contextualizing Health and Aging in the Americas* (pp. 249–269). Springer International Publishing. https://doi.org/10.1007/978-3-030-00584-9_12
Georges, D., Buber-Ennser, I., Rengs, B., Kohlenberger, J., & Doblhammer, G. (2021). Health determinants among refugees in Austria and Germany: A propensity-matched comparative study for Syrian, Afghan, and Iraqi refugees. *PloS One*, *16*(4), Article e0250821. https://doi.org/10.1371/journal.pone.0250821
Giannoni, M., Franzini, L., & Masiero, G. (2016). Migrant integration policies and health inequalities in Europe. *BMC Public Health*, *16*(1), Article 463. https://doi.org/10.1186/s12889-016-3095-9
Gkiouleka, A., & Huijts, T. (2020). Intersectional migration-related health inequalities in Europe: Exploring the role of migrant generation, occupational status & gender. *Social Science & Medicine*, *267*, Article 113218. https://doi.org/10.1016/j.socscimed.2020.113218
Gómez-Costilla, P., García-Prieto, C., & Somarriba-Arechavala, N. (2022). Aging and Gender Health Gap: A Multilevel Analysis for 17 European Countries. *Social Indicators Research*, *160*(2-3), 1051–1069. https://doi.org/10.1007/s11205-020-02595-2
Gorman, B. K., Lariscy, J. T., & Kaushik, C. (2014). Gender, acculturation, and smoking behavior among U.S. Asian and Latino immigrants. *Social Science & Medicine*, *106*, 110–118. https://doi.org/10.1016/j.socscimed.2014.02.002
Gorman, B. K., Read, J. G., & Krueger, P. M. (2010). Gender, acculturation, and health among Mexican Americans. *Journal of Health and Social Behavior*, *51*(4), 440–457. https://doi.org/10.1177/0022146510386792
Greene, J., & Baron, J. (2001). Intuitions about declining marginal utility. *Journal of Behavioral Decision Making*, *14*(3), 243–255. https://doi.org/10.1002/bdm.375
Guillot, M., Khlat, M., Elo, I., Solignac, M., & Wallace, M. (2018). Understanding age variations in the migrant mortality advantage: An international comparative perspective. *PloS One*, *13*(6), Article 0199669. https://doi.org/10.1371/journal.pone.0199669

Hankivsky, O. (2012). Women's health, men's health, and gender and health: Implications of intersectionality. *Social Science & Medicine*, *74*(11), 1712–1720. https://doi.org/10.1016/j.socscimed.2011.11.029

Hornby-Turner, Y. C., Peel, N. M., & Hubbard, R. E. (2017). Health assets in older age: A systematic review. *BMJ Open*, *7*(5), Article e013226. https://doi.org/10.1136/bmjopen-2016-013226

Hossin, M. Z. (2020). International migration and health: It is time to go beyond conventional theoretical frameworks. *BMJ Global Health*, *5*(2), Article e001938. https://doi.org/10.1136/bmjgh-2019-001938

Huddleston, T., Bilgili, Ö., Joki, A.-L., & Vankova, Z. (2015). *Migrant Integration Policy Index 2015*. Barcelona Centre for International Affairs (CIDOB); Migration Policy Group (MPG).

Huisman, M., Kunst, A. E., & Mackenbach, J. P. (2003). Socioeconomic inequalities in morbidity among the elderly; a European overview. *Social Science & Medicine*, *57*(5), 861–873. https://doi.org/10.1016/S0277-9536(02)00454-9

Jacobsen, F. F., Glasdam, S., Schopman, L. M., Sodemann, M., van den Muijsenbergh, M. E. T. C., & Ågotnes, G. (2023). Migration and health: Exploring healthy ageing of immigrants in European societies. *Primary Health Care Research & Development*, *24*, Article e10. https://doi.org/10.1017/S1463423623000014

Jasso, G., Massey, D. S., Rosenzweig, M. R., & Smith, J. P. (2004). Immigrant Health: Selectivity and Acculturation. In National Research Council (U.S.) (Ed.), *Understanding racial and ethnic differences in health in late life: A research agenda* (pp. 227–266). National Academies Press.

Karlson, K. B., Holm, A., & Breen, R. (2012). Comparing Regression Coefficients Between Same-sample Nested Models Using Logit and Probit. *Sociological Methodology*, *42*(1), 286–313. https://doi.org/10.1177/0081175012444861

Kneale, D., Henley, J., Thomas, J., & French, R. (2021). Inequalities in older LGBT people's health and care needs in the United Kingdom: A systematic scoping review. *Ageing and Society*, *41*(3), 493–515. https://doi.org/10.1017/S0144686X19001326

Kristiansen, M., Razum, O., Tezcan-Güntekin, H., & Krasnik, A. (2016). Aging and health among migrants in a European perspective. *Public Health Reviews*, *37*, Article 20. https://doi.org/10.1186/s40985-016-0036-1

Krobisch, V., Gebert, P., Gül, K., & Schenk, L. (2021). Women bear a burden: Gender differences in health of older migrants from Turkey. *European Journal of Ageing*, *18*(4), 467–478. https://doi.org/10.1007/s10433-020-00596-1

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

Lazarevič, P. (2023). Biases in Assertions of Self-Rated Health. *Comparative Population Studies*, *48*. https://doi.org/10.12765/CPoS-2023-04

Liversage, A., & Mizrahi Mirdal, G. (2017). Growing old in exile: A longitudinal study of migrant women from Turkey. *Journal of Ethnic and Migration Studies*, *43*(2), 287–302. https://doi.org/10.1080/1369183X.2016.1238910

Llácer, A., Zunzunegui, M. V., del Amo, J., Mazarrasa, L., & Bolumar, F. (2007). The contribution of a gender perspective to the understanding of migrants' health. *Journal of Epidemiology and Community Health*, *61*(Suppl. 2), ii4-10. https://doi.org/10.1136/jech.2007.061770

Loi, S., & Hale, J. M. (2019). Migrant health convergence and the role of material deprivation. *Demographic Research*, *40*, Article 32, 933–962. https://doi.org/10.4054/DemRes.2019.40.32

Loichinger, E., & Weber, D. (2016). Trends in Working Life Expectancy in Europe. *Journal of Aging and Health*, *28*(7), 1194–1213. https://doi.org/10.1177/0898264316656509

Lu, Y., & Li, X. (2020). Documentation status, gender, and health selection of immigrants: Evidence from Mexican–US migration. *Population, Space and Place*, *26*(7), Article e2333. https://doi.org/10.1002/psp.2333

Luca, G. de, Celidoni, M., & Trevisan, E. (2015). Item nonresponse and imputation strategies in SHARE Wave 5. In F. Malter & A. Börsch-Supan (Eds.), *Share wave 5: Innovations & Methodology* (pp. 85–100). Munich: Munich Research Institute for the Economics of Aging and SHARE Analyses (MEA).

Makwarimba, E., Stewart, M. J., Beiser, M., Neufeld, A., Simich, L., & Spitzer, D. (2010). Social support and health: Immigrants' and refugees' perspectives. *Diversity in Health and Care*, *7*(2), 91–103.

Malmusi, D., Borrell, C., & Benach, J. (2010). Migration-related health inequalities: Showing the complex interactions between gender, social class and place of origin. *Social Science & Medicine*, *71*(9), 1610–1619. https://doi.org/10.1016/j.socscimed.2010.07.043

Malter, F., & Sand, G. (2017). Fieldwork monitoring and survey participation in sixth wave of SHARE. In F. Malter & A. Börsch-Supan (Eds.), *SHARE Wave 6: Panel innovations and collecting Dried Blood Spots* (pp. 94–162). Munich: Munich Research Institute for the Economics of Aging and SHARE Analyses (MEA).

Malva, J. O., Bultje, D., Traver-Salcedo, V., Carriazo, A. M., Illario, M., Luca, V. de, Wozniak, L. A., Bousquet, J., Guldemond, N., Costa, I., Canhão, H., Costa, E., & Farrel, J. (2023). Flagship initiatives for healthy living and active aging in Europe: The European Innovation Partnership on Active and Healthy Ageing and the Reference Sites. In P. J. Oliveira & J. O. Malva (Eds.), *Aging: From fundamental biology to societal impact* (pp. 31–47). Academic Press. https://doi.org/10.1016/B978-0-12-823761-8.00008-2

Matthews, R. J., Jagger, C., & Hancock, R. M. (2006). Does socio-economic advantage lead to a longer, healthier old age? *Social Science & Medicine*, *62*(10), 2489–2499. https://doi.org/10.1016/j.socscimed.2005.11.019

McNeish, D., & Kelley, K. (2019). Fixed effects models versus mixed effects models for clustered data: Reviewing the approaches, disentangling the differences, and making recommendations. *Psychological Methods*, *24*(1), 20–35. https://doi.org/10.1037/met0000182

Moehring, K. (2012). *The fixed effects approach as an alternative to multilevel analysis for cross-national analyses* (GK SOCLIFE Working Paper Series). https://doi.org/10.31235/osf.io/3xw7v

Moullan, Y., & Jusot, F. (2014). Why is the 'healthy immigrant effect' different between European countries? *European Journal of Public Health*, *24*(Suppl. 1), 80–86. https://doi.org/10.1093/eurpub/cku112

Munich Research Institute for the Economics of Aging and SHARE Analyses (Ed.). (2015). *SHARE wave 6 questionnaire*. https://share-eric.eu/fileadmin/user_upload/Questionnaires/Q-Wave_6/w6_en_capi_main-Generic.pdf

Munich Research Institute for the Economics of Aging and SHARE Analyses (Ed.). (2017). *SHARE wave 7 questionnaire*. https://share-eric.eu/fileadmin/user_upload/Questionnaires/Q-Wave_7/w7_en_capi_main-Generic.pdf

Munich Research Institute for the Economics of Aging and SHARE Analyses (Ed.). (2022). *SHARE release guide 8.0.0*. https://share-eric.eu/fileadmin/user_upload/Release_Guides/SHARE_release_guide_8-0-0.pdf

Nagi, S. Z. (1976). An epidemiology of disability among adults in the United States. *The Milbank Memorial Fund Quarterly. Health and Society*, *54*(4), 439–467.

OECD (Ed.). (2020). *Compare your income: Methodology and conceptual issues*. https://www.oecd.org/statistics/Compare-your-income-methodology.pdf

Parker, M., Bucknall, M., Jagger, C., & Wilkie, R. (2020). Population-based estimates of healthy working life expectancy in England at age 50 years: Analysis of data from the English Longitudinal Study of Ageing. *The Lancet. Public Health*, *5*(7), e395-e403. https://doi.org/10.1016/S2468-2667(20)30114-6

Pedersen, A. B., Mikkelsen, E. M., Cronin-Fenton, D., Kristensen, N. R., Pham, T. M., Pedersen, L., & Petersen, I. (2017). Missing data and multiple imputation in clinical epidemiological research. *Clinical Epidemiology*, *9*, 157–166. https://doi.org/10.2147/CLEP.S129785

Ramos-Herrera, M. d. C., & Sosvilla-Rivero, S. (2020). Fiscal Sustainability in Aging Societies: Evidence from Euro Area Countries. *Sustainability*, *12*(24), Article 10276. https://doi.org/10.3390/su122410276

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*, *3*(4), 297–303. https://doi.org/10.1046/j.1365-3156.1998.00233.x

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), Article 57. https://doi.org/10.1186/s12916-018-1044-4

Rose, P. R. (2021). *Health equity, diversity, and inclusion: Context, controversies, and solutions* (2nd ed.). Jones & Bartlett Learning.

Roura, M. (2017). Unravelling migrants' health paradoxes: A transdisciplinary research agenda. *Journal of Epidemiology and Community Health*, *71*, 870–873. https://doi.org/10.1136/jech-2016-208439

Rouzet, D., Caldera Sánchez, A., & Renault, T. (2019). *Fiscal challenges and inclusive growth in ageing societies* (OECD Economic Policy Papers No. 27). https://www.oecd-ilibrary.org/docserver/c553d8d2-en.pdf?expires=1679915504&id=id&accname=guest&checksum=CD8E03115473D835CE77248D1378ADA6

Santos-Eggimann, B., Cuénoud, P., Spagnoli, J., & Junod, J. (2009). Prevalence of frailty in middle-aged and older community-dwelling Europeans living in 10 countries. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *64*(6), 675–681. https://doi.org/10.1093/gerona/glp012

Scherpenzeel, Annette, Axt, K., Bergmann, M., Douhou, S., Oepen, A., Sand, G., Schuller, K., Stuck, S., Wagner, M., & Börsch-Supan, A. (2020). Collecting survey data among the 50+ population during the COVID-19 outbreak: The Survey of Health, Ageing and Retirement in Europe (SHARE). *Survey Research Methods*, *14*(2), 217–221. https://doi.org/10.18148/SRM/2020.V14I2.7738

Schofield, T. (2018). *A Sociological Approach to Health Determinants*. Cambridge University Press. https://doi.org/10.1017/CBO9781316178713

Schouten, B. C., Cox, A., Duran, G., Kerremans, K., Banning, L. K., Lahdidioui, A., van den Muijsenbergh, M., Schinkel, S., Sungur, H., Suurmond, J., Zendedel, R., &

Krystallidou, D. (2020). Mitigating language and cultural barriers in healthcare communication: Toward a holistic approach. *Patient Education and Counseling*, *103*(12), 2604–2608. https://doi.org/10.1016/j.pec.2020.05.001

SHARE-ERIC (Ed.). (2023a). *Conditions of Use*. https://share-eric.eu/data/data-access/conditions-of-use

SHARE-ERIC (Ed.). (2023b). *FAQs & Support*. https://share-eric.eu/data/faqs-support

SHARE-ERIC (Ed.). (2023c). *SHARE Home*. https://share-eric.eu/

Sirven, N., Dumontet, M., & Rapp, T. (2020). The dynamics of frailty and change in socioeconomic conditions: Evidence for the 65+ in Europe. *European Journal of Public Health*, *30*(4), 715–719. https://doi.org/10.1093/eurpub/ckaa068

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

Sommet, N., & Morselli, D. (2017). Keep Calm and Learn Multilevel Logistic Modeling: A Simplified Three-Step Procedure Using Stata, R, Mplus, and SPSS. *International Review of Social Psychology*, *30*(1), 203–218. https://doi.org/10.5334/irsp.90

Stawarz, N., Ette, A., & Rüger, H. (2021). Healthy Migrants? Comparing Subjective Health of German Emigrants, Remigrants, and Non-Migrants. In M. Erlinghagen, A. Ette, N. F. Schneider, & N. Witte (Eds.), *IMISCOE Research Series. Global Lives of German Migrants* (pp. 205–225). Springer Nature. https://doi.org/10.1007/978-3-030-67498-4_12

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *BMJ*, *338*, Article b2393. https://doi.org/10.1136/bmj.b2393

Stoltzfus, J. C. (2011). Logistic regression: A brief primer. *Academic Emergency Medicine*, *18*(10), 1099–1104. https://doi.org/10.1111/j.1553-2712.2011.01185.x

Trappolini, E., & Giudici, C. (2021). Gendering health differences between nonmigrants and migrants by duration of stay in Italy. *Demographic Research*, *45*, Article 7, 221–258. https://doi.org/10.4054/DemRes.2021.45.7

United Nations (Ed.). (2023). *Human Development Index*. https://hdr.undp.org/data-center/human-development-index#/indicies/HDI

United Nations Development Programme (Ed.). (2016). *Human Development Report 2016: Human Development for Everyone*. New York. https://hdr.undp.org/system/files/documents/2016humandevelopmentreportpdf1pdf.pdf

United Nations Statistics Division (Ed.). (2023, August 22). *Methodology: Standard country or area codes for statistical use (M49)*. https://unstats.un.org/unsd/methodology/m49/

Vaillant, N. G., & Wolff, F.-C. (2010). Origin differences in self-reported health among older migrants living in France. *Public Health*, *124*(2), 90–98. https://doi.org/10.1016/j.puhe.2010.01.005

van der Linden, B. W. A., Cheval, B., Sieber, S., Orsholits, D., Guessous, I., Stringhini, S., Gabriel, R., Aartsen, M., Blane, D., Courvoisier, D., Burton-Jeangros, C., Kliegel, M., & Cullati, S. (2020). Life Course Socioeconomic Conditions and Frailty at Older Ages. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *75*(6), 1348–1357. https://doi.org/10.1093/geronb/gbz018

Verbrugge, L. M., & Jette, A. M. (1994). The disablement process. *Social Science & Medicine*, *38*(1), 1–14. https://doi.org/10.1016/0277-9536(94)90294-1

Viruell-Fuentes, E. A. (2007). Beyond acculturation: Immigration, discrimination, and health research among Mexicans in the United States. *Social Science & Medicine*, *65*(7), 1524–1535. https://doi.org/10.1016/j.socscimed.2007.05.010

Viruell-Fuentes, E. A., Miranda, P. Y., & Abdulrahim, S. (2012). More than culture: Structural racism, intersectionality theory, and immigrant health. *Social Science & Medicine*, *75*(12), 2099–2106. https://doi.org/10.1016/j.socscimed.2011.12.037

Vonneilich, N., Bremer, D., Knesebeck, O. von dem, & Lüdecke, D. (2021). Health Patterns among Migrant and Non-Migrant Middle- and Older-Aged Individuals in Europe: Analyses Based on Share 2004-2017. *International Journal of Environmental Research and Public Health*, *18*(22), Article 12047. https://doi.org/10.3390/ijerph182212047

Walkden, G. J., Anderson, E. L., Vink, M. P., Tilling, K., Howe, L. D., & Ben-Shlomo, Y. (2018). Frailty in older-age European migrants: Cross-sectional and longitudinal analyses of the Survey of Health, Aging and Retirement in Europe (SHARE). *Social Science & Medicine*, *213*, 1–11. https://doi.org/10.1016/j.socscimed.2018.07.033

Wallace, M., Khlat, M., & Guillot, M. (2019). Mortality advantage among migrants according to duration of stay in France, 2004-2014. *BMC Public Health*, *19*(1), 327–335. https://doi.org/10.1186/s12889-019-6652-1

Walton, E., Takeuchi, D. T., Herting, J. R., & Alegría, M. (2009). Does place of education matter? Contextualizing the education and health status association among Asian Americans. *Biodemography and Social Biology*, *55*(1), 30–51. https://doi.org/10.1080/19485560903054648

Wandschneider, L., Batram-Zantvoort, S., Razum, O., & Miani, C. (2020). Representation of gender in migrant health studies: A systematic review of the social epidemiological literature. *International Journal for Equity in Health*, *19*(1), Article 181. https://doi.org/10.1186/s12939-020-01289-y

White, P. (2007). Migrant Populations Approaching Old Age: Prospects in Europe. *Journal of Ethnic and Migration Studies*, *32*(8), 1283–1300. https://doi.org/10.1080/13691830600927708

Winship, C., & Radbill, L. (1994). Sampling Weights and Regression Analysis. *Sociological Methods & Research*, *23*(2), 230–257. https://doi.org/10.1177/0049124194023002004

Appendix

Contents

Table A1. Percentage of (non-)migrants across countries

	Non-migrant	Migrant	Missing migrant information	Total
Austria	90.34	8.30	1.36	100.00
	(2449)	(225)	(37)	(2711)
Germany	85.86	13.18	0.96	100.00
	(2683)	(412)	(30)	(3125)
Sweden	90.37	8.11	1.53	100.00
	(3077)	(276)	(52)	(3405)
Spain	94.10	3.14	2.76	100.00
	(4255)	(142)	(125)	(4522)
Italy	97.62	1.18	1.20	100.00
	(3813)	(46)	(47)	(3906)
France	89.28	9.51	1.21	100.00
	(2590)	(276)	(35)	(2901)
Denmark	96.11	3.24	0.65	100.00
	(2372)	(80)	(16)	(2468)
Greece	97.66	2.29	0.06	100.00
	(3502)	(82)	(2)	(3586)
Switzerland	83.08	15.70	1.23	100.00
	(1831)	(346)	(27)	(2204)
Belgium	91.74	7.62	0.65	100.00
	(3552)	(295)	(25)	(3872)
Czech Republic	94.72	4.34	0.94	100.00
	(3729)	(171)	(37)	(3937)
Poland	89.70	2.37	7.93	100.00
	(1211)	(32)	(107)	(1350)
Luxembourg	70.04	29.96	0.00	100.00
	(713)	(305)	(0)	(1018)
Portugal	97.97	1.79	0.23	100.00
	(1257)	(23)	(3)	(1283)
Slovenia	89.43	10.54	0.03	100.00
	(2901)	(342)	(1)	(3244)
Estonia	75.09	24.86	0.05	100.00
	(3114)	(1031)	(2)	(4147)
Croatia	81.98	18.02	0.00	100.00
	(1356)	(298)	(0)	(1654)
Total	90.01	8.88	1.11	100.00
	(44405)	(4382)	(546)	(49333)

Source: SHARE, wave 6, release 8.0.0, own calculations.
Respondents aged 60+, unweighted, imputed values included.
Number of observations in parentheses.

Table A2. Summary statistics by gender and migration status

	Non-migrant	Migrant	Male	Female	Total
Continuous variables: Mean (SD)					
Age	72.3	72.4	71.4	73.0	72.3
	(8.6)	(8.1)	(8.1)	(8.8)	(8.6)
Years since migration to interview country	.	50.6	50.6	50.5	50.6
	(.)	(18.6)	(18.6)	(18.6)	(18.6)
Household wealth, equivalized	172,551.5	166,123.9	187,525.0	158,974.7	171,634.4
	(282,177.3)	(270,871.3)	(331,005.5)	(231,064.8)	(280,176.5)
Total household income B, equivalized	22,150.0	27,417.5	23,512.3	21,722.7	22,516.2
	(35,255.6)	(44,520.3)	(36,739.6)	(35,527.0)	(36,080.3)
Categorical variables: Percent (N)					
Immigrant					
Non-migrant	100.0%	0.0%	90.6%	91.3%	91.0%
	(39,763)	(0)	(17,559)	(22,204)	(39,763)
Migrant	0.0%	100.0%	7.0%	7.2%	7.1%
	(0)	(3,101)	(1,358)	(1,743)	(3,101)
Missing	0.0%	0.0%	2.4%	1.5%	1.9%
	(0)	(0)	(459)	(376)	(835)
Gender					
Male	44.2%	43.8%	100.0%	0.0%	44.3%
	(17,559)	(1,358)	(19,377)	(0)	(19,377)
Female	55.8%	56.2%	0.0%	100.0%	55.7%
	(22,204)	(1,743)	(0)	(24,322)	(24,322)
Frailty type					
Robust	38.9%	37.9%	47.1%	32.5%	38.9%
	(15,461)	(1,175)	(9,125)	(7,895)	(17,020)
Pre-frail	48.2%	51.9%	44.1%	51.7%	48.3%
	(19,178)	(1,608)	(8,536)	(12,585)	(21,120)
Frail	12.9%	10.2%	8.9%	15.8%	12.7%
	(5,124)	(318)	(1,716)	(3,842)	(5,558)
Health problems in 0-3 dimensions					
0	16.0%	14.4%	18.5%	13.9%	15.9%
	(6,378)	(446)	(3,584)	(3,377)	(6,961)
1	58.3%	59.5%	61.6%	55.9%	58.4%
	(23,170)	(1,845)	(11,930)	(13,586)	(25,516)
2	13.8%	14.8%	11.1%	16.1%	13.9%
	(5,504)	(458)	(2,147)	(3,913)	(6,060)
3	11.6%	11.1%	8.6%	13.9%	11.6%
	(4,615)	(345)	(1,668)	(3,387)	(5,054)
Missing	0.2%	0.2%	0.2%	0.2%	0.2%

	(96)	(8)	(48)	(59)	(107)
Alcoholic drinks per week					
No drinks	50.0% (19,870)	48.9% (1,515)	34.6% (6,711)	62.3% (15,144)	50.0% (21,855)
1-2 drinks/week	14.1% (5,600)	16.9% (525)	12.5% (2,423)	15.6% (3,795)	14.2% (6,218)
3+ drinks/week	35.7% (14,179)	33.8% (1,047)	52.5% (10,172)	21.9% (5,323)	35.5% (15,496)
Missing	0.3% (113)	0.4% (13)	0.4% (71)	0.2% (59)	0.3% (130)
Smoking currently					
No	62.0% (24,642)	58.8% (1,822)	63.7% (12,342)	60.8% (14,787)	62.1% (27,130)
Yes	15.4% (6,118)	15.4% (476)	19.1% (3,703)	12.5% (3,038)	15.4% (6,740)
Missing	22.6% (9,003)	25.9% (802)	17.2% (3,331)	26.7% (6,497)	22.5% (9,828)
Employment status					
Retired	73.5% (29,240)	78.3% (2,429)	81.2% (15,727)	68.0% (16,551)	73.9% (32,278)
(Self-) Employed	9.9% (3,931)	9.8% (304)	12.3% (2,374)	8.0% (1,943)	9.9% (4,317)
Unemployed	1.6% (626)	2.3% (71)	2.3% (440)	1.1% (264)	1.6% (704)
Homemaker or perm. sick	12.0% (4,782)	7.5% (232)	2.6% (509)	19.0% (4,624)	11.7% (5,133)
Other	2.5% (994)	1.8% (55)	1.2% (231)	3.4% (834)	2.4% (1,065)
Missing	0.5% (190)	0.3% (9)	0.5% (96)	0.4% (107)	0.5% (203)
Formal education					
None/Primary	34.7% (13,784)	19.1% (591)	27.8% (5,392)	36.9% (8,983)	32.9% (14,375)
Secondary	45.5% (18,096)	49.5% (1,534)	44.6% (8,647)	45.2% (10,985)	44.9% (19,632)
Post-secondary	19.4% (7,694)	30.1% (934)	24.8% (4,798)	15.7% (3,831)	19.7% (8,629)
Other	0.4% (140)	1.3% (41)	0.3% (61)	0.5% (120)	0.4% (181)
Missing	0.1% (49)	0.0% (1)	2.5% (478)	1.7% (404)	2.0% (882)
Marital status					
Married or partnered	62.9% (25,017)	61.1% (1,896)	76.1% (14,736)	52.6% (12,799)	63.0% (27,535)
Widowed	24.4% (9,711)	23.9% (741)	11.2% (2,163)	34.8% (8,468)	24.3% (10,632)
Divorced	6.9%	9.5%	6.5%	7.4%	7.0%

	(2,744)	(294)	(1,256)	(1,798)	(3,054)
Never married	5.8%	5.5%	6.3%	5.2%	5.7%
	(2,291)	(169)	(1,221)	(1,257)	(2,477)
Attended a social club in past year					
No	72.7%	71.1%	71.2%	73.7%	72.6%
	(28,899)	(2,204)	(13,796)	(17,933)	(31,728)
Yes	22.4%	25.7%	24.3%	21.1%	22.5%
	(8,912)	(796)	(4,716)	(5,135)	(9,851)
Missing	4.9%	3.3%	4.5%	5.2%	4.9%
	(1,951)	(101)	(865)	(1,254)	(2,119)
Social network is empty					
No	89.3%	92.6%	89.4%	89.7%	89.6%
	(35,508)	(2,871)	(17,321)	(21,812)	(39,133)
Yes	2.5%	3.6%	2.8%	2.4%	2.6%
	(996)	(112)	(547)	(579)	(1,126)
Missing	8.2%	3.8%	7.8%	7.9%	7.9%
	(3,258)	(118)	(1,509)	(1,931)	(3,440)
Citizenship in country of interview					
No	0.1%	22.1%	1.9%	1.5%	1.7%
	(47)	(684)	(364)	(367)	(731)
Yes	99.8%	77.9%	95.7%	96.8%	96.4%
	(39,685)	(2,416)	(18,549)	(23,555)	(42,104)
Missing	0.1%	0.0%	2.4%	1.6%	2.0%
	(31)	(0)	(464)	(399)	(864)
HDI of origin country (immigrants)					
Wealthiest	0.0%	61.8%	4.3%	4.5%	4.4%
	(0)	(1,917)	(832)	(1,085)	(1,917)
Medium	0.0%	21.5%	1.4%	1.6%	1.5%
	(0)	(667)	(279)	(387)	(667)
Least wealthy	0.0%	9.4%	0.8%	0.6%	0.7%
	(0)	(290)	(156)	(134)	(290)
Missing	100.0%	7.3%	93.5%	93.4%	93.4%
	(39,763)	(227)	(18,110)	(22,715)	(40,824)
Geographical region					
North	4.4%	5.1%	4.6%	4.3%	4.4%
	(1,752)	(158)	(885)	(1,042)	(1,927)
East	13.4%	8.5%	12.8%	14.0%	13.4%
	(5,320)	(264)	(2,472)	(3,402)	(5,874)
South	37.0%	7.4%	35.1%	34.8%	34.9%
	(14,700)	(228)	(6,796)	(8,454)	(15,250)
West	45.2%	79.0%	47.6%	47.0%	47.2%
	(17,991)	(2,450)	(9,224)	(11,424)	(20,648)
MIPEX health care strand performance					

High inclusivity	25.6%	13.1%	24.8%	24.2%	24.5%
	(10,169)	(406)	(4,810)	(5,889)	(10,699)
Medium inclusivity	60.3%	77.9%	61.6%	61.1%	61.4%
	(23,969)	(2,414)	(11,941)	(14,872)	(26,813)
Low inclusivity	14.1%	9.0%	13.5%	14.6%	14.2%
	(5,625)	(281)	(2,625)	(3,561)	(6,186)

Source: SHARE, wave 6, release 8.0.0, own calculations.
Means (SD) and percent (N).
Respondents aged 60+, weighted, imputed values included.
Note: Differences in N between subgroups and total are due to weighting and rounding error and the omitted 'Missing migrant information' column.

Table A3. Fixed effects regression models with and without interaction of gender and migrant status

	Base model Pre-frail	Frail	Base model with interaction Pre-frail	Frail
Migrant status (ref. non-migrant)				
Migrant	1.13**	1.24**	1.09	1.28*
	(0.002)	(0.001)	(0.102)	(0.018)
Missing	0.90	1.03	0.79	1.05
	(0.277)	(0.839)	(0.092)	(0.849)
Female	1.51***	1.90***	1.50***	1.91***
	(0.000)	(0.000)	(0.000)	(0.000)
Female*migrant			1.05	0.96
			(0.494)	(0.760)
Female*missing migration info			1.30	1.03
			(0.195)	(0.936)
Age	1.04***	1.06***	1.04***	1.06***
	(0.000)	(0.000)	(0.000)	(0.000)
Health problems in 1-3 dimensions (ref. 0)				
1	2.23***	8.84***	2.23***	8.84***
	(0.000)	(0.000)	(0.000)	(0.000)
2	8.38***	108.89***	8.37***	108.86***
	(0.000)	(0.000)	(0.000)	(0.000)
3	32.04***	763.54***	32.04***	763.27***
	(0.000)	(0.000)	(0.000)	(0.000)
Missing	0.16***	2.34	0.16***	2.34
	(0.000)	(0.114)	(0.000)	(0.114)
Country (ref. Germany)				
Austria	1.06	1.79***	1.06	1.79***
	(0.337)	(0.000)	(0.344)	(0.000)
Sweden	1.03	1.00	1.03	1.00
	(0.588)	(0.996)	(0.592)	(0.995)
Spain	1.46***	1.98***	1.45***	1.98***
	(0.000)	(0.000)	(0.000)	(0.000)
Italy	2.28***	5.33***	2.27***	5.33***
	(0.000)	(0.000)	(0.000)	(0.000)
France	1.49***	2.37***	1.49***	2.37***
	(0.000)	(0.000)	(0.000)	(0.000)
Denmark	0.92	1.10	0.92	1.10
	(0.146)	(0.455)	(0.145)	(0.452)

Greece	2.77***	5.11***	2.77***	5.11***
	(0.000)	(0.000)	(0.000)	(0.000)
Switzerland	0.89	0.74*	0.89	0.74*
	(0.063)	(0.037)	(0.062)	(0.037)
Belgium	1.25***	2.06***	1.25***	2.06***
	(0.000)	(0.000)	(0.000)	(0.000)
Czech Republic	1.35***	1.81***	1.35***	1.81***
	(0.000)	(0.000)	(0.000)	(0.000)
Poland	2.41***	6.02***	2.42***	6.02***
	(0.000)	(0.000)	(0.000)	(0.000)
Luxembourg	1.13	1.49*	1.13	1.49*
	(0.138)	(0.013)	(0.139)	(0.013)
Portugal	1.62***	2.92***	1.62***	2.92***
	(0.000)	(0.000)	(0.000)	(0.000)
Slovenia	1.14*	1.32*	1.14*	1.32*
	(0.017)	(0.014)	(0.017)	(0.014)
Estonia	1.93***	2.40***	1.93***	2.40***
	(0.000)	(0.000)	(0.000)	(0.000)
Croatia	1.80***	4.71***	1.80***	4.71***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	49333		49333	

Exponentiated coefficients; *p*-values in parentheses
Source: SHARE, wave 6, release 8.0.0, own calculations.
Respondents aged 60+, unweighted, imputed values included.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A4. Fixed effects regression models including different groups of mediators (men)

	No mediators	Health behaviours	Economic factors	Social capital	All mediators
Pre-frail					
Migrant status (ref. non-migrant)					
Migrant	1.09 (0.121)	1.08 (0.189)	1.04 (0.482)	1.07 (0.204)	1.03 (0.639)
Missing	0.81 (0.131)	0.81 (0.125)	0.69 (0.485)	0.83 (0.177)	0.75 (0.599)
Age	1.04*** (0.000)	1.04*** (0.000)	1.04*** (0.000)	1.03*** (0.000)	1.04*** (0.000)
Health problems in 1-3 dimensions (ref. 0)					
1	2.15*** (0.000)	2.17*** (0.000)	2.13*** (0.000)	2.16*** (0.000)	2.16*** (0.000)
2	8.14*** (0.000)	8.05*** (0.000)	7.64*** (0.000)	8.07*** (0.000)	7.66*** (0.000)
3	39.71*** (0.000)	38.37*** (0.000)	35.87*** (0.000)	37.78*** (0.000)	34.49*** (0.000)
Missing	0.19** (0.002)	0.35 (0.096)	0.24* (0.017)	0.19** (0.002)	0.43 (0.212)
Country (ref. Germany)					
Austria	1.01 (0.876)	1.00 (0.993)	1.01 (0.887)	1.02 (0.855)	1.01 (0.886)
Sweden	1.01 (0.889)	1.02 (0.770)	1.09 (0.275)	1.07 (0.431)	1.15 (0.102)
Spain	1.38*** (0.000)	1.31*** (0.000)	1.23** (0.010)	1.24** (0.005)	1.12 (0.180)
Italy	2.10*** (0.000)	2.02*** (0.000)	1.94*** (0.000)	1.90*** (0.000)	1.77*** (0.000)
France	1.30** (0.002)	1.33** (0.001)	1.37*** (0.000)	1.28** (0.004)	1.37*** (0.000)
Denmark	1.00 (0.965)	1.02 (0.787)	1.09 (0.350)	1.05 (0.585)	1.14 (0.147)
Greece	2.77*** (0.000)	2.58*** (0.000)	2.39*** (0.000)	2.50*** (0.000)	2.14*** (0.000)
Switzerland	0.81* (0.020)	0.80* (0.018)	1.01 (0.893)	0.81* (0.022)	0.97 (0.769)
Belgium	1.26** (0.003)	1.26** (0.003)	1.40*** (0.000)	1.21* (0.017)	1.33*** (0.001)

Czech Republic	1.26^{**} (0.004)	1.22^{*} (0.012)	1.04 (0.672)	1.19^{*} (0.032)	1.00 (0.969)
Poland	2.20^{***} (0.000)	1.97^{***} (0.000)	1.57^{***} (0.000)	1.92^{***} (0.000)	1.37^{**} (0.007)
Luxembourg	1.00 (0.977)	1.02 (0.833)	1.33^{*} (0.020)	0.97 (0.772)	1.27 (0.053)
Portugal	1.60^{***} (0.000)	1.64^{***} (0.000)	1.32^{*} (0.014)	1.48^{***} (0.000)	1.33^{*} (0.014)
Slovenia	1.06 (0.458)	1.04 (0.622)	0.97 (0.721)	1.01 (0.887)	0.93 (0.422)
Estonia	2.08^{***} (0.000)	1.88^{***} (0.000)	1.72^{***} (0.000)	1.89^{***} (0.000)	1.52^{***} (0.000)
Croatia	1.71^{***} (0.000)	1.63^{***} (0.000)	1.39^{**} (0.002)	1.59^{***} (0.000)	1.32^{**} (0.009)
Smoking (ref. non-smoker)					
Yes		1.46^{***} (0.000)			1.36^{***} (0.000)
Missing		0.97 (0.444)			0.97 (0.520)
Alcoholic drinks (ref. none)					
1-2 drinks/week		0.79^{***} (0.000)			0.83^{***} (0.000)
3+ drinks/week		0.77^{***} (0.000)			0.84^{***} (0.000)
Missing		0.43^{*} (0.030)			0.45^{*} (0.042)
Education (ref. primary)					
Secondary			0.95 (0.271)		0.97 (0.545)
Post-secondary			0.89^{*} (0.017)		0.94 (0.198)
Other			0.89 (0.684)		0.91 (0.739)
Missing			1.14 (0.804)		1.08 (0.891)
Employment status (ref. retired)					

(Self-)Employed			1.16**		1.14**
			(0.004)		(0.008)
Unemployed			1.21		1.12
			(0.098)		(0.332)
Homemaker or perm. sick			2.29***		2.16***
			(0.000)		(0.000)
Other			1.65**		1.59**
			(0.001)		(0.003)
Missing			0.78		0.83
			(0.429)		(0.551)
Logged total household income, equivalized			0.92**		0.93**
			(0.001)		(0.005)
Cube-root transformed household wealth, equivalized			0.99***		0.99***
			(0.000)		(0.000)
Marital status (ref. married)					
Widowed				1.21**	1.18**
				(0.002)	(0.006)
Divorced				1.15*	1.04
				(0.031)	(0.528)
Never married				1.23**	1.13
				(0.005)	(0.103)
Attended social clubs in past year (ref. no)					
Yes				0.69***	0.75***
				(0.000)	(0.000)
Missing				0.89	0.85
				(0.223)	(0.118)
No social network (ref. has a social network)					
Yes				1.19	1.14
				(0.063)	(0.160)
Missing				1.08	1.10
				(0.274)	(0.208)
Frail					
Migrant status (ref. non-migrant)					
Migrant	1.24*	1.19	1.15	1.23	1.10
	(0.041)	(0.098)	(0.201)	(0.056)	(0.392)

Missing	1.11 (0.669)	1.12 (0.644)	21.38^{**} (0.004)	1.19 (0.481)	52.64^{**} (0.005)
Age	1.06^{***} (0.000)	1.06^{***} (0.000)	1.06^{***} (0.000)	1.06^{***} (0.000)	1.06^{***} (0.000)
Health problems in 1-3 dimensions (ref. 0)					
1	7.94^{***} (0.000)	8.12^{***} (0.000)	7.58^{***} (0.000)	7.98^{***} (0.000)	7.80^{***} (0.000)
2	116.32^{***} (0.000)	113.45^{***} (0.000)	98.84^{***} (0.000)	119.42^{***} (0.000)	103.34^{***} (0.000)
3	942.79^{***} (0.000)	867.12^{***} (0.000)	748.87^{***} (0.000)	$1,322.02^{***}$ (0.000)	$1,056.38^{***}$ (0.000)
Missing	0.00 (0.977)	0.00 (0.979)	0.00 (0.979)	0.00 (0.988)	0.00 (0.992)
Country (ref. Germany)					
Austria	1.58^{**} (0.009)	1.56^{*} (0.013)	1.49^{*} (0.025)	1.93^{***} (0.000)	1.87^{***} (0.001)
Sweden	1.00 (0.979)	1.01 (0.948)	1.16 (0.421)	1.18 (0.388)	1.28 (0.203)
Spain	1.81^{***} (0.000)	1.55^{**} (0.006)	1.50^{*} (0.016)	2.07^{***} (0.000)	1.58^{**} (0.009)
Italy	4.16^{***} (0.000)	3.77^{***} (0.000)	3.67^{***} (0.000)	4.43^{***} (0.000)	3.78^{***} (0.000)
France	1.96^{***} (0.000)	2.05^{***} (0.000)	2.10^{***} (0.000)	1.99^{***} (0.000)	2.29^{***} (0.000)
Denmark	1.25 (0.237)	1.30 (0.172)	1.49^{*} (0.041)	1.47 (0.051)	1.79^{**} (0.004)
Greece	4.79^{***} (0.000)	4.05^{***} (0.000)	3.65^{***} (0.000)	4.47^{***} (0.000)	3.17^{***} (0.000)
Switzerland	0.72 (0.142)	0.72 (0.144)	1.12 (0.621)	0.73 (0.172)	1.09 (0.712)
Belgium	1.78^{***} (0.000)	1.80^{***} (0.000)	2.16^{***} (0.000)	1.72^{**} (0.001)	2.12^{***} (0.000)
Czech Republic	1.70^{**} (0.001)	1.56^{**} (0.007)	1.22 (0.260)	1.66^{**} (0.003)	1.19 (0.344)
Poland	4.51^{***} (0.000)	3.52^{***} (0.000)	2.57^{***} (0.000)	3.80^{***} (0.000)	1.94^{**} (0.002)
Luxembourg	1.15 (0.597)	1.21 (0.461)	1.95^{*} (0.013)	1.12 (0.676)	1.86^{*} (0.027)

Portugal	2.49*** (0.000)	2.57*** (0.000)	1.76** (0.009)	2.63*** (0.000)	2.09** (0.001)
Slovenia	1.21 (0.286)	1.17 (0.363)	1.04 (0.831)	1.27 (0.193)	1.12 (0.562)
Estonia	2.44*** (0.000)	1.97*** (0.000)	1.90*** (0.000)	2.26*** (0.000)	1.57* (0.014)
Croatia	4.73*** (0.000)	4.14*** (0.000)	3.45*** (0.000)	4.29*** (0.000)	2.91*** (0.000)
Smoking (ref. non-smoker)					
Yes		2.19*** (0.000)			2.02*** (0.000)
Missing		0.98 (0.825)			1.07 (0.423)
Alcoholic drinks (ref. none)					
1-2 drinks/week		0.52*** (0.000)			0.53*** (0.000)
3+ drinks/week		0.55*** (0.000)			0.55*** (0.000)
Missing		0.28 (0.126)			0.48 (0.397)
Education (ref. primary)					
Secondary			0.81** (0.008)		0.85* (0.039)
Post-secondary			0.70*** (0.000)		0.78* (0.011)
Other			0.85 (0.751)		0.98 (0.965)
Missing			0.04** (0.004)		0.02** (0.006)
Employment status (ref. retired)					
(Self-)Employed			0.62** (0.003)		0.66* (0.011)
Unemployed			0.72 (0.275)		0.67 (0.183)
Homemaker or perm. sick			2.63*** (0.000)		2.69*** (0.000)

Other			1.36		1.38
			(0.335)		(0.324)
Missing			0.26		2.57
			(0.095)		(0.338)
Logged total household income, equivalized			0.91		0.91
			(0.053)		(0.058)
Cube-root transformed household wealth, equivalized			0.98***		0.99***
			(0.000)		(0.000)
Marital status (ref. married)					
Widowed				1.39**	1.30*
				(0.001)	(0.010)
Divorced				1.34*	1.04
				(0.023)	(0.782)
Never married				1.44**	1.17
				(0.009)	(0.275)
Attended social clubs in past year (ref. no)					
Yes				0.33***	0.39***
				(0.000)	(0.000)
Missing				0.05***	0.04***
				(0.000)	(0.000)
No social network (ref. has a social network)					
Yes				1.60**	1.42*
				(0.004)	(0.033)
Missing				0.88	0.86
				(0.359)	(0.308)
Observations	22322	22322	22322	22322	22322

Exponentiated coefficients; *p*-values in parentheses
Source: SHARE, wave 6, release 8.0.0, own calculations.
Respondents aged 60+, unweighted, imputed values included.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A5. Fixed effects regression models including different groups of mediators (women)

	No mediators	Health behaviours	Economic factors	Social capital	All mediators
Pre-frail					
Migrant status (ref. non-migrant)					
Migrant	1.16**	1.15**	1.12*	1.13*	1.10
	(0.005)	(0.009)	(0.034)	(0.017)	(0.069)
Missing	1.01	1.00	0.56	1.03	0.50
	(0.936)	(0.987)	(0.423)	(0.832)	(0.344)
Age	1.04***	1.04***	1.04***	1.04***	1.04***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Health problems in 1-3 dimensions (ref. 0)					
1	2.31***	2.31***	2.28***	2.34***	2.31***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
2	8.55***	8.46***	8.18***	8.62***	8.26***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
3	26.74***	25.93***	24.74***	26.77***	24.48***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing	0.14***	0.61	0.30*	0.15***	0.83
	(0.000)	(0.457)	(0.036)	(0.000)	(0.778)
Country (ref. Germany)					
Austria	1.10	1.06	1.09	1.06	1.03
	(0.233)	(0.467)	(0.290)	(0.475)	(0.719)
Sweden	1.05	1.05	1.16	1.09	1.16
	(0.522)	(0.534)	(0.062)	(0.261)	(0.062)
Spain	1.53***	1.46***	1.36***	1.40***	1.24**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.008)
Italy	2.49***	2.36***	2.31***	2.25***	2.05***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
France	1.68***	1.68***	1.78***	1.64***	1.71***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Denmark	0.85	0.88	0.93	0.90	0.96
	(0.054)	(0.125)	(0.366)	(0.231)	(0.667)
Greece	2.75***	2.59***	2.29***	2.43***	2.05***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Switzerland	0.97	0.95	1.20*	0.93	1.09
	(0.722)	(0.589)	(0.048)	(0.396)	(0.372)
Belgium	1.24**	1.27**	1.38***	1.19*	1.31***
	(0.004)	(0.002)	(0.000)	(0.021)	(0.001)

Czech Republic	1.43*** (0.000)	1.35*** (0.000)	1.20* (0.024)	1.28** (0.001)	1.09 (0.289)
Poland	2.66*** (0.000)	2.36*** (0.000)	1.99*** (0.000)	2.35*** (0.000)	1.76*** (0.000)
Luxembourg	1.27* (0.035)	1.32* (0.016)	1.61*** (0.000)	1.22 (0.082)	1.53*** (0.000)
Portugal	1.65*** (0.000)	1.69*** (0.000)	1.40** (0.004)	1.52*** (0.000)	1.39** (0.006)
Slovenia	1.22* (0.010)	1.19* (0.026)	1.13 (0.116)	1.19* (0.033)	1.10 (0.238)
Estonia	1.84*** (0.000)	1.69*** (0.000)	1.57*** (0.000)	1.68*** (0.000)	1.43*** (0.000)
Croatia	1.89*** (0.000)	1.81*** (0.000)	1.53*** (0.000)	1.68*** (0.000)	1.43*** (0.001)
Smoking (ref. non-smoker)					
Yes		1.31*** (0.000)			1.23*** (0.000)
Missing		0.92* (0.029)			0.92* (0.029)
Alcoholic drinks (ref. none)					
1-2 drinks/week		0.85*** (0.000)			0.91* (0.015)
3+ drinks/week		0.79*** (0.000)			0.88*** (0.000)
Missing		0.14*** (0.001)			0.20** (0.004)
Education (ref. primary)					
Secondary			0.92* (0.037)		0.93 (0.081)
Post-secondary			0.85** (0.001)		0.88* (0.011)
Other			0.82 (0.415)		0.83 (0.440)
Missing			1.68 (0.464)		1.92 (0.371)
Employment status (ref. retired)					

(Self-)Employed			1.19** (0.001)		1.15** (0.007)
Unemployed			1.42* (0.013)		1.34* (0.039)
Homemaker or perm. sick			1.11* (0.026)		1.11* (0.027)
Other			1.40** (0.003)		1.39** (0.004)
Missing			0.34** (0.001)		0.44* (0.022)
Logged total household income, equivalized			0.92*** (0.001)		0.93** (0.003)
Cube-root transformed household wealth, equivalized			0.99*** (0.000)		0.99*** (0.000)
Marital status (ref. married)					
Widowed				1.15*** (0.000)	1.07 (0.092)
Divorced				1.21*** (0.000)	1.11* (0.049)
Never married				1.21** (0.007)	1.16* (0.047)
Attended social clubs in past year (ref. no)					
Yes				0.73*** (0.000)	0.79*** (0.000)
Missing				0.84 (0.135)	0.89 (0.344)
No social network (ref. has a social network)					
Yes				1.52*** (0.000)	1.50*** (0.000)
Missing				0.95 (0.494)	0.99 (0.863)
Frail					
Migrant status (ref. non-migrant)					
Migrant	1.25** (0.007)	1.22* (0.015)	1.17 (0.057)	1.22* (0.019)	1.15 (0.109)
Missing	1.05 (0.839)	1.00 (0.996)	0.38 (0.267)	1.21 (0.415)	0.48 (0.442)

Age	1.06***	1.06***	1.06***	1.06***	1.06***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Health problems in 1-3 dimensions (ref. 0)					
1	9.52***	9.44***	9.03***	9.56***	9.13***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
2	105.41***	101.28***	92.78***	105.55***	93.43***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
3	634.18***	579.02***	526.07***	857.41***	697.11***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Missing	3.71*	35.53***	10.69***	44.51***	114.78***
	(0.021)	(0.000)	(0.000)	(0.000)	(0.000)
Country (ref. Germany)					
Austria	1.97***	1.77***	1.80***	1.95***	1.75***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Sweden	1.01	0.99	1.23	1.09	1.24
	(0.962)	(0.949)	(0.194)	(0.586)	(0.199)
Spain	2.16***	1.82***	1.72***	2.36***	1.70***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Italy	6.51***	5.58***	5.41***	6.75***	5.27***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
France	2.79***	2.76***	3.02***	2.65***	2.94***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Denmark	1.00	1.08	1.15	1.13	1.34
	(0.982)	(0.652)	(0.427)	(0.477)	(0.106)
Greece	5.39***	4.57***	3.69***	4.47***	2.98***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Switzerland	0.78	0.77	1.16	0.65*	0.92
	(0.187)	(0.175)	(0.458)	(0.028)	(0.690)
Belgium	2.26***	2.35***	2.75***	2.10***	2.58***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Czech Republic	1.94***	1.64***	1.30	1.58**	1.12
	(0.000)	(0.000)	(0.070)	(0.002)	(0.473)
Poland	7.47***	5.42***	4.17***	6.18***	3.28***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Luxembourg	1.83**	2.10***	2.84***	1.66*	2.54***
	(0.003)	(0.000)	(0.000)	(0.019)	(0.000)
Portugal	3.27***	3.53***	2.28***	3.19***	2.61***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Slovenia	1.43* (0.015)	1.35* (0.042)	1.17 (0.290)	1.41* (0.028)	1.21 (0.242)
Estonia	2.40*** (0.000)	1.84*** (0.000)	1.80*** (0.000)	1.91*** (0.000)	1.36* (0.043)
Croatia	4.85*** (0.000)	4.45*** (0.000)	3.06*** (0.000)	3.91*** (0.000)	2.68*** (0.000)
Smoking (ref. non-smoker)					
Yes		1.57*** (0.000)			1.41*** (0.000)
Missing		0.78*** (0.000)			0.87* (0.026)
Alcoholic drinks (ref. none)					
1-2 drinks/week		0.53*** (0.000)			0.57*** (0.000)
3+ drinks/week		0.46*** (0.000)			0.53*** (0.000)
Missing		0.03*** (0.000)			0.07** (0.009)
Education (ref. primary)					
Secondary			0.87* (0.018)		0.88* (0.042)
Post-secondary			0.65*** (0.000)		0.71*** (0.000)
Other			1.19 (0.601)		1.19 (0.597)
Missing			2.48 (0.280)		2.23 (0.385)
Employment status (ref. retired)					
(Self-)Employed			0.82 (0.145)		0.89 (0.413)
Unemployed			1.25 (0.416)		1.17 (0.571)
Homemaker or perm. sick			1.23** (0.002)		1.31*** (0.000)
Other			1.53** (0.004)		1.53** (0.005)

Missing			0.21**		3.86*
			(0.002)		(0.023)
Logged total household income, equivalized			0.84***		0.89**
			(0.000)		(0.002)
Cube-root transformed household wealth, equivalized			0.99***		0.99***
			(0.000)		(0.000)
Marital status (ref. married)					
Widowed				1.38***	1.20**
				(0.000)	(0.002)
Divorced				1.54***	1.32**
				(0.000)	(0.002)
Never married				1.16	1.09
				(0.229)	(0.510)
Attended social clubs in past year (ref. no)					
Yes				0.34***	0.42***
				(0.000)	(0.000)
Missing				0.04***	0.03***
				(0.000)	(0.000)
No social network (ref. has a social network)					
Yes				1.76***	1.70***
				(0.000)	(0.000)
Missing				0.61***	0.63***
				(0.000)	(0.000)
Observations	27011	27011	27011	27011	27011

Exponentiated coefficients; *p*-values in parentheses
Source: SHARE, wave 6, release 8.0.0, own calculations.
Respondents aged 60+, unweighted, imputed values included.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A6. Adjusted Odds Ratios for (pre-)frailty in migrants (KHB Method)

		Health behaviours	Economic factors	Social capital	All mediators
Pre-frailty	**Migrant men**				
	Total effect	1.10	1.09	1.09	1.09
		(0.107)	(0.141)	(0.123)	(0.124)
	Direct effect	1.08	1.04	1.07	1.03
		(0.189)	(0.482)	(0.204)	(0.639)
	Indirect effect	1.02	1.04**	1.02	1.06**
		(0.172)	(0.007)	(0.189)	(0.004)
	Percentage mediated	18.49	51.96	17.60	69.31
	Migrant women				
	Total effect	1.16**	1.16**	1.16**	1.15**
		(0.005)	(0.005)	(0.005)	(0.006)
	Direct effect	1.15**	1.12*	1.13*	1.10
		(0.009)	(0.034)	(0.017)	(0.069)
	Indirect effect	1.01	1.04*	1.02	1.05**
		(0.408)	(0.017)	(0.059)	(0.009)
	Percentage mediated	5.69	23.71	14.56	33.51
Frailty	**Migrant men**				
	Total effect	1.24*	1.26*	1.27*	1.26*
		(0.043)	(0.033)	(0.031)	(0.036)
	Direct effect	1.19	1.15	1.23	1.10
		(0.098)	(0.201)	(0.056)	(0.392)
	Indirect effect	1.04	1.09**	1.03	1.15*
		(0.139)	(0.008)	(0.616)	(0.036)
	Percentage mediated	18.21	39.69	10.94	59.01
	Migrant women				
	Total effect	1.26**	1.26**	1.29**	1.30**
		(0.005)	(0.006)	(0.003)	(0.002)
	Direct effect	1.22*	1.17	1.22*	1.15
		(0.015)	(0.057)	(0.019)	(0.109)
	Indirect effect	1.03	1.07*	1.06	1.13
		(0.270)	(0.021)	(0.316)	(0.056)
	Percentage mediated	12.30	30.09	21.44	46.95

Exponentiated coefficients; *p*-values in parentheses
Source: SHARE, wave 6, release 8.0.0, own calculations.
All models controlling for missing migrant information, age, health problems, and countries in addition to the mediators as specified in the model titles (coefficients not shown).
Respondents aged 60+, unweighted, imputed values included.
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table A7. Overview of explained percentage of migrant effect, by mediators (KHB Method)

	Percentage Mediated			
	Men pre-frail	Men frail	Women pre-frail	Women frail
Smoking	0.95	0.89	2.25	1.92
Smoking info missing	-0.66	-0.63	0.11	-0.10
1-2 drinks/week	1.64	5.12	1.97	2.52
3+ drinks/week	1.86	5.00	10.06	12.49
Missing drink info	-0.82	-0.73	-0.59	-0.20
Secondary education	3.32	3.23	2.80	6.40
Post-secondary education	-4.65	-6.94	-5.69	-8.22
Other education	-0.89	0.46	-0.72	-0.06
Missing education info	0.33	0.22	-0.06	1.24
(Self-)employed	-1.05	0.46	-0.64	0.75
Unemployed	1.10	0.33	1.49	-2.00
Homemaker or sick	0.01	0.01	4.72	2.25
Other employment status	1.39	0.99	0.75	0.20
Missing employment info	-0.07	0.06	-0.42	0.78
Income	2.42	2.17	1.93	0.92
Wealth	17.61	20.35	37.43	32.68
Widowed	0.23	0.35	-0.92	-0.53
Divorced	1.19	1.76	0.30	0.10
Never married	-0.85	-0.27	-3.23	-1.53
Attended social club	7.88	16.40	17.94	21.40
Missing social club info	-0.24	-4.01	-1.71	-12.80
Network empty	2.82	2.04	1.08	1.06

Missing network info	-0.02	-0.31	0.45	-0.27

Source: SHARE, wave 6, release 8.0.0, own calculations.
Based on models controlling for missing migrant information, age, health problems, countries, and all mediators (coefficients not shown).
Respondents aged 60+, unweighted, imputed values included.

Table A8. Migrant-only fixed effects regression models including migration-specific variables

	Men		Women	
	Pre-frail	Frail	Pre-frail	Frail
Age	1.02*	1.04*	1.02*	1.05***
	(0.017)	(0.019)	(0.024)	(0.000)
Health problems in 1-3 dimensions (ref. 0)				
1	2.37***	16.08**	2.08***	9.52***
	(0.000)	(0.007)	(0.000)	(0.000)
2	6.66***	219.20***	9.48***	117.34***
	(0.000)	(0.000)	(0.000)	(0.000)
3	52.83***	5,300.80***	146.34***	4,037.53***
	(0.000)	(0.000)	(0.000)	(0.000)
Missing	0.00	3.40e+06	6.86e+06	0.31
	(0.991)	(0.996)	(0.992)	(1.000)
Country (ref. Germany)				
Austria	1.13	3.56	0.89	2.32
	(0.695)	(0.069)	(0.665)	(0.073)
Sweden	1.30	6.18**	0.99	0.89
	(0.343)	(0.004)	(0.980)	(0.835)
Spain	1.14	2.56	0.91	0.75
	(0.706)	(0.335)	(0.775)	(0.718)
Italy	0.61	7.96*	1.97	6.30*
	(0.425)	(0.049)	(0.161)	(0.021)
France	1.37	8.86***	1.60	4.31**
	(0.293)	(0.001)	(0.095)	(0.002)
Denmark	3.68**	35.20***	0.82	2.34
	(0.007)	(0.000)	(0.606)	(0.248)
Greece	1.05	2.41	1.34	1.29
	(0.909)	(0.378)	(0.486)	(0.719)
Switzerland	1.42	3.81	0.86	0.98
	(0.202)	(0.057)	(0.550)	(0.966)
Belgium	1.82*	5.34*	1.11	2.25
	(0.030)	(0.012)	(0.701)	(0.098)
Czech Republic	0.89	2.21	0.78	1.75
	(0.741)	(0.261)	(0.444)	(0.287)
Poland	1.62	1.70	1.02	1.23
	(0.451)	(0.650)	(0.979)	(0.853)
Luxembourg	0.76	3.31	0.96	2.33
	(0.391)	(0.095)	(0.871)	(0.111)
Portugal	1.39	49.75**	1.16	0.41

	(0.672)	(0.007)	(0.847)	(0.523)
Slovenia	0.69 (0.162)	2.77 (0.091)	0.99 (0.980)	1.42 (0.498)
Estonia	1.50 (0.106)	2.86 (0.069)	1.37 (0.189)	2.60* (0.024)
Croatia	0.60 (0.116)	2.16 (0.276)	0.75 (0.329)	1.64 (0.352)
Years since migration to interview country	1.00 (0.463)	1.00 (0.926)	1.00 (0.286)	1.01 (0.357)
HDI level of origin country (ref. highest)				
Medium	1.41* (0.030)	1.91* (0.032)	1.42* (0.021)	1.42 (0.145)
Least wealthy	1.64* (0.036)	0.76 (0.593)	1.19 (0.442)	1.53 (0.290)
Missing	0.85 (0.631)	0.17 (0.128)	1.05 (0.899)	1.66 (0.338)
Citizenship in country of interview (ref. no)				
Yes	1.01 (0.935)	0.90 (0.724)	0.92 (0.523)	0.79 (0.231)
Missing	0.00 (0.993)	0.00 (0.998)	957071.14 (0.997)	0.10 (1.000)
Smoking (ref. non-smoker)				
Yes	1.31 (0.063)	2.67*** (0.000)	1.23 (0.176)	2.11** (0.003)
Missing	0.92 (0.575)	1.22 (0.561)	0.91 (0.490)	0.87 (0.546)
Alcoholic drinks (ref. none)				
1-2 drinks/week	0.65* (0.013)	0.31** (0.005)	1.01 (0.956)	0.51* (0.013)
3+ drinks/week	0.98 (0.859)	0.80 (0.387)	0.91 (0.458)	0.51** (0.009)
Missing	0.30 (0.325)	0.00 (0.990)	0.50 (0.639)	0.00 (0.996)
Education (ref. primary)				
Secondary	0.97	0.96	1.02	0.77

	(0.890)	(0.901)	(0.878)	(0.253)
Post-secondary	1.11 (0.598)	1.20 (0.614)	0.86 (0.375)	0.68 (0.141)
Other	2.00 (0.280)	1.58 (0.700)	0.54 (0.282)	2.61 (0.220)
Employment status (ref. retired)				
(Self-)Employed	1.32 (0.121)	0.32 (0.140)	1.04 (0.823)	0.60 (0.250)
Unemployed	1.19 (0.604)	0.35 (0.348)	2.25 (0.106)	2.75 (0.192)
Homemaker or perm. sick	2.95* (0.011)	3.76* (0.022)	0.93 (0.667)	1.28 (0.363)
Other	0.38 (0.169)	1.88 (0.512)	0.98 (0.945)	1.64 (0.315)
Missing	2.96 (0.277)	0.00 (0.985)	0.00 (0.988)	261.98 (0.571)
Logged total household income, equivalized	0.78** (0.006)	0.75 (0.141)	0.99 (0.950)	1.02 (0.881)
Cube-root transformed household wealth, equivalized	0.99** (0.002)	0.98** (0.003)	0.99*** (0.000)	0.98*** (0.000)
Marital status (ref. married)				
Widowed	1.48 (0.078)	2.33* (0.021)	0.90 (0.441)	0.91 (0.619)
Divorced	0.97 (0.897)	1.56 (0.259)	1.12 (0.497)	1.29 (0.344)
Never married	1.27 (0.459)	1.11 (0.887)	0.85 (0.527)	1.18 (0.680)
Attended social clubs in past year (ref. no)				
Yes	0.75* (0.036)	0.28*** (0.000)	0.79 (0.054)	0.41*** (0.001)
Missing	0.53* (0.047)	0.01*** (0.000)	0.47 (0.087)	0.02*** (0.000)
No social network (ref. has a social network)				
Yes	1.77 (0.081)	2.84* (0.042)	1.03 (0.923)	2.11 (0.072)

Missing	1.37	0.92	0.79	0.49
	(0.220)	(0.868)	(0.445)	(0.153)
Missing			0.45	0.00
			(0.595)	(0.993)
Observations	1882		2464	

Exponentiated coefficients; *p*-values in parentheses
Source: SHARE, wave 6, release 8.0.0, own calculations
Respondents aged 60+, unweighted, imputed values included.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A9. Regression models with country group interactions (no fixed effects)

	Men		Women	
	Pre-frail	Frail	Pre-frail	Frail
Migrant status (ref. non-migrant)				
Migrant	0.86	0.85	1.00	0.71
	(0.285)	(0.582)	(0.972)	(0.115)
Missing	1.09	45.33*	0.53	0.52
	(0.893)	(0.040)	(0.412)	(0.526)
Region (ref. West)				
North	0.97	0.72**	0.87**	0.56***
	(0.590)	(0.003)	(0.004)	(0.000)
East	0.78***	0.71**	0.86**	0.74***
	(0.000)	(0.002)	(0.004)	(0.001)
South	1.24***	1.29**	1.25***	1.32***
	(0.000)	(0.008)	(0.000)	(0.000)
North*migrant	1.46*	2.29*	1.06	1.38
	(0.044)	(0.024)	(0.709)	(0.272)
East*migrant	1.01	1.86	0.77	0.76
	(0.944)	(0.093)	(0.147)	(0.359)
South*migrant	0.94	1.10	0.84	0.43*
	(0.766)	(0.840)	(0.380)	(0.030)
Missing # North	0.52	0.80	1.27	0.64
	(0.161)	(0.855)	(0.650)	(0.713)
Missing # East	0.63	1.66	1.26	0.36
	(0.444)	(0.616)	(0.734)	(0.346)
Missing # South	0.62	0.71	0.68	0.63
	(0.198)	(0.659)	(0.313)	(0.437)
Health care inclusivity (ref. high inclusivity)				
Medium inclusivity	0.95	0.78**	0.94	0.74***
	(0.257)	(0.003)	(0.097)	(0.000)
Low inclusivity	1.28***	1.10	1.20***	0.98
	(0.000)	(0.360)	(0.000)	(0.777)
Med.inclusive*migrant	1.26	1.08	1.20	1.77*
	(0.133)	(0.821)	(0.176)	(0.020)
Low inclusive*migrant	0.90	0.60	1.23	1.67
	(0.589)	(0.165)	(0.219)	(0.086)
Missing # Medium inclusivity	0.85	1.56	1.00	1.00
	(0.631)	(0.561)	(1.000)	(0.997)
Missing # Low	1.60	1.15	0.84	4.00

inclusivity	(0.473)	(0.903)	(0.827)	(0.259)
Age	1.04***	1.06***	1.04***	1.06***
	(0.000)	(0.000)	(0.000)	(0.000)
Health problems in 1-3 dimensions (ref. 0)				
1	2.14***	7.56***	2.30***	8.95***
	(0.000)	(0.000)	(0.000)	(0.000)
2	7.61***	98.88***	8.22***	90.78***
	(0.000)	(0.000)	(0.000)	(0.000)
3	34.26***	1,008.70***	24.52***	689.75***
	(0.000)	(0.000)	(0.000)	(0.000)
Missing	0.44	0.00	0.85	102.79***
	(0.225)	(0.992)	(0.817)	(0.000)
Smoking (ref. non-smoker)				
Yes	1.37***	2.03***	1.22***	1.41***
	(0.000)	(0.000)	(0.000)	(0.000)
Missing	0.97	1.09	0.89**	0.84**
	(0.393)	(0.319)	(0.001)	(0.003)
Alcoholic drinks (ref. none)				
1-2 drinks/week	0.84***	0.53***	0.91*	0.57***
	(0.000)	(0.000)	(0.016)	(0.000)
3+ drinks/week	0.85***	0.57***	0.89***	0.56***
	(0.000)	(0.000)	(0.001)	(0.000)
Missing	0.46*	0.43	0.21**	0.08*
	(0.049)	(0.332)	(0.005)	(0.011)
Education (ref. primary)				
Secondary	0.97	0.81**	0.88**	0.77***
	(0.476)	(0.007)	(0.001)	(0.000)
Post-secondary	0.93	0.73**	0.84***	0.65***
	(0.144)	(0.001)	(0.000)	(0.000)
Other	0.90	0.97	0.80	1.12
	(0.724)	(0.950)	(0.354)	(0.730)
Missing	1.04	0.02*	1.89	2.29
	(0.944)	(0.015)	(0.365)	(0.361)
Employment status (ref. retired)				
(Self-)Employed	1.14**	0.65**	1.14*	0.84

	(0.009)	(0.008)	(0.015)	(0.199)
Unemployed	1.09	0.68	1.29	1.21
	(0.449)	(0.201)	(0.068)	(0.498)
Homemaker or perm. sick	2.13***	2.64***	1.08	1.29***
	(0.000)	(0.000)	(0.070)	(0.000)
Other	1.58**	1.36	1.38**	1.66***
	(0.003)	(0.351)	(0.004)	(0.001)
Missing	0.85	2.49	0.45*	4.20*
	(0.611)	(0.352)	(0.026)	(0.014)
Logged total household income, equivalized	0.90***	0.86**	0.90***	0.85***
	(0.000)	(0.001)	(0.000)	(0.000)
Cube-root transformed household wealth, equivalized	0.99***	0.99***	1.00***	0.99***
	(0.000)	(0.000)	(0.000)	(0.000)
Marital status (ref. married)				
Widowed	1.19**	1.31**	1.08	1.22***
	(0.005)	(0.007)	(0.053)	(0.001)
Divorced	1.05	1.04	1.12*	1.30**
	(0.468)	(0.749)	(0.032)	(0.004)
Never married	1.12	1.18	1.16*	1.11
	(0.129)	(0.244)	(0.044)	(0.404)
Attended social clubs in past year (ref. no)				
Yes	0.74***	0.38***	0.78***	0.41***
	(0.000)	(0.000)	(0.000)	(0.000)
Missing	0.81*	0.04***	0.86	0.03***
	(0.038)	(0.000)	(0.186)	(0.000)
No social network (ref. has a social network)				
Yes	1.17	1.45*	1.55***	1.74***
	(0.106)	(0.023)	(0.000)	(0.000)
Missing	1.09	0.87	1.00	0.69**
	(0.227)	(0.303)	(0.973)	(0.002)
Observations	22322		27011	

Exponentiated coefficients; *p*-values in parentheses
Source: SHARE, wave 6, release 8.0.0, own calculations.
Respondents aged 60+, unweighted, imputed values included.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A10. Variance inflation factors for the three full sets of variables

	Main model	Region model	Migrant model
Migrant status (ref. non-migrant)			
Migrant	1.09	1.05	
Missing	**13.84**	**13.82**	
Female	1.28	1.27	1.29
Smoking (ref. non-smoker)			
Yes	1.15	1.14	1.14
Missing	1.30	1.17	1.33
Alcoholic drinks (ref. none)			
1-2 drinks/week	1.20	1.19	1.20
3+ drinks/week	1.46	1.42	1.47
Missing	2.24	2.24	2.38
Education (ref. primary)			
Secondary	1.96	1.83	2.34
Post-secondary	2.06	1.93	2.57
Other	1.02	1.01	1.07
Missing	**13.89**	**13.88**	1.02
Employment status (ref. retired)			
(Self-) Employed	1.27	1.25	1.30
Unemployed	1.05	1.04	1.10
Homemaker or perm. sick	1.23	1.21	1.20
Other	1.04	1.03	1.07
Missing	1.67	1.67	1.71
Logged total household income, equivalized	2.18	1.74	2.97
Cube-root transformed household wealth, equivalized	1.55	1.41	1.66
Marital status (ref. married)			
Widowed	1.28	1.28	1.32
Divorced	1.09	1.08	1.11
Never married	1.04	1.04	1.06
Attended social clubs in past year (ref. no)			
Yes	1.21	1.18	1.17
Missing	1.73	1.67	1.89
No social network (ref. has a social network)			
Yes	1.02	1.02	1.04
Missing	1.73	1.55	1.84
Age	1.60	1.57	1.86
Health problems in 1-3 dimensions (ref. 0)			
1	1.91	1.91	2.06
2	1.74	1.74	1.81
3	1.74	1.73	1.84
Missing	2.58	2.58	3.02

Country (ref. Germany)			
Austria	1.82		1.58
Sweden	2.04		1.72
Spain	2.59		1.56
Italy	2.36		1.15
France	1.94		1.88
Denmark	1.77		1.23
Greece	2.22		1.25
Switzerland	1.81		2.10
Belgium	2.20		1.90
Czech Republic	2.41		1.57
Poland	1.58		1.14
Luxembourg	1.43		2.22
Portugal	1.54		1.13
Slovenia	2.12		1.77
Estonia	2.60		4.09
Croatia	1.69		2.44
Health care inclusivity (ref. high inclusivity)			
Medium inclusivity		1.70	
Low inclusivity		2.20	
Region (ref. West)			
North		1.65	
East		2.11	
South		1.96	
Years since migration to interview country			1.72
HDI level of origin country (ref. highest)			
Medium			1.66
Least wealthy			1.33
Missing			1.08
Citizenship in country of interview (ref. no)			
Yes			1.49
Missing			1.01

Source: SHARE, wave 6, release 8.0.0, own calculations.
Respondents aged 60+, unweighted, imputed values included.
Values over 10 bolded.

Table A11. Fixed effects regression models with all mediators (weighted)

	Men		Women	
	Pre-frail	Frail	Pre-frail	Frail
Migrant	1.10	0.90	1.25*	1.02
	(0.362)	(0.621)	(0.033)	(0.912)
Missing	1.68	23.11***	0.18**	0.21
	(0.435)	(0.000)	(0.005)	(0.115)
Age	1.04***	1.07***	1.04***	1.07***
	(0.000)	(0.000)	(0.000)	(0.000)
Health problems in 1-3 dimensions (ref. 0)				
1	2.08***	8.83***	2.24***	10.36***
	(0.000)	(0.000)	(0.000)	(0.000)
2	7.27***	117.28***	8.72***	115.06***
	(0.000)	(0.000)	(0.000)	(0.000)
3	24.51***	872.14***	17.77***	599.76***
	(0.000)	(0.000)	(0.000)	(0.000)
Missing	0.06*	0.00***	1.22	41.80**
	(0.032)	(0.000)	(0.768)	(0.001)
Country (ref. Germany)				
Austria	0.97	1.52*	0.99	1.94***
	(0.771)	(0.045)	(0.879)	(0.000)
Sweden	1.15	1.05	1.12	1.31
	(0.123)	(0.818)	(0.213)	(0.134)
Spain	1.14	1.21	1.23	1.98***
	(0.229)	(0.413)	(0.062)	(0.000)
Italy	1.65***	2.95***	1.79***	4.65***
	(0.000)	(0.000)	(0.000)	(0.000)
France	1.39***	1.95**	1.63***	2.94***
	(0.001)	(0.001)	(0.000)	(0.000)
Denmark	1.14	1.53	0.91	1.32
	(0.162)	(0.058)	(0.347)	(0.150)
Greece	2.11***	2.38***	2.04***	3.06***
	(0.000)	(0.000)	(0.000)	(0.000)
Switzerland	0.98	1.13	0.99	0.86
	(0.861)	(0.650)	(0.892)	(0.493)
Belgium	1.24*	1.72**	1.16	2.36***
	(0.017)	(0.005)	(0.101)	(0.000)
Czech Republic	0.87	0.77	1.05	1.04
	(0.174)	(0.233)	(0.624)	(0.818)
Poland	1.22	1.50	1.71***	3.52***

	(0.145)	(0.109)	(0.000)	(0.000)
Luxembourg	1.28	1.59	1.46**	3.05***
	(0.083)	(0.127)	(0.009)	(0.000)
Portugal	1.06	1.14	0.86	2.62**
	(0.763)	(0.680)	(0.497)	(0.003)
Slovenia	0.83	0.79	0.95	1.11
	(0.072)	(0.281)	(0.578)	(0.570)
Estonia	1.46***	1.39	1.32**	1.40
	(0.000)	(0.141)	(0.010)	(0.063)
Croatia	1.33*	2.64***	1.28*	2.65***
	(0.018)	(0.000)	(0.048)	(0.000)
Smoking (ref. non-smoker)				
Yes	1.30***	1.96***	1.13	1.05
	(0.000)	(0.000)	(0.142)	(0.733)
Missing	0.94	1.08	0.87*	0.81
	(0.389)	(0.591)	(0.038)	(0.066)
Alcoholic drinks (ref. none)				
1-2 drinks/week	0.80**	0.52***	0.90	0.55***
	(0.007)	(0.000)	(0.178)	(0.000)
3+ drinks/week	0.79***	0.53***	0.88	0.49***
	(0.000)	(0.000)	(0.051)	(0.000)
Missing	0.58	1.14	0.17*	0.16
	(0.480)	(0.844)	(0.020)	(0.151)
Education (ref. primary)				
Secondary	1.04	0.86	0.92	0.96
	(0.559)	(0.195)	(0.232)	(0.669)
Post-secondary	0.99	0.75	0.82*	0.73*
	(0.939)	(0.073)	(0.021)	(0.042)
Other	1.64	2.13	0.88	1.09
	(0.243)	(0.311)	(0.766)	(0.890)
Missing	0.46	0.03***	3.02	4.06
	(0.236)	(0.000)	(0.064)	(0.141)
Employment status (ref. retired)				
(Self-)Employed	1.16	0.50*	1.15	1.08
	(0.093)	(0.021)	(0.133)	(0.748)
Unemployed	1.31	0.82	0.92	1.19

	(0.154)	(0.633)	(0.691)	(0.672)
Homemaker or perm. sick	1.60*	1.42	1.05	1.24
	(0.019)	(0.239)	(0.520)	(0.050)
Other	1.12	1.23	1.72**	2.04**
	(0.630)	(0.609)	(0.003)	(0.002)
Missing	1.35	3.84	0.17**	0.86
	(0.544)	(0.499)	(0.006)	(0.938)
Logged total household income, equivalized	0.94	0.89	0.94	0.92
	(0.153)	(0.160)	(0.125)	(0.194)
Cube-root transformed household wealth, equivalized	0.99***	0.98***	0.99***	0.99***
	(0.000)	(0.000)	(0.000)	(0.000)
Marital status (ref. married)				
Widowed	1.05	1.08	1.10	1.32**
	(0.637)	(0.642)	(0.147)	(0.005)
Divorced	1.09	0.95	1.17	1.13
	(0.485)	(0.828)	(0.117)	(0.497)
Never married	1.05	1.26	1.03	0.77
	(0.665)	(0.300)	(0.805)	(0.257)
Attended social clubs in past year (ref. no)				
Yes	0.77***	0.39***	0.75***	0.35***
	(0.000)	(0.000)	(0.000)	(0.000)
Missing	0.94	0.09***	0.95	0.03***
	(0.734)	(0.000)	(0.815)	(0.000)
No social network (ref. has a social network)				
Yes	1.22	0.95	1.09	1.50
	(0.225)	(0.870)	(0.636)	(0.116)
Missing	1.28*	0.82	1.21	0.65*
	(0.041)	(0.414)	(0.153)	(0.049)
Observations	22315		27007	

Exponentiated coefficients; p-values in parentheses
Source: SHARE, wave 6, release 8.0.0, own calculations.
Respondents aged 60+, imputed values included.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A12. Random intercept regression models with and without interaction of gender and migrant status

	Base model		Base model with interaction	
	Pre-frail	Frail	Pre-frail	Frail
Migrant status (ref. non-migrant)				
Migrant	1.11**	1.22**	1.08	1.26*
	(0.005)	(0.002)	(0.163)	(0.027)
Missing	0.89	1.03	0.79	1.04
	(0.255)	(0.844)	(0.084)	(0.858)
Female	1.51***	1.90***	1.50***	1.91***
	(0.000)	(0.000)	(0.000)	(0.000)
Female*migrant			1.05	0.96
			(0.472)	(0.758)
Female*missing migration info			1.31	1.03
			(0.188)	(0.924)
Age	1.04***	1.06***	1.04***	1.06***
	(0.000)	(0.000)	(0.000)	(0.000)
Health problems in 1-3 dimensions (ref. 0)				
1	2.23***	8.85***	2.23***	8.85***
	(0.000)	(0.000)	(0.000)	(0.000)
2	8.40***	108.86***	8.39***	108.79***
	(0.000)	(0.000)	(0.000)	(0.000)
3	31.96***	761.44***	31.93***	760.62***
	(0.000)	(0.000)	(0.000)	(0.000)
Missing health problem info	0.16***	2.32	0.16***	2.32
	(0.000)	(0.117)	(0.000)	(0.117)
Observations	49333		49333	

Exponentiated coefficients; *p*-values in parentheses
Source: SHARE, wave 6, own calculations.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A13. Fixed effects regression models with and without interaction of gender and migrant status: Excluding imputed values

	Base model		Base model with interaction	
	Pre-frail	Frail	Pre-frail	Frail
Migrant status (ref. non-migrant)				
Migrant	1.19*** (0.000)	1.27** (0.006)	1.12 (0.091)	1.28 (0.054)
Missing	0.95 (0.701)	1.26 (0.244)	0.90 (0.487)	1.33 (0.286)
Female	1.47*** (0.000)	1.86*** (0.000)	1.45*** (0.000)	1.87*** (0.000)
Female*migrant			1.13 (0.221)	1.01 (0.968)
Female*missing migration info			1.16 (0.544)	0.92 (0.833)
Age	1.04*** (0.000)	1.06*** (0.000)	1.04*** (0.000)	1.06*** (0.000)
Health problems in 1-3 dimensions (ref. 0)				
1	2.29*** (0.000)	11.65*** (0.000)	2.29*** (0.000)	11.65*** (0.000)
2	8.29*** (0.000)	143.24*** (0.000)	8.28*** (0.000)	143.19*** (0.000)
3	36.26*** (0.000)	1,112.07*** (0.000)	36.24*** (0.000)	1,111.35*** (0.000)
Missing	0.30* (0.014)	8.56*** (0.001)	0.30* (0.014)	8.54*** (0.001)
Country (ref. Germany)				
Austria	1.09 (0.224)	1.66*** (0.000)	1.09 (0.231)	1.66*** (0.000)
Sweden	0.98 (0.823)	1.01 (0.941)	0.98 (0.808)	1.01 (0.940)
Spain	1.48*** (0.000)	1.88*** (0.000)	1.48*** (0.000)	1.88*** (0.000)
Italy	2.34*** (0.000)	5.28*** (0.000)	2.34*** (0.000)	5.28*** (0.000)
France	1.35*** (0.000)	1.99*** (0.000)	1.35*** (0.000)	2.00*** (0.000)
Denmark	0.97 (0.634)	1.10 (0.547)	0.96 (0.625)	1.10 (0.544)

Greece	2.62***	4.48***	2.62***	4.48***
	(0.000)	(0.000)	(0.000)	(0.000)
Switzerland	0.83*	0.64*	0.83*	0.64*
	(0.021)	(0.019)	(0.021)	(0.019)
Belgium	1.17*	1.99***	1.17*	1.99***
	(0.027)	(0.000)	(0.027)	(0.000)
Czech Republic	1.27***	1.71***	1.27***	1.71***
	(0.000)	(0.000)	(0.000)	(0.000)
Poland	2.41***	5.27***	2.41***	5.26***
	(0.000)	(0.000)	(0.000)	(0.000)
Luxembourg	0.98	1.28	0.97	1.28
	(0.841)	(0.317)	(0.827)	(0.317)
Portugal	1.60***	3.09***	1.60***	3.09***
	(0.000)	(0.000)	(0.000)	(0.000)
Slovenia	1.12	1.28	1.12	1.28
	(0.116)	(0.099)	(0.113)	(0.098)
Estonia	1.87***	2.15***	1.87***	2.15***
	(0.000)	(0.000)	(0.000)	(0.000)
Croatia	1.87***	5.05***	1.87***	5.05***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	27692		27692	

Exponentiated coefficients; *p*-values in parentheses
Source: SHARE, wave 6, release 8.0.0, own calculations.
Respondents aged 60+, unweighted.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A14. Fixed effects regression models with and without interaction of gender and migrant status: Excluding missing values

	Base model		Base model with interaction	
	Pre-frail	Frail	Pre-frail	Frail
Migrant status (ref. non-migrant)				
Migrant	1.24***	1.52***	1.21*	1.54**
	(0.000)	(0.000)	(0.016)	(0.004)
Female	1.50***	1.76***	1.49***	1.76***
	(0.000)	(0.000)	(0.000)	(0.000)
Female*migrant			1.05	0.99
			(0.672)	(0.943)
Age	1.03***	1.06***	1.03***	1.06***
	(0.000)	(0.000)	(0.000)	(0.000)
Health problems in 1-3 dimensions (ref. 0)				
1	2.25***	10.92***	2.25***	10.92***
	(0.000)	(0.000)	(0.000)	(0.000)
2	7.74***	127.52***	7.74***	127.50***
	(0.000)	(0.000)	(0.000)	(0.000)
3	22.24***	1,047.13***	22.24***	1,047.10***
	(0.000)	(0.000)	(0.000)	(0.000)
Country (ref. Germany)				
Austria	1.04	2.01***	1.04	2.01***
	(0.671)	(0.000)	(0.675)	(0.000)
Sweden	0.98	1.14	0.98	1.14
	(0.816)	(0.502)	(0.810)	(0.501)
Spain	1.34***	2.88***	1.34***	2.88***
	(0.001)	(0.000)	(0.001)	(0.000)
Italy	2.09***	7.66***	2.09***	7.66***
	(0.000)	(0.000)	(0.000)	(0.000)
France	1.26**	2.25***	1.25**	2.25***
	(0.009)	(0.000)	(0.009)	(0.000)
Denmark	0.93	1.21	0.93	1.21
	(0.385)	(0.318)	(0.381)	(0.317)
Greece	2.57***	4.87***	2.57***	4.88***
	(0.000)	(0.000)	(0.000)	(0.000)
Switzerland	0.75**	0.67	0.75**	0.67
	(0.001)	(0.067)	(0.001)	(0.067)
Belgium	1.09	2.23***	1.09	2.23***
	(0.286)	(0.000)	(0.286)	(0.000)

Czech Republic	1.24**	2.03***	1.24**	2.03***
	(0.008)	(0.000)	(0.008)	(0.000)
Poland	2.13***	6.15***	2.13***	6.15***
	(0.000)	(0.000)	(0.000)	(0.000)
Luxembourg	0.87	1.79	0.87	1.79
	(0.406)	(0.102)	(0.409)	(0.103)
Portugal	1.84***	6.96***	1.83***	6.97***
	(0.000)	(0.000)	(0.000)	(0.000)
Slovenia	1.02	1.41	1.02	1.42
	(0.835)	(0.095)	(0.830)	(0.094)
Estonia	1.86***	2.58***	1.85***	2.58***
	(0.000)	(0.000)	(0.000)	(0.000)
Croatia	1.78***	6.06***	1.78***	6.06***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	19171		19171	

Exponentiated coefficients; *p*-values in parentheses
Source: SHARE, wave 6, release 8.0.0, own calculations.
Respondents aged 60+, unweighted.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

STATA code

Data preparation

```
/*============================================================================
File name:	Data preparation
Task:			Combine the necessary datasets of SHARE
Author:			Carlotta Bochert
============================================================================*/



/*------------------------------------------------------------------------------
		0. Purpose, preparations and ToC
------------------------------------------------------------------------------*/
/* This do-file merges and appends modules and waves of SHARE data.
After configuring the do-file in section 1, in section 2, the necessary modules
are first prepared for merging, leaving only the necessary variables and dealing
with multiple imputations (2.1). Then, the modules of each wave are merged
together (2.2) and older waves are appended to wave 6 (2.3). This is in
preparation of "importing" values of variables that are only filled once, in a
respondent's first survey (which might have been any of the waves 1, 2, 4, 5, or
6). Values are imported accordingly in section 2.4, and the cases from other waves
dropped. The resulting data set is called w6_originalvars.dta.

To run this do-file, please:
	- adapt the working directory (section 1)
	- set up a folder structure as referred to in section 1

Table of Contents:
	0. Purpose, preparations and ToC
	1. Configuring the do-file
	2. Data preparation
		2.1 Preparing modules
		2.2 Merging modules
		2.3 Appending waves
		2.4 Importing values

*/

/*------------------------------------------------------------------------------
		1. Configuring the dofile
------------------------------------------------------------------------------*/

* Setup
version 17.0		// Stata version control
clear all		// clear working memory
macro drop _all		// clear macros

* Working directory
global wdir "[your/directory]"


* Define paths to subdirectories
global data		"$wdir\0_data"		// folder for original data
global code		"$wdir\1_dofiles"	// folder for do-files
global posted		"$wdir\2_posted"	// adjusted data
global temp		"$wdir\3_temp"			// folder for temporary files
global table		"$wdir\4_tables"	// folder for table output
global graph		"$wdir\5_graphs"		// folder for graph output
```

```
*install necessary ado files
	*ssc install xtmrho
	*ssc install sepscatter

/*---------------------------------------------------------------------------
		2. Data preparation
---------------------------------------------------------------------------*/

/*---------------------------------------------------------------------------
		2.1 Preparing modules
---------------------------------------------------------------------------*/
/*Prepare the modules of wave 6 in a loop, capturing unexecutable commands that
 refer to variables not included in the module currently being prepared: */

 foreach m in ac br cv_r dn ep gv_imputations gv_health gv_isced gv_weights ///
	hc mh ph sn sp technical_variables {
	* Load module
		use "$data/sharew6_rel8-0-0_`m'", clear

	*keep only the necessary variables
		capture keep mergeid ac035d5
		capture keep mergeid br016_ br002_ br003_ br029_ br039 br040 br623
		capture keep mergeid hhsize
		capture keep mergeid country dn002_ dn003_ dn004_ dn005c dn006 dn007	///
			dn014_ dn042_
		capture keep mergeid ep005_
		capture keep mergeid gender age bmi mstat hnetw thinc thinc2 isced empstat	///
			gender_f age_f bmi_f mstat_f hnetw_f thinc_f thinc2_f isced_f empstat_f
		capture keep mergeid adl iadl chronicw6c maxgrip
		capture keep mergeid isced1997_r
		capture keep mergeid cchw_w6 cciw_w6
		capture keep mergeid hc114_ hc115_
		capture keep mergeid mh013_ mh011_ mh012_
		capture keep mergeid ph012_ ph013_ ph048d1 ph048d4
		capture keep mergeid sn012 sn017
		capture keep mergeid sp002
		capture keep mergeid mn024

	*sort the data by mergeid (person ID)
		sort mergeid

	*Save the prepared module
		save "$posted/w6_`m'_prep", replace
	}

	*deal with imputations from gv_imputations

		use "$posted/w6_gv_imputations_prep", clear

		*gen var identifying duplicates (on mergeid)
		sort mergeid
		by mergeid: gen dupid = cond(_N==1,0,_n)
		*browse if dupid > 0

		tab dupid, m
		*drop if dupid > 1 & gender_f == 3 & age_f == 3 & bmi_f == 3 & mstat_f == 3
		//this leaves all imputed values in to make e.g. a mean of them.
```

```
		/*generate variables for mode/mean and standard deviation
		*as a loop to easily add variables:
		*/

		*first, metric vars
		foreach v in age bmi hnetw thinc thinc2 {
			capture drop `v'_mean
			bysort mergeid: egen `v'_mean = mean(`v')
			capture drop `v'_sd
			bysort mergeid: egen `v'_sd = sd(`v')
			/*For some reason even if there are only identical regular
			observations, sometimes sd is not 0 – force it?*/
		}

		*mode for categorical vars
		foreach v in gender mstat {
			capture drop `v'_mode
			bysort mergeid: egen `v'_mode = mode(`v'), minmode
			/*minmode decides between two values that are both the
			most common and takes the lower one. Bias, of course.
			No way to do this randomly, I don't think. */
		}


		/*Since I could not use multiple imputation procedures, I delete all multiples*/
			drop if dupid > 1

		*Save the prepared module
			save "$posted/w6_gv_imputations_prep", replace


/* Loop to load values from other waves and their modules
(e.g., migrant status is only asked about in the first interview and is then
 missing in all subsequent datasets, so has to be "imported" like this) */

	foreach w in 1 2 4 5 {
		foreach m in br ep dn ph {
			use "$data\sharew`w'_rel8-0-0_`m'", clear

			*keep only the necessary variables
				capture keep mergeid hhid`w' br016_ br002_ br003_ br029_
				capture keep mergeid hhid`w' ep005_
				capture keep mergeid hhid`w' dn004_ dn005c dn006 dn007 dn014_
				capture keep mergeid hhid`w' ph013_

			*sort the data by mergeid (person ID)
				sort mergeid

			*Save the prepared module
			save "$posted\w`w'_`m'_prep", replace
		}
	}


/*------------------------------------------------------------------------
		2.2 Merging modules
--------------------------------------------------------------------------*/
*Wave 6
	*use prepared wave 6, dn module
		use "$posted\w6_dn_prep", clear
```

```
	*sort by mergeid
		sort mergeid
	*merge modules
		foreach m in ac br cv_r dn ep gv_health gv_imputations gv_isced  ///
				gv_weights hc mh ph sn sp technical_variables {
			merge 1:1 mergeid using "$posted/w6_`m'_prep", keep(1 3)
			drop _merge
		}
/*if using multiple imputations:
	*merge gv_imputations separately (1:m)
			merge 1:m mergeid using "$posted/w6_gv_imputations_prep", keep(1 3)
			*/
	*save resulting dataset
		save "$posted/w6_merged", replace
		use "$posted/w6_merged", clear
		*sort mergeid


*Merge br, ep, ph and dn for the rest of the waves
	foreach w in 1 2 4 5 {
		use "$posted/w`w'_dn_prep", clear
		*sort by mergeid
			sort mergeid
		*merge modules
			merge 1:1 mergeid using "$posted/w`w'_br_prep"
			drop _merge
			merge 1:1 mergeid using "$posted/w`w'_ep_prep"
			drop _merge
			merge 1:1 mergeid using "$posted/w`w'_ph_prep"
			drop _merge
		*save resulting dataset
			save "$posted/w`w'_merged", replace
	}


/*------------------------------------------------------------------------
		2.3 Appending waves
------------------------------------------------------------------------*/
*use wave 6 as the base
	use "$posted/w6_merged", clear
	*append waves in a loop
		foreach w in 5 4 2 1 {
			append using "$posted/w`w'_merged", gen(wave`w't6)
		}

	*sort by mergeid and wave
		sort mergeid wave5t6 wave4t6 wave2t6 wave1t6

	*save appended set
		save "$posted/w1_2_4_5_6", replace


/*------------------------------------------------------------------------
		2.4 Importing values
------------------------------------------------------------------------*/
*use appended dataset
	use "$posted/w1_2_4_5_6", clear
*impute later value (i.e., earlier wave) for variables in list
	foreach v in br001 br002_ br003_ dn004_ dn005c dn006 dn007 dn014_ ph013_ {
		replace `v' = `v'[_n+1] if mergeid==mergeid[_n+1] & `v' == .
```

```
		replace `v' = `v'[_n+1] if mergeid==mergeid[_n+1] & `v' == .
		replace `v' = `v'[_n+1] if mergeid==mergeid[_n+1] & `v' == .
		replace `v' = `v'[_n+1] if mergeid==mergeid[_n+1] & `v' == .

	/*	replace `v' = `v'[_n+1] if mergeid==mergeid[_n+1] & `v' == .
		replace `v' = `v'[_n+1] if mergeid==mergeid[_n+1] & `v' == .
		replace `v' = `v'[_n+1] if mergeid==mergeid[_n+1] & `v' == .
		replace `v' = `v'[_n+1] if mergeid==mergeid[_n+1] & `v' == . */
		/*the same code has to be run as many times as waves were appended,
		plus, if using, the four extra imputed observations in wave 6;
		those are no longer in the dataset, though. */
	}

*save imputed set
	save "$posted/w1_2_4_5_6_imp", replace
	use "$posted/w1_2_4_5_6_imp", clear


*drop observations not from wave 6
	keep if wave5t6 == 0

*add numerical values to value labels
	numlabel, add

*save dataset with non-recoded variables
	save "$posted/w6_originalvars", replace
	use "$posted/w6_originalvars", clear
```

Re-coding variables

```
/*===========================================================================
File name:	Recoding variables
Task:			Recode SHARE variables for analysis
Author:			Carlotta Bochert
=============================================================================*/


/*---------------------------------------------------------------------------
		0. Purpose, preparations and ToC
---------------------------------------------------------------------------*/
/* This do-file prepares the variables and sample for analysis. After configuring
 the do-file in section 1, Variables are recoded in section 2.1 to 2.3; additional
 specifications of missing variables are coded in section 2.4. Finally, in
 section 3., the analytical sample is restricted and alternative sample
 specifications (no imputations, no missing values) are saved, as well.

To run this do-file, please:
	- adapt the working directory (section 1)
	- set up a folder structure as referred to in section 1

Table of Contents:
	0. Purpose and ToC
	1. Configuring the dofile
	2. Recoding and generating variables
		2.1 Main variables
		2.2 Controls and moderators
		2.3 Mediators
		2.4 "Missing"-specifications
	3. Restricting the analytical sample(s)
```

```
*/

/*------------------------------------------------------------------------------
                1. Configuring the dofile
------------------------------------------------------------------------------*/

* Setup
version 17.0        // Stata version control
clear all           // clear working memory
macro drop _all     // clear macros

* Working directory
global wdir  "[your/directory]"


* Define paths to subdirectories
global data         "$wdir\0_data"          // folder for original data
global code         "$wdir\1_dofiles"               // folder for do-files
global posted       "$wdir\2_posted"                // adjusted data
global temp         "$wdir\3_temp"                  // folder for temporary files
global table        "$wdir\4_tables"        // folder for table output
global graph        "$wdir\5_graphs"                // folder for graph output

*install necessary ado files
        ssc install egenmore

/*-----------------------------------------------------------------------------
                2. Recoding and generating variables
------------------------------------------------------------------------------*/
*load the prepared dataset
        use "$posted/w6_originalvars", clear


/*-----------------------------------------------------------------------------
                2.1 Main variables
------------------------------------------------------------------------------*/

*Frailty index
        *1) Exhaustion: Positive answer on mh013_
                gen f_ex = .
                replace f_ex = 1 if mh013_ == 1 //yes if yes
                replace f_ex = 0 if mh013_ == 5 //no if no
                replace f_ex = 98 if inlist(mh013_, -1, -2) //uncodable answers
                replace f_ex = 99 if mh013_ == . //missing values

                *Labeling
                label define yn_c 1 "Yes" 0 "No" 97 "Not applicable" 98 "Uncodable" ///
                99 "Missing" 997 "Not applicable" 998 "Uncodable" 999 "Missing"    , replace
                label values f_ex yn_c
                label var f_ex "Frailty 1: Exhaustion"

                *Check recoded var
                tab mh013_ f_ex if wave5t6 == 0, m

        *2) Shrinking: Diminution in desire for food (mh011_) or eating less (mh012_)
                        *Note: mh012 is only asked if mh011 is uncodable.
                gen f_sh = .
                replace f_sh = 1 if mh011_ == 1 | mh012_ == 1 //yes if yes
                replace f_sh = 0 if mh011_ == 2 | inlist(mh012_, 2, 3) //no if no
```

```
replace f_sh = 98 if inlist(mh011_, -1, -2) | inlist(mh012_, -1, -2)
        //uncodable answers
replace f_sh = 99 if mh011_ == . //missing values

*Labeling
label values f_sh yn_c
label var f_sh "Frailty 2: Shrinking"

*Check recoded var
tab mh011_ f_sh if wave5t6 == 0 , m
tab mh012_ f_sh if wave5t6 == 0 , m


*3) Weakness: Hand grip strength (maxgrip) below cutoff-point by bmi and gender
    *This uses bmi from gv_imputations. So it's sometimes an imputation (see flag, bmi_f).
    gen f_we = .
    *First, men:
    replace f_we = 0 if dn042_ == 1 & (bmi_mean <= 24 & maxgrip > 29)          ///
                        | ((bmi_mean > 24 & bmi_mean <= 28) & maxgrip > 30)     ///
                        | (bmi_mean > 28 & maxgrip > 32)
                        //no if maxgrip is over the threshold
    replace f_we = 1 if dn042_ == 1 & (bmi_mean <= 24 & maxgrip <= 29)         ///
                        | ((bmi_mean > 24 & bmi_mean <= 28) & maxgrip <= 30)    ///
                        | (bmi_mean > 28 & maxgrip <= 32)
                        //yes if maxgrip is under the threshold
    *Then, women:
    replace f_we = 0 if dn042_ == 2 & (bmi_mean <= 23 & maxgrip > 17)          ///
                        | ((bmi_mean > 23 & bmi_mean <= 26) & maxgrip > 17.3)   ///
                        | ((bmi_mean > 26 & bmi_mean <= 29) & maxgrip > 18)     ///
                        | (bmi_mean > 29 & maxgrip > 21)
                        //no if maxgrip is over the threshold
    replace f_we = 0 if dn042_ == 2 & (bmi_mean <= 23 & maxgrip <= 17)         ///
                        | ((bmi_mean > 23 & bmi_mean <= 26) & maxgrip <= 17.3)  ///
                        | ((bmi_mean > 26 & bmi_mean <= 29) & maxgrip <= 18)    ///
                        | (bmi_mean > 29 & maxgrip <= 21)
                        //yes if maxgrip is under the threshold


    *regardless of gender:

    replace f_we = 98 if inlist(maxgrip, -1, -2) //uncodable answers
    //bmi and gender have no refusal option; bmi has been imputed for missings

    replace f_we = 99 if dn042_ == . | bmi == . | maxgrip == . //missing values

    *Labeling
    label values f_we yn_c
    label var f_we "Frailty 3: Weakness"

    *Check recoded var
    tab gender f_we if wave5t6 == 0, m row
    *browse dn042_ maxgrip bmi bmi_f if f_we == . & wave5t6 == 0

*4) Slowness: Selection of either ph048d1 or ph048d4
    gen f_sl = .
    replace f_sl = 0 if ph048d1 == 0 & ph048d4 == 0 //no if both are no
    replace f_sl = 1 if ph048d1 == 1 | ph048d4 == 1 //yes if either is yes
    replace f_sl = 98 if inlist(ph048d1, -1, -2) | inlist(ph048d4, -1, -2)
            //uncodable answers
    replace f_sl = 99 if ph048d1 == . & ph048d4 == . //missing values
```

```
		*Labeling
		label values f_sl yn_c
		label var f_sl "Frailty 4: Slowness"

		*Check recoded var
		tab ph048d1 f_sl if wave5t6 == 0 , m
		tab ph048d4 f_sl if wave5t6 == 0 , m

*5) Low activity: Answers 3 and 4 on br016_
		gen f_lo = .
		replace f_lo = 1 if inlist(br016_, 3, 4) //yes if 3 times a month or fewer
		replace f_lo = 0 if inlist(br016_, 1, 2) //no if more often
		replace f_lo = 98 if inlist(br016_, -1, -2) //uncodable answers
		replace f_lo = 99 if br016_ == . //missing values

		*Labeling
		label values f_lo yn_c
		label var f_lo "Frailty 5: Low activity"

		*Check recoded var
		tab br016_ f_lo if wave5t6 == 0 , m

/*Now, dichotomous variables for being frail (3+ criteria = 1)
and for pre-frail (1-2 criteria) */
		egen count_frail = anycount(f_*), values(1)

		*Pre-frail
			gen prefrail = .
			replace prefrail = 1 if inrange(count_frail, 1, 2)
			replace prefrail = 0 if inlist(count_frail, 0, 3, 4, 5)

			tab count_frail prefrail if wave5t6 == 0 , m
		*Frail
			gen frail = .
			replace frail = 1 if inrange(count_frail, 3, 5)
			replace frail = 0 if inlist(count_frail, 0, 1, 2)

			tab count_frail frail if wave5t6 == 0 , m
			tab frail prefrail if wave5t6 == 0, m
			tab frail prefrail if wave5t6 == 0 & age >= 60, m

/*Now, one variable for being robust (no criteria fulfilled), pre-frail
(1-2 fulfilled) and frail (3-5 fulfilled) */

		gen frailtype = .
			replace frailtype = 2 if inrange(count_frail, 3, 5)
			replace frailtype = 1 if inrange(count_frail, 1, 2)
			replace frailtype = 0 if count_frail == 0

*label variable and values
		label var frailtype "Frailty type"
		label define frailtype_l 0 "Robust" 1 "Pre-frail" 2 "Frail"
		label values frailtype frailtype_l

		tab frailtype if age >=60, m

/*Result: ~7000 respondents frail, ~30,000 each pre-frail and robust.
		For 60+, 6000 frail, 25000 pre, 21000 robust. */
```

```
*Migration experience
	gen mig = .
		replace mig = 0 if dn004_ == 1 //no if yes
		replace mig = 1 if dn004_ == 5 // yes if no
		replace mig = 98 if inlist(dn004_, -1, -2) //uncodable answers
		replace mig = 99 if dn004_ == .

		label values mig yn_c
		label var mig "Immigrant"

		tab dn004_ mig, m

*Gender recode
	tab gender, m
	recode gender (1=0 "Male") (2=1 "Female"), gen(gendercat)

	label var gendercat "Gender"
	tab gender gendercat, m

/*---------------------------------------------------------------------------
		2.2 Controls and moderators
---------------------------------------------------------------------------*/

*Age: age as from gv_imputations
	codebook age // no missings
	tab age_f, m // very few imputations

	*generate age squared
		gen age2 = age^2
		codebook age2

*Health:
	*(Instrumental) Activities of Daily Living, Chronic conditions
		tab1 adl iadl chronicw6c, m
			foreach v in adl iadl chronicw6c {
				gen `v'_n = .
				replace `v'_n = `v' if inrange(`v', 0, 15)
				replace `v'_n = 98 if inlist(`v', -2, -1)
				replace `v'_n = 99 if `v' == .

				gen `v'_2 = .
				replace `v'_2 = 0 if `v' == 0
				replace `v'_2 = 1 if inrange(`v', 1, 15)
				replace `v'_2 = 98 if inlist(`v', -2, -1)
				replace `v'_2 = 99 if `v' == .

				label values `v'_2 yn_c

			*and with missings as missing
				gen `v'_2_m = .
				replace `v'_2_m = 0 if `v' == 0
				replace `v'_2_m = 1 if inrange(`v', 1, 15)
				replace `v'_2_m = . if inlist(`v', -2, -1, .)
				}


				label var adl_n "Number of ADL limitations"
				label var iadl_n "Number of IADL limitations"
				label var chronicw6c_n "Number of chronic conditions"
				label var adl_2 "ADL limitations (y/n)"
```

```
				label var iadl_2 "IADL limitations (y/n)"
				label var chronicw6c_2 "Chronic conditions (y/n)"

				tab adl adl_n, m
				tab iadl iadl_2, m
				tab chronicw6c chronicw6c_2, m

			*combine into one index
				*using whether someone has one (_2_m) because their numbers/the scales
of _n aren't really comparable
				egen health = rowtotal(adl_2_m iadl_2_m chronicw6c_2_m), missing
				replace health = 99 if chronicw6c_2 ==. | adl_2_m == . | iadl_2_m == .

				label var health "Health problems in 0-3 dimensions"
				label define health_l 99 "Missing"
				label values health health_l

				tab health chronicw6c_2_m, m

*Duration of stay
		tab dn006 dn004, m
			*improbable value of immigration in 1800s
			*(check with birth year and put into uncodable)

		gen stay = .
		replace stay = 2015-dn006 if dn004 == 5
		replace stay = .b if inlist(dn006, -1, -2) | inlist(dn004, -1, -2)
		replace stay = .a if dn004 == . | (dn006 == . & dn004 != 1)
		replace stay = .a if dn006 < dn003 //implausible migration years
		replace stay = .c if dn004 == 1 //nonmigrants

		label define threedigitmissing .c "Not applicable" .b "Uncodable"	///
											.a "Missing", replace
		label values stay threedigitmissing
		label var stay "Years since migration to interview country"

		codebook stay
		tab stay if dn004 == 5, m
		tab stay, m


*HDI origin countries
		codebook dn005c
		tab dn005c
			*157 unique values,
			/*incl some for former German territories (cutoff of timing of
			immigration, see above?) either way, a lot to research/code... */
			gen hdi = .

			replace hdi = 1 if inlist(dn005c, 32, 36, 40, 42, 56, 124, 152, 191, ///
			196, 200, 203, 208, 233, 246, 250, 276, 300, 344, 348, 352, 372, 376, ///
			380,392, 410, 428, 438, 440, 442, 528, 554, 578, 616, 620, 642, 643, ///
			702, 703, 705, 724, 752, 756, 826, 840)
			//HDI 0.827 and over

			replace hdi = 2 if inlist(dn005c, 8, 12, 31, 51, 70, 76, 100, 112, ///
			144, 156, 170, 188, 192, 214, 218, 268, 364, 398, 400, 422, 434,  ///
			458, 480, 484, 499, 604, 688, 740, 764, 788, 792, 804, 807, 810, ///
			858, 860, 862, 890, 1040)
			 //HDI between 0.700 and 0.827
```

```
			replace hdi = 3 if inlist(dn005c, 4, 24, 68, 108, 116, 120, 132, ///
			180, 204, 222, 226, 230, 231, 232, 266, 270, 288, 324, 332, 340, ///
			356, 360, 368, 384, 404, 417, 418, 430, 450, 466, 478,  498, 504, ///
			508, 558, 566, 586, 600, 608, 624, 646, 678, 686, 704, 706, 710, ///
			729, 760, 762, 768, 795, 818, 834, 887, 894, 1010)
			 //HDI under 0.700

			replace hdi = 98 if inlist(dn005c, -4, -2, -1)		///
			| inlist(dn004, -1, -2) | inlist(dn005c, 2, 158, 234, 254, 258, ///
						275, 312, 474, 531, 1024, 1030, 1031, 1050, 1070, 1080)
				/*recorded countries and areas without clear HDI: Africa, Taiwan,
				Faroe islands, French Guiana, French Polynesia,
				occupied Palestinian territory, Guadeloupe, Martinique, Curacao,
				Afghan-Turkish, Former German Territories (+Eastern), Minor Asia,
				Former Austria-Hungary, Kurdistan */

			replace hdi = 99 if dn005c == . & dn004 != 1 | dn004 == .
			replace hdi = 97 if dn004 == 1 //nonmigrants

			label var hdi "HDI of origin country (immigrants)"
			label define hdi_l 1 "Wealthiest" 2 "Medium" 3 "Least wealthy" ///
			97 "Not applicable" 98 "Uncodable" 99 "Missing"
			label values hdi hdi_l

			tab dn005c if hdi ==2

			tab hdi, m


*Citizenship
		tab dn007 if dn004 == 5, m
		gen citizen = .
		replace citizen = 0 if dn007 == 5 //no
		replace citizen = 1 if dn007 == 1 //yes
		replace citizen = 98 if inlist(dn007, -2, -1)
		replace citizen = 99 if dn007 == .

		label var citizen "Citizenship in country of interview"
		label values citizen yn_c

		tab citizen, m

		*and citizenship with non-migrants coded as n/a
			gen imcitizen = citizen
			replace imcitizen = 97 if dn004 == 1
			replace imcitizen = 99 if dn004 == .

			label var imcitizen "Citizenship in country of interview (immigrants)"
			label values imcitizen yn_c

			tab imcitizen citizen, m

*Country categories

	*MIPEX health care strand performance
		gen mipex = .
			replace mipex = 1 if inlist(country, 11, 13, 16, 20)
			replace mipex = 2 if inlist(country, 12, 15, 17, 18, 23, 28, 31, 33)
```

```
replace mipex = 3 if inlist(country, 19, 29, 34, 35, 47)

label define mipex_l 1 "High inclusivity" 2 "Medium inclusivity"        ///
3 "Low inclusivity"
label values mipex mipex_l
label var mipex "MIPEX health care strand performance"

tab country mipex, m

*Geographical Region (NESW)

gen nesw = .
replace nesw = 1 if inlist(country, 13, 18, 35)
replace nesw = 2 if inlist(country, 28, 29, 34, 47)
replace nesw = 3 if inlist(country, 15, 16, 19, 33)
replace nesw = 4 if inlist(country, 11, 12, 17, 20, 23, 31)

label define nesw_l 1"North" 2"East" 3"South" 4"West"
label values nesw nesw_l
label var nesw "Geographical region"

tab country nesw, m

/*---------------------------------------------------------------------------
2.3 Mediators
---------------------------------------------------------------------------*/
*Health behaviour mediators

*Smoking
*recode br002, smoking presently
gen smokenow = .
replace smokenow = 1 if br002 == 1 //yes if yes
replace smokenow = 0 if br002 == 5 // no if no
replace smokenow = 98 if inlist(br002, -1, -2) //uncodable answers
replace smokenow = 99 if br002 == .

label values smokenow yn_c
label var smokenow "Smoking currently"

tab br002 smokenow, m //almost 18,000 missing


*Alcohol consumption
*last 7 days: br039 (if) and br040 (how many units)
tab br040 br039, m

/*code a three-category variable: three or more drinks per week, less
than 3 drinks per week, and non-drinkers*/
gen drink73 = .
replace drink73 = 1 if inrange(br040, 1, 2)   //1-2 drinks
replace drink73 = 2 if inrange(br040, 3, 1000)         //3+ drinks
replace drink73 = 99 if br040 == . | br039 == .         //missings
replace drink73 = 0 if br039 == 5 | br040 == 0         //no drinks
replace drink73 = 98 if inlist(br040, -1, -2) | ///
inlist(br039,         -1,         -2)
//uncodable answers

label define drink73_l  0 "No drinks" 1 "1-2 drinks/week" ///
2 "3+ drinks/week" 98 "Uncodable" 99 "Missing"
label var drink73 "Alcoholic drinks per week"
```

```
				label values drink73 drink73_l

				tab drink73 br039, m



*Economic mediators

		*Household income and wealth need to be equivalized (by household size)
			tab hhsize, m //exclude those who say their household has size 0?
			tab1 hnetw thinc thinc2 if hhsize == 0, m //their income/wealth is . anyway


			/*thinc, thinc2 and hnetw by default include imputations (see _f vars).
			 around 40% for net worth, 35% for thinc and 15% for thinc2.
			 */
			 tab1 hnetw_f thinc_f thinc2_f

			*equivalize the measures using square root of the household size
				*this uses the mean of imputations for those with imputed values
				gen hnetweq = hnetw_mean/sqrt(hhsize)
					label var hnetweq "Household wealth, equivalized"
				gen thinceq = thinc_mean/sqrt(hhsize)
					label var thinceq "Total household income (agg.), equivalized"
				gen thinc2eq = thinc2_mean/sqrt(hhsize)
					label var thinc2eq "Total household income, equivalized"
				codebook hnetweq thinceq thinc2eq

			browse hhsize hnetw hnetweq thinc thinceq thinc2 thinc2eq

			*check for (close to) normal distributions
				*Histograms
					hist thinceq, name(hist_thinceq, replace)
					hist thinc2eq, name(hist_thinc2eq, replace)
					hist hnetweq, name(hist_hnetweq, replace)
				*-> All highly right-skewed

				*generate logged variables:
				foreach v in thinceq thinc2eq hnetweq {
					capture drop log_`v'
					gen log_`v' = log(`v')
					hist log_`v', name(hist_log_`v', replace) normal kdensity
				}
				*-> Better, though wealth is now a bit left-skewed.
				*Also, the negative values of wealth are not well-served with this transformation
(automatically set to missing)
				*thinc(2) are slightly bimodal
				label var log_hnetweq "Logged household wealth, equivalized"
				label var log_thinceq "Logged total household income (agg.), equivalized"
				label var log_thinc2eq "Logged total household income, equivalized"

				*cube-root transform wealth
				gen c3_hnetweq = sign(hnetweq) * abs(hnetweq)^(1/3)
				hist c3_hnetweq, name(hist_c3_hnetweq, replace) normal kdensity
					*indeed better than log, takes into account negative values as well
				label var c3_hnetweq "Cube-root transformed household wealth, equivalized"


		*Employment status
			codebook ep005_
```

```
/*
                         Freq.   Numeric  Label
                                                   70        -2  Refusal
                                                   40        -1  Don't know
                                            185,609         1  Retired
                                             26,685         2  Employed or self-employed
                                                                              (including
working for family
                                                                              business)
                                                 3,195         3  Unemployed
                                                 5,025         4  Permanently sick or disabled
                                             20,910         5  Homemaker
                                                 4,330        97  Other
                                                   450         .
*/
*recode into: Retired, working, homemaker or sick, other, 98, 99
	gen empl = .
		replace empl = 1 if ep005_ == 1 //retired
		replace empl = 2 if ep005_ == 2 //employed
		replace empl = 3 if ep005_ == 3 //unemployed
		replace empl = 4 if inlist(ep005_, 4, 5) //homemaker and disabled
		replace empl = 5 if ep005_ == 97
		replace empl = 98 if inlist(ep005_, -1, -2) //uncodable answers
		replace empl = 99 if ep005_ == .

		label define empl_l 1 "Retired" 2 "(Self-)Employed" ///
		3 "Unemployed" 4 "Homemaker or perm. sick" 5 "Other"  ///
		98 "Uncodable" 99 "Missing"
		label values empl empl_l
		label var empl "Employment status"

		tab ep005_ empl, m

*Education
/*based on ISCED 1997 categories, divided into three categories:
None or primary (ranks 0, 1); secondary (ranks 2, 3); and post-secondary
(ranks 4-6)*/
tab isced1997_r, m // adds None, Still in School, Other, and Refusal/Don't know
gen edu = .
	replace edu = 0 if inlist(isced1997_r, 0, 1)
	replace edu = 1 if inlist(isced1997_r, 2, 3)
	replace edu = 2 if inlist(isced1997_r, 4, 5, 6)
	replace edu = 3 if inlist(isced1997_r, 95, 97)
	replace edu = 98 if inlist(isced1997_r, -2, -1)
	replace edu = 99 if isced1997_r == .

	label var edu "Formal education"
	label define edu_l 0 "None/Primary" 1 "Secondary" 2 "Post-secondary"		///
	3 "Other" 98 "Uncodable" 99 "Missing"
	label values edu edu_l

	tab isced1997_r edu, m

*Social mediators

*Marital status
	tab mstat, m
	*married or in a registered partnership, widowed, divorced, never married
	recode mstat (1 2 3 = 1 "Married or partnered") (6 = 2 "Widowed")		///
	(5 = 3 "Divorced") (4 = 4 "Never married"), gen(marital) label(marital_l)
```

```
        label var marital "Marital status"
        tab mstat marital, m

    *Club activities: Participated in the last year
        tab ac035, m
        *recode the missing values
            gen socialclub = .
                replace socialclub = 1 if ac035 == 1 //yes if yes
                replace socialclub = 0 if ac035 == 0 // no if no
                replace socialclub = 98 if inlist(ac035, -1, -2) //uncodable answers
                replace socialclub = 99 if ac035 == .

                label values socialclub yn_c
                label var socialclub "Attended a social club in past year"

                tab ac035 socialclub, m

    *Empty network: generate a variable for whether respondents' network is empty
        gen netempty = .
            replace netempty = 0 if sn012 != .
            replace netempty = 1 if sn017 != .
            replace netempty = 99 if sn012 == . & sn017 == .

            label values netempty yn_c
            label var netempty "Social network is empty"

            tab netempty, m


/*---------------------------------------------------------------------------
            2.4 "Missing"-specifications
---------------------------------------------------------------------------*/

*generate the vars that have imputed missing values as missing for those values

        *frailty, marital, income, wealth, age.

        *Frailty
        gen frailtype_ni = frailtype
        replace frailtype_ni = 99 if bmi_f != 1 & bmi_f != 3
        label define frailtype_l 99 "Missing", add
        label values frailtype_ni frailtype_l

        *Marital status
        gen marital_ni = marital
        replace marital_ni = 99 if mstat_f != 1 & mstat_f != 3
        label define marital_l 99 "Missing", add
        label values marital_ni marital_l

        tab marital_ni mstat_f, m

        *Monetary variables
            foreach v in thinc thinc2 hnetw {
                foreach s in `v'eq log_`v'eq {
                    capture drop `s'_ni
                    gen `s'_ni = `s'
                    replace `s'_ni = . if `v'_f != 1 & `v'_f != 3

                    local label: variable label `s'
```

```
					label variable `s'_ni `"`label' (excl. imputations)"'
					codebook `s'_ni
					}
				}

					gen c3_hnetweq_ni = c3_hnetweq
					replace c3_hnetweq_ni = . if hnetw_f  != 1 & hnetw_f != 3
					label variable c3_hnetweq_ni ///
					"Cube-root transformed household wealth, equivalized (excl. im-
putations)"
					codebook c3_hnetweq_ni

		*Age
			foreach v in age age2 {
				capture drop `v'_ni
				gen `v'_ni = `v'
				replace `v'_ni = . if age_f != 1 & age_f != 3
			}

	*define globals for the lists of variables

	global contvarsimp age stay c3_hnetweq log_thinceq log_thinc2eq
		*all of these have their missings either imputed or coded .x

	global contvarsnoimp age_ni stay c3_hnetweq_ni log_thinceq_ni log_thinc2eq_ni

	global catvars frailtype gendercat mig nesw mipex hdi imcitizen 	citizen	///
		drink73 smokenow empl edu marital socialclub netempty health



	*Code 97, 98, 99 for the categorical variables as missing
		foreach v in $catvars {
			capture drop `v'_m
			gen `v'_m = `v'
			replace `v'_m = . if inlist(`v', 97, 98, 99, 997, 998, 999)
			local label: variable label `v'
			label variable `v'_m `"`label'"'
			labellist `v'
			label values `v'_m `"`r(lblname)'"'
			codebook `v'_m
		}
		tab mig mig_m, m

		*and one where all missing/uncodeable etc ones are put into the same category
		foreach v in $catvars {
			capture drop `v'_99
			gen `v'_99 = `v'
			replace `v'_99 = 99 if inlist(`v', 97, 98, 99, 997, 998, 999)
			local label: variable label `v'
			label variable `v'_99 `"`label'"'
			labellist `v'
			label values `v'_99 `"`r(lblname)'"'
			codebook `v'_99
		}
		tab mig_m mig_99, m


*Weights
	*mean-standardizing weight (so that mean is 1 and scale is adjusted)
```

```
		quietly sum cciw_w6
		gen weightm = cciw_w6/r(mean)
		sum weightm



*save dataset with the prepared variables
	save "$posted/w6_general", replace
	use "$posted/w6_general", clear

/*--------------------------------------------------------------------------
		3. Restricting the analytical sample(s)
--------------------------------------------------------------------------*/

	*Drop Israeli respondents
		drop if country == 25 //2035 respondents dropped

	*Drop those below age 60
		drop if age < 60 //16091 respondents dropped

	*Drop those living in nursing homes
		drop if mn024 == 2 //626 respondents dropped

		*49,333 observations remaining

	*drop numlabels
	numlabel, remove

*save dataset with the restricted sample
	save "$posted/w6_restricted", replace
	use "$posted/w6_restricted", clear


*Dataset without imputed values
	foreach v in age_ni c3_hnetweq_ni log_thinc2eq_ni marital_ni frailtype_ni {
		drop if `v' == .
	}
		*2 for age; ~22.000 for monetary; then no more for marital status or bmi/frailty


	save "$posted/w6_noimp", replace
	use "$posted/w6_noimp", clear

*And deleting missing values of categorical variables

	global catvarsmis mig_99 gendercat frailtype health_99		///
		drink73_99 smokenow_99 empl_99 edu_99				///
		marital_99 socialclub_99 netempty_99			///
		citizen_99

	foreach v in $catvarsmis {
		drop if `v' == 99
	}

	save "$posted/w6_nomis", replace
	use "$posted/w6_nomis", clear
```

Descriptives

```
/*================================================================================
File name:        Descriptives
Task:                     Compute descriptive and bivariate measures
Author:                   Carlotta Bochert
================================================================================*/

/*------------------------------------------------------------------------------
                0. Purpose, preparations and ToC
------------------------------------------------------------------------------*/
/* This do-file computes and tabulates/graphs (bivariate) summary statistics of
 the analytical sample. After configuring the do-file in section 1, section 2
 creates tables of weighted summary statistics a) by migration status and gender
 and b) by frailty state, as well as a table of the share of migrants per country.
 Section 3 creates graphs of some bivariate distributions.

To run this do-file, please:
            - adapt the working directory (section 1)
            - set up a folder structure as referred to in section 1

Table of Contents:
            0. Purpose and ToC
            1. Configuring the dofile
            2. Descriptive/bivariate tables
            3. Bivariate graphs

*/

/*------------------------------------------------------------------------------
                1. Configuring the dofile
------------------------------------------------------------------------------*/

* Setup
version 17.0         // Stata version control
clear all            // clear working memory
macro drop _all      // clear macros

* Working directory
global wdir "[your/directory]"

* Define paths to subdirectories
global data            "$wdir\0_data"            // folder for original data
global code            "$wdir\1_dofiles"         // folder for do-files
global posted          "$wdir\2_posted"          // adjusted data
global temp            "$wdir\3_temp"            // folder for temporary files
global table           "$wdir\4_tables"          // folder for table output
global graph           "$wdir\5_graphs"          // folder for graph output

*install necessary ado files
            *ssc install estout

/*----------------------------------------------------------------------------
                2. Descriptive/bivariate tables
------------------------------------------------------------------------------*/
*load the prepared dataset
            use "$posted/w6_restricted", clear

*define globals for the lists of variables with recoded missing values

            global catvars99 mig_99 gendercat frailtype health_99          ///
```

```
            drink73_99 smokenow_99 empl_99 edu_99                                    ///
            marital_99 socialclub_99 netempty_99                             ///
            citizen_99 hdi_99 nesw mipex

    global catvarsm mig_m gendercat frailtype health_m                   ///
            drink73_m smokenow_m empl_m edu_m                                        ///
            marital_m socialclub_m netempty_m                                        ///
            citizen_m hdi_m nesw mipex

    global contvars age stay hnetweq thinc2eq //includes imputed

    global contvarsm age_m stay hnetweq_m thinc2eq_m //excludes imputed



*One table with descriptive statistics by migration and gender groups
    collect clear
    table ( mig_99 ) (gendercat) () [pw=weightm],                        ///
            statistic(mean $contvars)                                                ///
            statistic(sd $contvars)                                                  ///
            statistic(fvpercent $catvars99)                                  ///
            statistic(fvfrequency $catvars99)

    collect dims

    collect label list colname, all //to find the names of the labels to rename them
    collect label list result, all
    collect label list statcmd, all
    collect label list mig_m, all //.m for the total
    collect label list var, all

    *styling
            collect style cell result, nformat(%11.1fc)
            collect style cell result[fvfrequency N min max], nformat(%11.0fc)
            collect style cell result[fvpercent], sformat("%s%%")
            collect style cell result[sd fvfrequency], sformat("(%s)")
                            //put some results in brackets

            /*relabel some of the results as well as migrant status
             and remove the superfluous info from the header. */
                    collect label levels mig_99 . "Missing" 0 "Non-migrant" 1 "Migrant", modify
                    collect label levels result fvfrequency "N" fvpercent "Percent (N)" ///
                    mean "Mean (SD)" sd "SD" min "Minimum" max "Maximum", modify

            *modifying the header
                    collect style header mig_99, title(hide) level(label)
                    collect style header gendercat, title(hide) level(label)
                    collect style header result, level(hide)

            // STACK THE ROW LEVEL LABELS
                    collect style row stack, nobinder spacer
            // REMOVE THE VERTICAL LINE
                    collect style cell border_block, border(right, pattern(nil))

    *rearranging
            collect layout (var#result) (mig_99[0 1]#gendercat[.m] gendercat#mig_99[.m])

            collect preview

            collect style putdocx,  ///
```

```
		title("Table x. Summary statistics by gender and migration status") ///
		note("Source: SHARE, wave 6, release 8.0.0, own calculations. Means (SD) and percent (N). Respondents aged 60+, weighted, imputed values included. Note: Differences in N between subgroups and total are due to weighting and rounding error and the omitted 'Missing migrant information' column.")

		*export
		collect export $table\desc\desc_miggender , as(docx) replace

*Summary statistics by frailty states

	global catvars99_nonfrail mig_99 gendercat health_99		///
		drink73_99 smokenow_99 empl_99 edu_99				///
		marital_99 socialclub_99 netempty_99			///
		citizen_99 hdi_99 nesw mipex

		collect clear
	table ( frailtype ) () () [pw=weightm],		///
		statistic(mean $contvars)				///
		statistic(sd $contvars)				///
		statistic(fvpercent $catvars99_nonfrail)			///
		statistic(fvfrequency $catvars99_nonfrail)

	collect dims

	collect label list colname, all //to find the names of the labels to rename them*/
	collect label list result, all
	collect label list statcmd, all
	collect label list mig_m, all //.m for the total
	collect label list var, all

	*styling
		collect style cell result, nformat(%11.1fc)
		collect style cell result[fvfrequency N min max], nformat(%11.0fc)
		collect style cell result[fvpercent], sformat("%s%%")
		collect style cell result[sd fvfrequency], sformat("(%s)")
				//put some results in brackets

		/*relabel some of the results as well as migrant status
		 and remove the superfluous info from the header. */
			collect label levels mig_99 . "Missing" 0 "Non-migrant" 1 "Migrant", modify
			collect label levels result fvfrequency "N" fvpercent "Percent (N)" ///
			mean "Mean (SD)" sd "SD" min "Minimum" max "Maximum", modify

		*modifying the header
			collect style header mig_m, title(hide) level(label)
			collect style header gendercat, title(hide) level(label)
			collect style header result, level(hide)

		// STACK THE ROW LEVEL LABELS
			collect style row stack, nobinder spacer
		// REMOVE THE VERTICAL LINE
			collect style cell border_block, border(right, pattern(nil))

	*rearranging
		collect layout (var#result) (frailtype)

		collect preview

		collect style putdocx,  ///
			title("Table x. Summary statistics by frailty state") ///
```

```
                    note("Source: SHARE, wave 6, release 8.0.0, own calculations. Means (SD) and
percent (N). Respondents aged 60+, weighted, imputed values included. Note: Differences in N between sub-
groups and total are due to weighting and rounding error and the omitted 'Missing migrant information' col-
umn.")

            *export
            collect export $table\desc\desc_byfrailty , as(docx) replace


*Percentage of migrants by country
    *weighted
     *to check what the results should be:
            tab country mig_99, m row

                svyset mergeid [pw=weightm]
            svy: tab country mig_99, obs count row percent

     *Laid out table
            eststo clear
            eststo migcountry_w: estpost svy: tabulate country mig_99, row percent

            esttab migcountry_w using $table\desc\migcountry_w.rtf,                    ///
            main(b %9.2f) aux(obs %9.0f)
                ///
            title("{\b Table X.} Percentage of (non-)migrants across countries, weighted") ///
            mtitle("")
                                    ///
            label eqlabels("Non-migrant" "Migrant" "Missing migrant information")      ///
            noobs nostar nonumbers nonotes replace nogaps onecell unstack              ///
            addnotes("Source: SHARE, wave 6, release 8.0.0, own calculations."  ///
                    "Respondents aged 60+, weighted, imputed values included." ///
                    "Number of (unweighted) observations in parentheses.")

    *unweighted

    svyset mergeid //unweighted

    eststo migcountry_unw: estpost svy: tabulate country mig_99 , row percent

    tabulate country mig_99, row

            esttab migcountry_unw using $table\desc\migcountry_unw.rtf,
    ///
            main(b %9.2f) aux(obs %9.0f)
                ///
            title("{\b Table X.} Percentage of (non-)migrants across countries") ///
            mtitle("")
                                    ///
            label eqlabels("Non-migrant" "Migrant" "Missing migrant information")      ///
            noobs nostar nonumbers nonotes replace nogaps onecell unstack              ///
            addnotes("Source: SHARE, wave 6, release 8.0.0, own calculations."  ///
                    "Respondents aged 60+, unweighted, imputed values included."       ///
                    "Number of observations in parentheses.")



/*--------------------------------------------------------------------------
        3. Bivariate graphs
--------------------------------------------------------------------------*/
```

* frailty by migration

```
graph bar (percent) if inlist(mig_m, 0, 1) [pweight=cciw_w6],            ///
by(mig_m) over(frailtype, relabel(1 "Robust" 2 "Prefrail" 3 "Frail")) ///
yscale(range(0(5)55)) ytitle("Percent")        ///
blabel(bar, format(%9.1f) size(medium)) nolabel name(bar_migfrail, replace)

graph export "$graph/bar_migfrail.png", replace
```

*frailty by gender

```
graph bar (percent) if inlist(mig, 0, 1) [pweight=cciw_w6],    ///
by(gender) over(frailtype, relabel(1 "Robust" 2 "Prefrail" 3 "Frail")) ///
yscale(range(0(5)55)) ytitle("Percent")        ///
blabel(bar, format(%9.1f) size(medium)) nolabel name(bar_genderfrail, replace)

graph export "$graph/bar_genderfrail.png", replace
```

*frailty by age

*as a stacked histogram

```
twoway__histogram_gen age if frailtype == 0, ///
gen(freq_robust x, replace) freq start(0) width(1)
twoway__histogram_gen age if frailtype == 1, ///
gen(freq_prefrail x2, replace) freq start(0) width(1)
twoway__histogram_gen age if frailtype == 2, ///
gen(freq_frail x3, replace) freq start(0) width(1)

qui gen freqrp = freq_robust + freq_prefrail
qui gen freqrpf = freq_robust + freq_prefrail + freq_frail

twoway bar freq_robust x if freqrpf < ., barw(1) bcolor(navy) xsc(r(60 .)) || ///
rbar freq_robust freqrp x2 if freqrpf < ., barw(1) bcolor(lavender) ||     ///
rbar freqrp freqrpf x3 if freqrpf < ., barw(1) bcolor(cranberry) ///
legend(order(1 "Robust" 2 "Prefrail" 3 "Frail") pos(2) col(1) ring(0)) ///
xtitle("Age") ytitle("Number of respondents") xla(60(5)105) yla(, ang(h))

graph export "$graph/hist_agefrail.png", replace
```

*as an overlayed histogram

```
twoway  (histogram age if frailtype == 1,     ///
        start(60) width(1) freq color(lavender%70))            ///
        (histogram age if frailtype == 0,     ///
        start(60) width(1) freq color(navy%30))                ///
        (histogram age if frailtype == 2,     ///
        start(60) width(1) freq color(cranberry%50)),                 ///
        legend(order(1 "Robust" 2 "Prefrail" 3 "Frail")         ///
        col(1) bplacement(neast) ring(0))                      ///
        ytitle("Number of respondents") name(hist_frailage, replace)

        graph export "$graph/hist_frailage.png", replace
```

*age by gender

*as an overlayed histogram

```
twoway  (histogram age if gender==2,        ///
        start(60) width(1) freq color(navy%50))                ///
        (histogram age if gender==1,        ///
        start(60) width(1) freq color(green%40)),              ///
```

```
legend(order(1 "Women" 2 "Men")                              ///
col(1) bplacement(neast) ring(0))                            ///
ytitle("Number of respondents") name(hist_genderage, replace)

graph export "$graph/hist_genderage.png", replace
```

Multivariate analyses

```
/*============================================================================
File name:      Multivariate
Task:                   Estimate multivariate models
Author:                 Carlotta Bochert
============================================================================*/

/*-----------------------------------------------------------------------------
                0. Purpose, preparations and ToC
-----------------------------------------------------------------------------*/
/* This do-file analyses the previously prepared, coded, and described data
 multivariately. After preparing the do-file in section 1, in section 2, data
 are analysed. First, in section 2.1, logistic regression models are estimated
 using the mlogit command. In section two, this is supplemented with some
 sensitivity checks and mediation analysis through the KHB command. In section
 2.3, multi-level models are estimated, and section 2.4 concludes with model
 diagnostic and robustness checks.

To run this do-file, please:
        - adapt the working directory (section 1)
        - set up a folder structure as referred to in section 1

Table of Contents:
        0. Purpose and ToC
        1. Configuring the dofile
        2. Data analysis
                2.1 Fixed effects models (mlogit)
                2.2 KHB analysis
                2.3 Multi-level models
                2.4 Diagnostics and robustness checks

*/

/*-----------------------------------------------------------------------------
                1. Configuring the dofile
-----------------------------------------------------------------------------*/

* Setup
version 17.0        // Stata version control
clear all           // clear working memory
macro drop _all     // clear macros

* Working directory
global wdir “[your/directory]”

* Define paths to subdirectories
global data         "$wdir/0_data"         // folder for original data
global code         "$wdir/1_dofiles"      // folder for do-files
global posted       "$wdir/2_posted"       // data ready for analysis
global temp         "$wdir/3_temp"                 // folder for temporary files
global table        "$wdir/4_tables"       // folder for table output
global graph        "$wdir/5_graphs"               // folder for graph output
```

```
*ssc install boxtid
*ssc install estout
*ssc install khb  //install necessary ado files

/*---------------------------------------------------------------------------
                 2. Data analysis
---------------------------------------------------------------------------*/
*use prepared data for wave 6 (including imported values from earlier waves and imputations)
            use "$posted/w6_restricted", clear


/*---------------------------------------------------------------------------
                 2.1 Fixed effects models (mlogit)
---------------------------------------------------------------------------*/

*Models 1 and 2
            eststo clear


            *Model 1
                        eststo m1: mlogit frailtype i.mig_99 i.gendercat        ///
                        c.age i.health_99 ib(12).country, base(0)

            *Model 2
                        eststo m2: mlogit frailtype i.mig_99##i.gendercat     ///
                        c.age i.health_99 ib(12).country, base(0)


            *tables
            *Full list of coefficients
                        esttab m1 m2 using $table\mlogit\mlogit_12full.rtf , replace              ///
                        eform noconstant label compress nobase nonumbers                                   ///
                        b(%9.2fc) p(%9.3fc) par   onecell unstack
            ///
                        coeflabels(1.mig_99 Migrant 99.mig_99 "Missing"
            ///
                        age "Age" 1.mig_99#1.gendercat "Female*migrant"                                  ///
                        99.mig_99#1.gendercat "Female*missing migration info"                           ///
                        1.gendercat "Female"
                                                ///
                        99.health_99 "Missing")
                                    ///
                        drop(Robust: _cons)
                                                            ///
                        refcat(1.mig_99 "Migrant status (ref. non-migrant)"                         ///
                        11.country "Country (ref. Germany)"
                        ///
                        1.health_99 "Health problems in 1-3 dimensions (ref. 0)", nolabel)       ///
                        order(1.mig_99 99.mig_99 1.gendercat 1.mig_99#1.gendercat                       ///
                        99.mig_99#1.gendercat )
                                    ///
                        mtitles("Base model" "Base model with interaction")                                ///
                        title("{\b Table x.} Fixed effects regression models with and without interaction of gender
and migrant status")
                        ///
                        addnote("Source: SHARE, wave 6, release 8.0.0, own calculations."    ///
                        "Respondents aged 60+, unweighted, imputed values included.")
```

```
*Excluding control coefficients and missing categories
        esttab m1 m2 using $table\mlogit\mlogit_12min.rtf , replace                ///
        eform noconstant label compress nobase nonumbers                            ///
        b(%9.2fc) p(%9.3fc) par   onecell unstack
///
        coeflabels(1.mig_99 "Migrant"
            ///
        1.mig_99#1.gendercat "Female*migrant"
        ///
        1.gendercat "Female")
                ///
        drop(*.health_99 age *.country 99.mig_99 99.mig_99#1.gendercat Robust: _cons)      ///
        order(1.mig_99 1.gendercat 1.mig_99#1.gendercat )                              ///
        mtitles("Base model" "Base model with interaction")                            ///
        title("{\b Table x.} Fixed effects regression models with and without interaction of gender
and migrant status")
        ///
        addnote("Source: SHARE, wave 6, release 8.0.0, own calculations."    ///
        "All models controlling for missing migrant information, age, health problems, and countries
(coefficients not shown)."
///
        "Respondents aged 60+, unweighted, imputed values included.")

*Mediation models 3a1 to 3b5
    *Mediation models separately for men and women
        eststo clear

        *loop by gender, generating 5 sets of estimates each
        foreach g in 0 1 {
            eststo: mlogit frailtype i.mig_99 c.age i.health_99 ib(12).country              ///
            if gendercat== `g', base(0)

                *health behaviours
            eststo: mlogit frailtype i.mig_99 c.age i.health_99 ib(12).country              ///
             i.smokenow_99 i.drink73_99 if gendercat== `g', base(0)

                *Economic factors
            eststo: mlogit frailtype i.mig_99 c.age i.health_99 ib(12).country              ///
            i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq if gendercat== `g', base(0)

                *Social network
            eststo: mlogit frailtype i.mig_99 c.age i.health_99 ib(12).country              ///
            i.marital_99 i.socialclub_99 i.netempty_99 if gendercat== `g', base(0)

                *all mediators
            eststo: mlogit frailtype i.mig_99 c.age i.health_99 ib(12).country        ///
            i.smokenow_99 i.drink73_99
                    ///
            i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
            ///
            i.marital_99 i.socialclub_99 i.netempty_99
        ///
            if gendercat== `g', base(0)
        }


        *tables for men

            *Full list of coefficients
```

```
esttab est1 est2 est3 est4 est5 ///
using $table\mlogit\mlogit_menmedfull.rtf , replace ///
eform  noconstant label compress nobase     nonumbers ///
b(%9.2fc) p(%9.3fc) par   onecell ///
coeflabels(1.mig_99 Migrant 99.mig_99 "Missing") ///
drop(Robust: _cons) ///
refcat(1.mig_99 "Migrant status (ref. non-migrant)" ///
11.country "Country (ref. Germany)" ///
1.health_99 "Health problems in 1-3 dimensions (ref. 0)" ///
1.smokenow_99 "Smoking (ref. non-smoker)" ///
1.drink73_99 "Alcoholic drinks (ref. none)" ///
1.edu_99 "Education (ref. primary)" ///
2.empl_99 "Employment status (ref. retired)" ///
2.marital_99 "Marital status (ref. married)" ///
1.netempty_99 "No social network (ref. has a social network)" ///
1.socialclub_99 "Attended social clubs in past year (ref. no)", nolabel) ///
mtitles("No mediators" "Health behaviours" "Economic factors" ///
 "Social capital" "All mediators") ///
title("{\b Table x.} Fixed effects regression models including different groups of mediators (men)") ///
addnote("Source: SHARE, wave 6, release 8.0.0, own calculations." ///
"Respondents aged 60+, unweighted, imputed values included.")


*Minimum table for men
esttab est1 est2 est3 est4 est5 ///
using $table\mlogit\mlogit_menmed_min.rtf , replace ///
eform  noconstant label compress nobase     nonumbers ///
b(%9.2fc) p(%9.3fc) par   onecell ///
coeflabels(1.mig_99 Migrant ) ///
keep(Pre_frail:1.mig_99 Frail:1.mig_99) ///
mtitles("No mediators" "Health behaviours" "Economic factors" ///
 "Social capital" "All mediators") ///
title("{\b Table x.} Fixed effects regression models including different groups of mediators (men)") ///
addnote("Source: SHARE, wave 6, release 8.0.0, own calculations." ///
"All models controlling for missing migrant information, age, health problems, and countries in addition to the mediators as specified in the model titles (coefficients not shown)." ///
"Respondents aged 60+, unweighted, imputed values included.")
```

```
*tables for women
	*Full list of coefficients
		esttab est6 est7 est8 est9 est10
///
		using $table\mlogit\mlogit_womenmedfull.rtf , replace		///
		eform  noconstant label compress nobase     nonumbers
///
		b(%9.2fc) p(%9.3fc) par   onecell
///
		coeflabels(1.mig_99 Migrant 99.mig_99 "Missing")			///
		drop(Robust: _cons)
			///
		refcat(1.mig_99 "Migrant status (ref. non-migrant)"		///
		11.country "Country (ref. Germany)"
///
		1.health_99 "Health problems in 1-3 dimensions (ref. 0)"	///
		1.smokenow_99 "Smoking (ref. non-smoker)"
///
		1.drink73_99 "Alcoholic drinks (ref. none)"
///
		1.edu_99 "Education (ref. primary)"
	///
		2.empl_99 "Employment status (ref. retired)"
///
		2.marital_99 "Marital status (ref. married)"				///
		1.netempty_99 "No social network (ref. has a social network)"	///
		1.socialclub_99 "Attended social clubs in past year (ref. no)", nolabel) ///
		mtitles("No mediators" "Health behaviours" "Economic factors" ///
		 "Social capital" "All mediators")
///
		title("{\b Table x.} Fixed effects regression models including different
groups of mediators (women)")
	///
		addnote("Source: SHARE, wave 6, release 8.0.0, own calculations."	///
		"Respondents aged 60+, unweighted, imputed values included.")
		//evtl health problem dims noch labeln wie oben


	*Minimum table for women
		esttab est6 est7 est8 est9 est10
///
		using $table\mlogit\mlogit_womenmed_min.rtf , replace		///
		eform  noconstant label compress nobase     nonumbers
///
		b(%9.2fc) p(%9.3fc) par   onecell
///
		coeflabels(1.mig_99 Migrant )
	///
		keep(Pre_frail:1.mig_99 Frail:1.mig_99)
///
		mtitles("No mediators" "Health behaviours" "Economic factors" ///
		 "Social capital" "All mediators")
///
		title("{\b Table x.} Fixed effects regression models including different
groups of mediators (women)")
	///
		addnote("Source: SHARE, wave 6, release 8.0.0, own calculations."	///
		"All models controlling for missing migrant information, age, health prob-
lems, and countries in addition to the mediators as specified in the model titles (coefficients not shown)." ///
```

```
                        "Respondents aged 60+, unweighted, imputed values included.")

*margins

        *men
                mlogit frailtype i.mig_99 c.age i.health_99 ib(12).country             ///
                        i.smokenow_99 i.drink73_99
                        ///
                        i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
        ///
                        i.marital_99 i.socialclub_99 i.netempty_99
///
                        if gendercat== 0, base(0)

                margins mig_99 if mig_99 == 0 | mig_99 == 1, at(age=(60(10)90))

                marginsplot, name(margins_3a5, replace) plotopts(msize(small))       ///
                title("Men's probability of frailty states by age and migration experience") ///
                subtitle("Including all controls and mediators", size(small))   ///
                ytitle("Predicted probability", size(small))                              ///
                ylabel(0(.1).6)
                            ///
                note("Source: SHARE, wave 6, release 8.0.0, own calculations")       ///
                legend(order(- "Robust:" 1 4 - "Pre-frail:" 2 5 - "Frail:" 3 6) ///
                cols(3) size(small)
                            ///
                label(1 "Non-migrant") label(4 "Migrant")
///
                label(2 "Non-migrant") label(5 "Migrant")
///
                label(3 "Non-migrant") label(6 "Migrant"))

                *export
                        graph export "$graph/margins_3a5.png", replace


        *women
                mlogit frailtype i.mig_99 c.age i.health_99 ib(12).country             ///
                        i.smokenow_99 i.drink73_99
                        ///
                        i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
        ///
                        i.marital_99 i.socialclub_99 i.netempty_99
///
                        if gendercat== 1, base(0)

                margins mig_99 if mig_99 == 0 | mig_99 == 1, at(age=(60(10)90))

                marginsplot, name(margins_3b5, replace)
///
                plotopts(msize(small))
                        ///
                title("Women's probability of frailty states by age and migration experience") ///
                subtitle("Including all controls and mediators", size(small))   ///
                ytitle("Predicted probability", size(small))                              ///
                ylabel(0(.1).6)
                            ///
                note("Source: SHARE, wave 6, release 8.0.0, own calculations")       ///
                legend(order(- "Robust:" 1 4 - "Pre-frail:" 2 5 - "Frail:" 3 6) ///
```

```
                    cols(3) size(small)
                            ///
                    label(1 "Non-migrant") label(4 "Migrant")
    ///
                    label(2 "Non-migrant") label(5 "Migrant")
    ///
                    label(3 "Non-migrant") label(6 "Migrant"))

                    *export
                            graph export "$graph/margins_3b5.png", replace



*Migrant-only models comparing men and women
        *all mediators

                *loop by gender
                foreach g in 0 1 {
                    eststo mig_`g': mlogit frailtype
                ///
                    c.age i.health_99 ib(12).country
                ///
                    c.stay i.hdi_99     i.imcitizen_99
                    ///
                    i.smokenow_99 i.drink73_99
                            ///
                    i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
                ///
                    i.marital_99 i.socialclub_99 i.netempty_99
    ///
                    if mig_99 == 1 & gendercat== `g', base(0)
                    }


                    *With control coefficients
                            esttab mig_0 mig_1 ///
                            using $table\mlogit\mlogit_migrantfull.rtf , replace            ///
                            eform  noconstant label compress nobase     nonumbers
    ///
                            b(%9.2fc) p(%9.3fc) par   onecell unstack
    ///
                            coeflabels(1.mig_99 Migrant 99.mig_99 "Missing")                    ///
                            drop(Robust: _cons)
                                    ///
                            refcat(11.country "Country (ref. Germany)"
                ///
                            1.health_99 "Health problems in 1-3 dimensions (ref. 0)"      ///
                            1.smokenow_99 "Smoking (ref. non-smoker)"
                ///
                            1.drink73_99 "Alcoholic drinks (ref. none)"
    ///
                            1.edu_99 "Education (ref. primary)"
                    ///
                            2.empl_99 "Employment status (ref. retired)"
    ///
                            2.marital_99 "Marital status (ref. married)"                          ///
                            1.netempty_99 "No social network (ref. has a social network)" ///
                            1.socialclub_99 "Attended social clubs in past year (ref. no)" ///
                            2.hdi_99 "HDI level of origin country (ref. highest)"                ///
                            1.imcitizen_99 "Citizenship in country of interview (ref. no)", nolabel) ///
```

```
mtitles("Men" "Women")
///
title("{\b Table x.} Migrant-only fixed effects regression models including
migration-specific variables")                                    ///
addnotes("Source: SHARE, wave 6, release 8.0.0, own calculations" ///
"Respondents aged 60+, unweighted, imputed values included.")


*Minimum table
esttab mig_0 mig_1
///
using $table\mlogit\mlogit_migrantmin.rtf , replace                    ///
eform noconstant label compress nobase nonumbers                       ///
b(%9.2fc) p(%9.3fc) par   onecell unstack
///
coeflabels(1.mig_99 Migrant )
///
keep(Pre_frail:stay Pre_frail:2.hdi_99 Pre_frail:3.hdi_99       ///
Pre_frail:1.imcitizen_99 Frail:stay Frail:2.hdi_99                    ///
Frail:3.hdi_99 Frail:1.imcitizen_99)
///
refcat(2.hdi_99 "HDI level of origin country (ref. highest)" ///
1.imcitizen_99 "Citizenship in country of interview (ref. no)", nolabel) ///
mtitles("Men" "Women")
///
title("{\b Table x.} Migrant-only fixed effects regression models including
migration-specific variables")
///
addnote("Source: SHARE, wave 6, release 8.0.0, own calculations."    ///
"All models controlling for age, health problems, and countries as well as
all mediators and missing categories for HDI level and citizenship (coefficients not shown)." ///
"Respondents aged 60+, unweighted, imputed values included.")



*Models without country fixed effects, instead mipex and region
eststo clear


*Country-group interactions with migrant status

foreach g in 0 1 {
eststo country_`g': mlogit frailtype
///
i.mig_99##ib(4).nesw i.mig_99##i.mipex
///
c.age i.health_99
///
i.smokenow_99 i.drink73_99
///
i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
///
i.marital_99 i.socialclub_99 i.netempty_99
///
if gendercat== `g', base(0)
}

*Table with full list of coefficients
esttab country_0 country_1
///
```

```
using $table\mlogit\mlogit_c_full.rtf , replace ///
eform  noconstant label compress nobase     nonumbers ///
b(%9.2fc) p(%9.3fc) par   onecell
///
coeflabels(1.mig_99 Migrant 99.mig_99 "Missing" ///
1.mig_99#1.nesw "North*migrant"
///
1.mig_99#2.nesw "East*migrant"
///
1.mig_99#3.nesw "South*migrant"
///
1.mig_99#2.mipex "Med.inclusive*migrant"
///
1.mig_99#3.mipex "Low inclusive*migrant")
///
drop(Robust: _cons) unstack
///
refcat(1.mig_99 "Migrant status (ref. non-migrant)" ///
1.health_99 "Health problems in 1-3 dimensions (ref. 0)" ///
1.smokenow_99 "Smoking (ref. non-smoker)"
///
1.drink73_99 "Alcoholic drinks (ref. none)" ///
1.edu_99 "Education (ref. primary)"
///
2.empl_99 "Employment status (ref. retired)" ///
2.marital_99 "Marital status (ref. married)" ///
1.netempty_99 "No social network (ref. has a social network)" ///
1.socialclub_99 "Attended social clubs in past year (ref. no)" ///
1.nesw "Region (ref. West)"
///
2.mipex "Health care inclusivity (ref. high inclusivity)", nolabel) ///
mtitles("Men" "Women" )
///
title("{\b Table x.} Regression models with country group interactions (no fixed ef-
fects)") ///
addnotes("Source: SHARE, wave 6, release 8.0.0, own calculations."  ///
        "Respondents aged 60+, unweighted, imputed values included.")

*table of just the main effects
esttab country_0 country_1
        ///
using $table\mlogit\mlogit_c_min.rtf , replace
///
eform noconstant label compress nobase nonumbers ///
b(%9.2fc) p(%9.3fc) par   onecell unstack
///
coeflabels(1.mig_99 "Migrant"
///
1.mig_99#1.nesw "North*migrant"
///
1.mig_99#2.nesw "East*migrant"
///
1.mig_99#3.nesw "South*migrant"
///
1.mig_99#2.mipex "Med.inclusive*migrant"
///
1.mig_99#3.mipex "Low inclusive*migrant")
///
drop(Robust: _cons 99.mig_99* *.health_99
///
```

```
				age *.smokenow_99 *.drink73_99 *.marital_99 *.socialclub_99 *.netempty_99 ///
				log_thinc2eq c3_hnetweq *.edu_99 *.empl_99 )
	///
				refcat(1.nesw "Geographical region (ref. West)"
	///
				2.mipex "Health care inclusivity (ref. high inclusivity)", nolabel)	///
				order(1.mig_99 *.nesw *.mipex)
				///
				mtitles("Men" "Women")
					///
				title("{\b Table x.} Regression models with country group interactions (no fixed ef-
fects)")		///
				addnotes("Source: SHARE, wave 6, release 8.0.0, own calculations."	///
						"All models controlling for missing migrant information, age,
health problems, and all mediators (coefficients not shown)."		///
						"Respondents aged 60+, unweighted, imputed values included.")



/*--------------------------------------------------------------------------
		2.2 KHB analysis
----------------------------------------------------------------------------*/
*in preparation of khb command needs: shorten varnames
	gen h99 = health_99
	gen wm = weightm

*Sensitivity to different variable specifications and weighting

	*Looking at the difference between models with weights
		khb mlogit frailtype i.mig_99
				///
				|| c.wm c.wm#i.h99 c.wm#i.country c.wm#c.age,
	///
				concomitant(c.age i.h99 ib(12).country)
	///
				base(0) summary disentangle outcome(1)
		//the difference made by including the weight interaction is not significant

		khb mlogit frailtype i.mig_99
				///
				|| c.wm c.wm#i.h99 c.wm#i.country c.wm#c.age,
	///
				concomitant(c.age i.h99 ib(12).country)
	///
				base(0) summary disentangle outcome(2)
		//the difference made by including the weight interaction is not significant



	*Looking at the difference between models with differently specified cont vars
		*age
				khb mlogit frailtype i.mig_99
				///
								|| c.age#c.age,
					///
								concomitant(c.age   i.health_99   ib(12).country)
	///
								base(0) summary disentangle outcome(1)

					//the difference made by including the squared term is not significant
```

```
khb mlogit frailtype i.mig_99
///
                                        || c.age#c.age,
          ///
                                        concomitant(c.age   i.health_99   ib(12).country)
///
                                        base(0) summary disentangle outcome(2)

          //the difference made by including the squared term is not significant


     *income
          khb mlogit frailtype i.mig_99
          ///
                    || c.thinc2eq,
                    ///
                    concomitant(c.age i.health_99 ib(12).country)
///
                    base(0) summary disentangle outcome(1)

          khb mlogit frailtype i.mig_99
          ///
                    || c.log_thinc2eq,
          ///
                    concomitant(c.age i.health_99 ib(12).country)
///
                    base(0) summary disentangle outcome(1)


     *wealth
          khb mlogit frailtype i.mig_99
          ///
                    || c.hnetweq,
                    ///
                    concomitant(c.age i.health_99 ib(12).country)
///
                    base(0) summary disentangle outcome(1)

          khb mlogit frailtype i.mig_99
          ///
                    || c.c3_hnetweq,
          ///
                    concomitant(c.age i.health_99 ib(12).country)
///
                    base(0) summary disentangle outcome(1)



*Mediation models
     eststo clear

     *loop over gender (0 female 1 male) and outcome (1 pre-frail, 2 frail)

          foreach g in 0 1 {
               foreach o in 1 2 {

                    *health behaviours
               eststo hb_`g'_`o': khb mlogit frailtype i.mig_99 ||                ///
```

```
            i.smokenow_99 i.drink73_99 if gendercat== `g',
///
            concomitant(c.age i.health_99 ib(12).country)
///
            base(0) or outcome(`o') summary

                *Economic factors
            eststo econ_`g'_`o': khb mlogit frailtype i.mig_99 ||                ///
            i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq if gendercat== `g',       ///
            concomitant(c.age i.health_99 ib(12).country)
///
            base(0) or outcome(`o') summary

                *Social network
            eststo soc_`g'_`o': khb mlogit frailtype i.mig_99 ||                 ///
            i.marital_99 i.socialclub_99 i.netempty_99 if gendercat== `g',       ///
            concomitant(c.age i.health_99 ib(12).country)
///
            base(0) or outcome(`o') summary

                *all mediators
            eststo full_`g'_`o': khb mlogit frailtype i.mig_99 ||                ///
            i.smokenow_99 i.drink73_99
                ///
            i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
        ///
            i.marital_99 i.socialclub_99 i.netempty_99 if gendercat== `g' ///
            , concomitant(c.age i.health_99 ib(12).country)                          ///
            base(0) or outcome(`o') summary
            }
        }

        *Export results into tables
            *Men pre-frail
            esttab hb_0_1 econ_0_1 soc_0_1 full_0_1
        ///
            using "$table\khb\khb_pf_men.rtf", replace
///
            transform(e*: exp(@) exp(@) exp(@))
        ///
            nogaps noabbrev nobaselevels noobs nonote
        ///
            nonumbers nopar drop(0b.mig_99: 99.mig_99:) collabels(none)              ///
            cells(b(star fmt(2)) p(fmt(3) par))
        ///
            mtitles("Health behaviours" "Economic factors"
///
                        "Social capital" "All mediators")
        ///
            varlabels(Full "Direct effect" Reduced "Total effect"                ///
                                Diff "Indirect effect" )
            ///
            eqlabels("Migrant men")
            ///
            title("{\b Table x.} Adjusted Odds Ratios for pre-frailty in male migrants (KHB
Method)")                       ///
            s(pct_1_mig_99, label("Percentage mediated"))
///
            addnote("Exponentiated coefficients; {\i p}-values in parentheses" ///
            "Source: SHARE, wave 6, release 8.0.0, own calculations."            ///
```

```
"All models controlling for missing migrant information, age, health problems, and countries in addition to the mediators as specified in the model titles (coefficients not shown)." ///
"Respondents aged 60+, unweighted, imputed values included."          ///
"* {\i p}<0.05, ** {\i p}<0.01, *** {\i p}<0.001")


*Men frail
esttab hb_0_2 econ_0_2 soc_0_2 full_0_2
///
using "$table\khb\khb_f_men.rtf", replace
///
transform(e*: exp(@) exp(@) exp(@))
///
nogaps noabbrev nobaselevels noobs nonote
///
nonumbers nopar drop(0b.mig_99: 99.mig_99:) collabels(none)                    ///
cells(b(star fmt(2)) p(fmt(3) par))
///
mtitles("Health behaviours" "Economic factors"
///
                "Social capital" "All mediators")
///
varlabels(Full "Direct effect" Reduced "Total effect"                    ///
                        Diff "Indirect effect" )
///
eqlabels("Migrant men" )
///
title("{\b Table x.} Adjusted Odds Ratios for frailty in male migrants (KHB Method)")                              ///
s(pct_1_mig_99, label("Percentage mediated"))
///
addnote("Exponentiated coefficients; {\i p}-values in parentheses" ///
"Source: SHARE, wave 6, release 8.0.0, own calculations."              ///
"All models controlling for missing migrant information, age, health problems, and countries in addition to the mediators as specified in the model titles (coefficients not shown)." ///
"Respondents aged 60+, unweighted, imputed values included."          ///
"* {\i p}<0.05, ** {\i p}<0.01, *** {\i p}<0.001")


*Women pre-frail
esttab hb_1_1 econ_1_1 soc_1_1 full_1_1
///
using "$table\khb\khb_pf_women.rtf", replace
///
transform(e*: exp(@) exp(@) exp(@))
///
nogaps noabbrev nobaselevels noobs nonote
///
nonumbers nopar drop(0b.mig_99: 99.mig_99:) collabels(none)                    ///
cells(b(star fmt(2)) p(fmt(3) par))
///
mtitles("Health behaviours" "Economic factors"
///
                "Social capital" "All mediators")
///
varlabels(Full "Direct effect" Reduced "Total effect"                    ///
                        Diff "Indirect effect" )
///
eqlabels("Migrant women")
    ///
```

```
title("{\b Table x.} Adjusted Odds Ratios for pre-frailty in female migrants (KHB Method)")                                   ///
s(pct_1_mig_99, label("Percentage mediated"))        ///
addnote("Exponentiated coefficients; {\i p}-values in parentheses" ///
"Source: SHARE, wave 6, release 8.0.0, own calculations."              ///
"All models controlling for missing migrant information, age, health problems, and countries in addition to the mediators as specified in the model titles (coefficients not shown)." ///
"Respondents aged 60+, unweighted, imputed values included."          ///
"* {\i p}<0.05, ** {\i p}<0.01, *** {\i p}<0.001")

*Women frail
esttab hb_1_2 econ_1_2 soc_1_2 full_1_2
///
using "$table\khb\khb_f_women.rtf", replace
///
transform(e*: exp(@) exp(@) exp(@))
///
nogaps noabbrev nobaselevels noobs nonote
///
nonumbers nopar drop(0b.mig_99: 99.mig_99:) collabels(none)                    ///
cells(b(star fmt(2)) p(fmt(3) par))
///
mtitles("Health behaviours" "Economic factors"
///
        "Social capital" "All mediators")
///
varlabels(Full "Direct effect" Reduced "Total effect"                    ///
        Diff "Indirect effect" )
///
eqlabels("Migrant women")
///
title("{\b Table x.} Adjusted Odds Ratios for frailty in female migrants (KHB Method)")                                   ///
s(pct_1_mig_99, label("Percentage mediated"))
///
addnote("Exponentiated coefficients; {\i p}-values in parentheses" ///
"Source: SHARE, wave 6, release 8.0.0, own calculations."              ///
"All models controlling for missing migrant information, age, health problems, and countries in addition to the mediators as specified in the model titles (coefficients not shown)." ///
"Respondents aged 60+, unweighted, imputed values included."          ///
"* {\i p}<0.05, ** {\i p}<0.01, *** {\i p}<0.001")


*Estimates of the percentage contributed by variable
    *men pre-frail
    khb mlogit frailtype i.mig_99 ||
    ///
        i.smokenow_99 i.drink73_99
        ///
        i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
    ///
        i.marital_99 i.socialclub_99 i.netempty_99
///
        if gendercat== 1, concomitant(c.age i.health_99 ib(12).country) ///
        base(0) or outcome(1) disentangle

    *look at disentangle-matrix
        mat list e(disentangle)
    /*save it into an actual matrix and keep only the fourth column and the
    rows below the (empty) reference category*/
```

```
		matrix dismpf = e(disentangle)
		matrix dismpf = dismpf[24..46, 4]
		mat list dismpf

	*men frail
	khb mlogit frailtype i.mig_99 ||
	///
		i.smokenow_99 i.drink73_99
			///
		i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
	///
		i.marital_99 i.socialclub_99 i.netempty_99
///
		if gendercat== 1, concomitant(c.age i.health_99 ib(12).country) ///
		base(0) or outcome(2) disentangle

	*look at disentangle-matrix
		mat list e(disentangle)
	*save it into an actual matrix and keep only the relevant rows
		matrix dismf = e(disentangle)
		matrix dismf = dismf[24..46, 4]
		mat list dismf

*women pre-frail
	khb mlogit frailtype i.mig_99 ||
	///
		i.smokenow_99 i.drink73_99
			///
		i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
	///
		i.marital_99 i.socialclub_99 i.netempty_99
///
		if gendercat== 0, concomitant(c.age i.health_99 ib(12).country) ///
		base(0) or outcome(1) disentangle

	*look at disentangle-matrix
		mat list e(disentangle)
	/*save it into an actual matrix and keep only the fourth column and the
	rows below the (empty) reference category*/
		matrix diswpf = e(disentangle)
		matrix diswpf = diswpf[24..46, 4]
		mat list diswpf

	*women frail
	khb mlogit frailtype i.mig_99 ||
	///
		i.smokenow_99 i.drink73_99
			///
		i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
	///
		i.marital_99 i.socialclub_99 i.netempty_99
///
		if gendercat== 0, concomitant(c.age i.health_99 ib(12).country)		///
		base(0) or outcome(2) disentangle

	*look at disentangle-matrix
		*mat list e(disentangle)
	*save it into an actual matrix and keep only the relevant rows
		matrix diswf = e(disentangle)
		matrix diswf = diswf[24..46, 4]
```

```
			mat list diswf

	/*could also do the above as a loop like this:
			foreach g in 0 1 {
				foreach o in 1 2 {

			khb mlogit frailtype i.mig_99 || ///
				i.nesw i.mipex if gendercat== `g',		///
				concomitant(c.age i.health_99 i.smokenow_99 i.drink73_99   ///
					i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
		///
					i.marital_99 i.socialclub_99 i.netempty_99)
	///
				base(0) or outcome(`o') disentangle

			*look at disentangle-matrix
				mat list e(disentangle)
			/*save it into an actual matrix and keep only the fourth column and
			the rows below the (empty) reference category*/
				matrix dism_`g'_`o' = e(disentangle)
				matrix dism_`g'_`o' = dism_`g'_`o'[24..46, 4]
				mat list dism_`g'_`o'
				}
			}
			*and then use those matrix names in the row-naming and joining below.
			foreach m in dism_0_1 dism_0_2 dism_1_1 dism_1_2 etc etc
			*/

	foreach m in dismpf dismf diswpf diswf {
		mat rownames `m' = "Smoking" "Missing smoking info"				///
		"1-2 drinks/week" "3+ drinks/week" "Missing drink info"			///
		"Secondary education" "Post-secondary education"				///
		"Other education" "Missing education info"
		///
		"(Self-)employed" "Unemployed" "Homemaker or sick"
	///
		"Other employment status" "Missing employment info"				///
		"Income" "Wealth"
					///
		"Widowed" "Divorced" "Never married"
		///
		"Attended social club" "Missing social club info"				///
		"Network empty" "Missing network info"
	}

		matrix coljoinbyname dis = dismpf dismf diswpf diswf
		mat list dis

		*and export the estimates into a table
			esttab matrix(dis, fmt(2))
			///
			using "$table/khb/khb_disentangle.rtf", unstack replace		///
			noabbrev nobaselevels noobs nonumbers nonote mlabels("")		///
			coll("Men pre-frail" "Men frail" "Women pre-frail" "Women frail") ///
			nopar eql("Percentage Mediated")
		///
			title("{\b Table x.} Overview of explained percentage of migrant effect, by media-
tors (KHB Method)")	///
			addnote("Source: SHARE, wave 6, release 8.0.0, own calculations." ///
```

```
                              "Based on models controlling for missing migrant information, age, health prob-
lems, countries, and all mediators (coefficients not shown)."                        ///
                              "Respondents aged 60+, unweighted, imputed values included.")


/*---------------------------------------------------------------------------
              2.3 Multi-level models
---------------------------------------------------------------------------*/
*set countries as the higher level
        xtset country

*Model 0: Empty multilevel multinomial logistic regression with countries as level 1

        eststo empty: xtmlogit frailtype, base(0) cov(unstructured)
                *highly significant LR test
                estat ic
                estat sd


*Model 1: Adding migration, gender, and controls

        eststo xtbase1: xtmlogit frailtype i.mig_99 i.gendercat                          ///
                                        c.age i.health_99 , base(0) cov(unstructured)


*Model 2: Adding an interaction between migration and gender
        eststo xtbase2:    xtmlogit frailtype i.mig_99##i.gendercat                      ///
                                        c.age i.health_99 , base(0) cov(unstructured)

        *Coefficient table:
         esttab xtbase1 xtbase2 using $table\xt\xt_12.rtf , replace                       ///
                eform noconstant label compress nobase nonumbers                                 ///
                b(%9.2fc) p(%9.3fc) par   onecell unstack
        ///
                coeflabels(1.mig_99 Migrant 99.mig_99 "Missing"
        ///
                age "Age" 1.mig_99#1.gendercat "Female*migrant"                                  ///
                99.mig_99#1.gendercat "Female*missing migration info"                            ///
                1.gendercat "Female"
                        ///
                99.health_99 "Missing")
                      ///
                drop(_cons Robust: /:)
                        ///
                refcat(1.mig_99 "Migrant status (ref. non-migrant)"                        ///
                1.health_99 "Health problems in 1-3 dimensions (ref. 0)", nolabel)      ///
                order(1.mig_99 99.mig_99 1.gendercat 1.mig_99#1.gendercat                        ///
                99.mig_99#1.gendercat )
                      ///
                mtitles("Base model" "Base model with interaction")                              ///
                addnote("Source: SHARE, wave 6, release 8.0.0, own calculations.")  ///
                title("{\b Table x.} Random intercept regression models with and without interaction of gen-
der and migrant status")


*For ICC estimation: emptpy binomial models
                *only robust and pre-frail
                        melogit frailtype if frailtype != 2 || country:
                                estat icc //ICC 0.031
```

```
*only robust and frail
        melogit frailtype if frailtype != 1 || country:
                estat icc //ICC 0.098

*only pre-frail and frail
        /*(in var "frail", robust and pre-frail people are coded 0;
        robust are left out through the if-clause)*/
        melogit frail if frailtype != 0 || country:
                estat icc //ICC 0.028



/*-------------------------------------------------------------------------
            2.4 Diagnostics and robustness checks
--------------------------------------------------------------------------*/

*Collinearity diagnostics
    /*Since multicollinearity is a feature of the variables, not the models,
    I can run an OLS regression and look at the VIFs there. */

        reg frailtype i.mig_99 i.gendercat
        ///
                i.smokenow_99 i.drink73_99
                        ///
                i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
        ///
                i.marital_99 i.socialclub_99 i.netempty_99
    ///
                c.age i.health_99 ib(12).country
                        eststo vifcoun: estadd vif


    *and again using mipex/region instead of country
        reg frailtype i.mig_99 i.gendercat
        ///
                i.smokenow_99 i.drink73_99
                        ///
                i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
        ///
                i.marital_99 i.socialclub_99 i.netempty_99
    ///
                c.age i.health_99 i.mipex ib(4).nesw
                        eststo vifreg: estadd vif

    *and again for the migrant-only model
        reg frailtype i.gendercat ///
                i.smokenow_99 i.drink73_99
                        ///
                i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
        ///
                i.marital_99 i.socialclub_99 i.netempty_99
    ///
                c.stay i.hdi_99     i.imcitizen_99
                ///
                c.age i.health_99 ib(12).country if mig_99 == 1
                        eststo vifmig: estadd vif

    *and into a table...
        esttab vifcoun vifreg vifmig
                ///
```

```
			using $table\vif.rtf, replace
			///
			noconstant label compress nobase nonumbers noobs				///
			cells(vif(fmt(2)))
						///
			coeflabels(1.mig_99 Migrant 99.mig_99 "Missing")				///
			refcat(1.mig_99 "Migrant status (ref. non-migrant)"			///
			1.health_99 "Health problems in 1-3 dimensions (ref. 0)"		///
			1.smokenow_99 "Smoking (ref. non-smoker)"
		///
			1.drink73_99 "Alcoholic drinks (ref. none)"
	///
			1.edu_99 "Education (ref. primary)"
			///
			2.empl_99 "Employment status (ref. retired)"
	///
			2.marital_99 "Marital status (ref. married)"				///
			1.netempty_99 "No social network (ref. has a social network)"		///
			1.socialclub_99 "Attended social clubs in past year (ref. no)"  ///
			11.country "Country (ref. Germany)"
		///
			1.nesw "Region (ref. West)"
				///
			2.mipex "Health care inclusivity (ref. high inclusivity)"		///
			2.hdi_99 "HDI level of origin country (ref. highest)"			///
			1.imcitizen_99 "Citizenship in country of interview (ref. no)", nolabel) ///
			mtitles("Main model" "Region model" "Migrant model")			///
			collabels("", none)
						///
			title("{\b Table x.} Variance inflation factors for the three full sets of variables")
		///
			addnote("Source: SHARE, wave 6, release 8.0.0, own calculations." ///
			"Respondents aged 60+, unweighted, imputed values included.")


*Robustness checks


	*robustness to weighting the data

		*Men's full mediation model
		eststo men_w: mlogit frailtype i.mig_99 c.age i.health_99 ib(12).country ///
			i.smokenow_99 i.drink73_99
				///
			i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
		///
			i.marital_99 i.socialclub_99 i.netempty_99
	///
			if gendercat== 0 [pweight=weightm], base(0)

		*Women's full mediation model
		eststo women_w: mlogit frailtype i.mig_99 c.age i.health_99 ib(12).country ///
			i.smokenow_99 i.drink73_99
				///
			i.edu_99 i.empl_99 log_thinc2eq c3_hnetweq
		///
			i.marital_99 i.socialclub_99 i.netempty_99
	///
			if gendercat== 1 [pweight=weightm], base(0)
```

```
*Table of coefficients
esttab men_w women_w ///
	using $table\mlogit\mlogit_weighted.rtf , replace	///
	eform  noconstant label compress nobase     nonumbers
///
	b(%9.2fc) p(%9.3fc) par   onecell unstack
///
	coeflabels(1.mig_99 Migrant 99.mig_99 "Missing")	///
	drop(Robust: _cons)
		///
	refcat(11.country "Country (ref. Germany)"
///
	1.health_99 "Health problems in 1-3 dimensions (ref. 0)"	///
	1.smokenow_99 "Smoking (ref. non-smoker)"
///
	1.drink73_99 "Alcoholic drinks (ref. none)"
///
	1.edu_99 "Education (ref. primary)"
	///
	2.empl_99 "Employment status (ref. retired)"
///
	2.marital_99 "Marital status (ref. married)"	///
	1.netempty_99 "No social network (ref. has a social network)"	///
	1.socialclub_99 "Attended social clubs in past year (ref. no)"  ///
	2.hdi_99 "HDI level of origin country (ref. highest)"	///
	1.imcitizen_99 "Citizenship in country of interview (ref. no)", nolabel) ///
	mtitles("Men" "Women")
		///
	title("{\b Table x.} Fixed effects regression models with all mediators (weighted)")
///
	addnote("Source: SHARE, wave 6, release 8.0.0, own calculations." ///
		"Respondents aged 60+, imputed values included.")


*robustness to omitting imputations
	use "$posted/w6_noimp", clear

	*Model 1
		eststo m1noimp: mlogit frailtype i.mig_99 i.gendercat	///
		c.age i.health_99 ib(12).country, base(0)

	*Model 2
		eststo m2noimp: mlogit frailtype i.mig_99##i.gendercat	///
		c.age i.health_99 ib(12).country, base(0)

	*Full list of coefficients
	esttab m1noimp m2noimp
		///
	using $table\mlogit\noimp_mlogit_12full.rtf , replace	///
	eform noconstant label compress nobase nonumbers	///
	b(%9.2fc) p(%9.3fc) par   onecell unstack
	///
	coeflabels(1.mig_99 Migrant 99.mig_99 "Missing"
///
	age "Age" 1.mig_99#1.gendercat "Female*migrant"	///
	99.mig_99#1.gendercat "Female*missing migration info"	///
	1.gendercat "Female"
		///
	99.health_99 "Missing")
		///
```

```
drop(Robust: _cons)
///
refcat(1.mig_99 "Migrant status (ref. non-migrant)"                    ///
11.country "Country (ref. Germany)"
///
1.health_99 "Health problems in 1-3 dimensions (ref. 0)", nolabel)    ///
order(1.mig_99 99.mig_99 1.gendercat 1.mig_99#1.gendercat              ///
99.mig_99#1.gendercat )
///
mtitles("Base model" "Base model with interaction")                    ///
title("{\b Table x.} Fixed effects regression models with and without interaction of gender and migrant status: Excluding imputed values")    ///
addnote("Source: SHARE, wave 6, release 8.0.0, own calculations."    ///
"Respondents aged 60+, unweighted.")



*Robustness to omitting missing values
use "$posted/w6_nomis", clear

*Model 1
eststo m1mis: mlogit frailtype i.mig_99 i.gendercat ///
c.age i.health_99 ib(12).country, base(0)

*Model 2
eststo m2mis: mlogit frailtype i.mig_99##i.gendercat ///
c.age i.health_99 ib(12).country, base(0)

*Full list of coefficients
esttab m1mis m2mis
///
using $table\mlogit\nomis_mlogit_12full.rtf , replace                  ///
eform noconstant label compress nobase nonumbers                       ///
b(%9.2fc) p(%9.3fc) par   onecell unstack
///
coeflabels(1.mig_99 Migrant
///
age "Age" 1.mig_99#1.gendercat "Female*migrant"                        ///
1.gendercat "Female")
///
drop(Robust: _cons)
///
refcat(1.mig_99 "Migrant status (ref. non-migrant)"                    ///
11.country "Country (ref. Germany)"
///
1.health_99 "Health problems in 1-3 dimensions (ref. 0)", nolabel)    ///
order(1.mig_99 1.gendercat 1.mig_99#1.gendercat)                       ///
mtitles("Base model" "Base model with interaction")                    ///
title("{\b Table x.} Fixed effects regression models with and without interaction of gender and migrant status: Excluding missing values") ///
addnote("Source: SHARE, wave 6, release 8.0.0, own calculations."    ///
"Respondents aged 60+, unweighted.")
```