University of Groningen

Faculty of Spatial Sciences

Population Studies

Master's thesis

Regional mortality inequalities in the Netherlands and the role of internal migration

Author:

Ernst van der Hoeven

e.a.van.der.hoeven@student.rug.nl

S3197646

Supervisor:

Adrien Remund

Table of Contents

Table of Contents ii
List of figures iii
List of tables iii
List of abbreviations iv
Abstract v
1. Introduction 1
1.1 Background 1
1.2 Objective and Research Questions 1
1.3 Structure 2
2. Theoretical framework and conceptual model 3
2.1 Explaining regional mortality by regional context and a rural-urban approach 3
2.2 Explaining regional mortality with selection effects 4
2.2.1 Breeder and drifter hypothesis 4
2.2.2 Migration, health and deprivation 4
2.2.3 Migration of the elderly 5
2.3 Conceptual model 5
2.4 Hypotheses 6
3. Data & Methods 7
3.1 Setting 7
3.2 Measuring life expectancy 7
3.3 Analytical strategy 7
3.3.1 Operationalisations 8
3.3.2 Three perspectives of mapping life expectancy 8
3.3.3 Analyses of the regional context 10
3.4 Ethical considerations 10
4. Results 11
4.1 Life expectancy in the Netherlands: 2015-2019 11
4.2 Hypothetical 'no-migration' scenario 13
4.2.1 Spatial patterns of the 'no-migration' scenario 13
4.3 In-migrants, out-migrants and stayers 14
4.3.1 Spatial patterns: comparing the 'no-migration' scenario and the in-migrants, out-migrants and stayers 20
4.3.2 Age-specific mortality rates of stayers and migrants 22
4.4 Regional analysis: prosperity 23
4.4.1 Descriptive statistics 24
4.4.2 Testing regional prosperity and life expectancy 24
5. Conclusion 27
5.1 Conclusion 27

5.2 Reflection 27
5.3 Recommendations 28
References 29

List of figures

Figure 1: Conceptual model 6
Figure 2: Mapping life expectancy at birth in the 40 COROP regions of the Netherlands, males and females, own calculations.. 12
Figure 3: Mapping the no-migration scenario: the change of life expectancy at birth in the no-migration scenario compared to the 'real observations,' in the 40 COROP regions of the Netherlands, males and females, own calculations.. 17
Figure 4: Confidence intervals of the four sub-groups, of life expectancy at birth (e0), the Netherlands, 40 COROP regions, females, own calculations.. 18
Figure 5: Confidence intervals of the four sub-groups, of life expectancy at birth (e0), the Netherlands, 40 COROP regions, males, own calculations.. 19
Figure 6 Comparing the gap between in-migrants and stayers to the 'no-migration' scenario, for the 40 COROP regions, own calculations 21
Figure 7 Comparing the gap between out-migrants and stayers to the 'no-migration' scenario, for the 40 COROP regions, own calculations 21
Figure 8: Age-specific mortality rates in the Netherlands (2015-2019), males and females, own calculations 22
Figure 9: Age-specific mortality rates in the Netherlands (2015-2019), institutional care homes filtered out, own calculations 23

List of tables

Table 1: the three perspectives of mapping life expectancy 9
Table 2: Population of the sub-groups per COROP region, 5-year average 2015-2019, males and females 16
Table 3: Descriptive statistics 24
Table 4: linear regression of average life expectancy (real observation) and average prosperity percentiles, averages for 40 COROP regions, males and females.. 25
Table 5: linear regression of difference in life expectancy between in-migrants and stayers and the average prosperity percentiles, averages for the 40 COROP regions, males and females 25
Table 6: linear regression of difference in life expectancy between out-migrants and stayers and the average prosperity percentiles, averages for the 40 COROP regions, males and females 26

List of abbreviations

CBS	Statistics Netherlands
COROP	Coördinatiecommissie Regionaal Onderzoeksprogramma
EU	European Union
GDP	Gross Domestic Product
HMD	Human Mortality Database
NUTS	Nomenclature of territorial units for statistics

Abstract

Identifying regional differences in life expectancy across the Netherlands is important to policy makers, especially for the allocation of regional health budgets. Internal migration can cause selection effects, leading to a distortion in the observed regional mortality. I use mortality and population data for the period 2015-2019 to calculate and analyse life expectancy in the 40 NUTS-3 regions of the Netherlands. I use aggregated data as an approach for this analysis. I assess what drives the spatial inequalities in life expectancy in the Netherlands, accounting for the role of internal migrants. By creating a hypothetical 'no-migration' scenario, potential selection effects are tested. Additionally, I compare life expectancy outcomes for in-migrants, out-migrants and stayers. Finally, using linear regressions, I investigate the relationship between regional characteristics and regional life expectancy outcomes. I demonstrate that regional life expectancy inequalities exist in the Netherlands, for both males and females. In contrast to the hypotheses and existing literature, results show that selection effects barely impact regional life expectancy differences and that non-migrants show a structurally higher life expectancy than migrants in almost all COROP regions, for both males and females. There is a positive linear relationship between the average prosperity of a region and life expectancy, suggesting the importance of regional welfare on life expectancy outcomes.

Keywords: mortality, life expectancy, inequalities, selection effects, internal migration

1. Introduction

1.1 Background

In the last years approximately 150.000 people died annually in the Netherlands. The age at which people are dying is higher than ever. In fact, life expectancy has seen a secular increase and is expected to continue to rise (CBS, 2020). Life expectancy in the Netherlands was higher in 2019 than the average in the European Union and also higher than neighbouring countries Belgium and Germany. Life expectancy in the Netherlands can be considered as high. However, national life expectancies are an average and do not show underlying inequalities in mortality. It is thought that regional mortality inequalities are large and tend to be increasing within industrialized countries (van Raalte et al., 2019).

Regional mortality, and thus regional life expectancy, reflects the overall health in an area or region (Kibele & Janssen, 2013). In a perfectly equal situation, life expectancy would be exactly the same for each region in the Netherlands. This would suggest that wherever one lives or is born, the span of life would be the same as for any other person, meaning that no region would fall behind in things such as quality of or access to health care. However, this perfect situation is not reality. The reality is that even in a relatively small country such as the Netherlands disparities in life expectancy tend to exist. Finding the underlying reasons why certain regions show lower life expectancies can substantially help policy makers in identifying focus points to battle certain problems in certain areas. For example, knowing about regional mortality differences could influence the allocation of health budgets (Janssen et al., 2016), since life expectancy differences are supposed to be avoidable and could be a result of an unequal distribution of resources or opportunities (van Raalte et al., 2019).

Considering the size of the Netherlands, migration from one side of the country to the other side is not uncommon. As a result of movement between regions, migration patterns could have an impact on the mortality in an area. For instance, healthy migrants tend to move to places with favourable living conditions, this is considered positive health selection (Janssen et al., 2016, Bentham, 1988). The other side of this process is that the relatively less healthy people stay behind. Migration patterns and life expectancy differences go hand in hand, making regional migration an additional focus point of this thesis.

Many studies regarding life expectancy disparities have focused on cross-country comparisons (Omran, 1998, Gerry et al., 2018, Jasilionis et al, 2011), whereas sub-national trends are covered less in this field of research, though the usefulness of these sub-national trends are acknowledged (Vallin & Meslé, 2005). National averages might hide variation across sub-regions (Kibele et al., 2015), and could therefore potentially be misleading. For instance, if the rural areas in the Netherlands show high mortality, while the urban areas show low mortality, then the average does not represent either of the regions. Therefore, in-depth regional analyses can possibly expose misleading averages and trends. Much of the comparison-studies are based on historical trends (Vallin & Meslé, 2005). In this study, historical trends in life expectancy are more or less disregarded and an in-depth focus on the period 2015-2019 is chosen. Because this is the most recent time period possible, results might be useful for the near future.

1.2 Objective and Research Questions

In this thesis, the main objective is to identify which spatial differences across the Netherlands exist and how these spatial differences might be impacted by selective processes of internal migration. These aims are covered by three research questions.

(i) What are the regional differences in life expectancy?
(ii) To what extent does migration affect regional life expectancy?
(iii) How are regional contextual factors affecting regional life expectancy differences?

1.3 Structure

This thesis is structured as follows: first a theoretical framework is constructed in which relevant literature in the study field of the topic is discussed and a conceptual model is developed as a bridge between the theoretical framework and the data and results sections. After this section, data and methods will be discussed, including a description of datasets used and indications of sources. Following this is the result section, where the main results of the analysis are set out and where results relevant to the research questions are highlighted. The last chapter of the thesis includes the conclusion, reflection and recommendations. The conclusion consists of an overview and summary of the research aim and relevant results, whereas the reflection is included to evaluate data, methods and obtained results.

.

2. Theoretical framework and conceptual model

Life expectancy and mortality can be approached from multiple perspectives. It is an area of interest for multiple fields of studies, such as health sciences, geography, biology, actuarial sciences and genetics. In this thesis, literature from the field of demography will be the central focus, but since mortality is such a broad topic, it is important to approach it from multiple perspectives. Given the research aim of identifying the spatial differences across the Netherlands and the potential role of selective internal migration this literature review includes three main perspectives: rural-urban relations and the urban gradient, selection effects of migrants and the role of migrants of old age in this process.

Studies on spatial mortality inequalities can be of a diachronic nature, such as the study done in Germany by Kibele et al. (2015). They present trends in life expectancy at birth over a period of a full century. In the Netherlands, Janssen et al. (2016) study mortality trends over the period 1988 to 2009, focusing on the convergence of life expectancy between regions. Their findings suggest a convergence of life expectancy between the regions during the study-period. In Austria, a similar study was done for the period 1969-2004 by Gächter & Theurl (2011). In the USA, increased inequalities in life expectancy have been found on both the county as state level (van Raalte et al., 2019). Bonnet & d'Albis (2020) provide an overview of longitudinal studies on regional convergence of life expectancy, they state that often a stagnation or increase in spatial inequality of mortality is found.

Although this thesis is not of a diachronically nature, it is important to realize that historical trends have in principle led to the current state of spatial life expectancy patterns, and therefore making results of those studies very relevant.

2.1 Explaining regional mortality by regional context and a rural-urban approach

Looking at the regional context is an obvious approach to regional mortality. Regional health outcomes might be a consequence of individuals with certain health behaviour clustering together. For instance, the prevalence of smokers in an area might stimulate others in that area to smoke as well (Kibele et al., 2015). This way, individual-level determinants of mortality actually contribute to the regional context and regional mortality. In all EU countries, mortality is influenced by socioeconomic factors such as income, employment and education level (Corsini, 2015). While socioeconomic status is in fact an individual-level determinant, people that live in areas with low socioeconomic status have higher mortality than people living in areas with higher socioeconomic status (Meijer et al., 2014). This suggests that the regional context can lead to specific health outcomes, where regional contextual factors influence the individual health outcome. Therefore, the regional context could be seen as an overarching set of opportunities for an individual, or rather lack of opportunities. For example, the labour participation in Zuid-Limburg and Oost-Groningen are respectively 63.3% and 63.4%, while it is 71.1% for Utrecht (CBS, 2020). Presumably, labour participation is low because of a lack of job opportunities, and vice versa. Therefore, regional context and its opportunities impacts health outcomes, as the lack of labour participation can be seen as contextual factor, leading to a higher presence of people with low socioeconomic status.

Moreover, a distinction in life expectancy between rural and urban areas has often been identified. Although the focus in this thesis is not to research rural-urban trends, literature regarding rural-urban aspects of mortality can still be relevant for regional differentials. Mortality can be impacted by environmental factors. For instance, urban populations might suffer from bad health as a consequence of air pollution, smog, traffic accidents or be affected by the heat island effect (Lerch et al., 2017). Rural populations are less likely to suffer from any of those factors. This would be the approach from the environmental hypothesis. The environment is part of the regional context. Thus, approaching regional life expectancy from the environmental hypothesis would suggest that rural populations have better health outcomes than urban populations, based on regional context. However, findings from Lerch et al. (2017) in Switzerland, actually show that mortality is higher in both rural areas and city centres than in urban agglomeration belts. The environmental hypothesis can therefore be dismissed in this case, though considering the historical hygiene conditions of urban areas it is a logical hypothesis to test. In fact, the mortality advantage of people in periurban areas in Switzerland is largely related to the spatial concentration of highly educated people (Lerch et al., 2017). The absence of an urban-rural gap in life expectancy was also found in Germany (Rau & Schmertmann, 2020), even though quite large disparities in regional mortality were found. District-level socioeconomic indicators that fit with the poorer segments of the population are a better explanation for life expectancy differences than for example population density, density of physicians or the gross domestic product (GDP) in a region (Rau & Schmertmann, 2020). Highlighting once more some of the important aspects of the regional context.

2.2 Explaining regional mortality with selection effects

2.2.1 Breeder and drifter hypothesis

Another way of approaching life expectancy and mortality is by looking at selection effects and selective spatial mobility. The breeder and drifter hypotheses is one approach of this theory of selection effects. The breeder hypothesis entails that someone's health is impacted by locality bound exposure and behaviour (Verheij, 1996, Verheij et al., 1998). This breeder hypothesis consists of three sets of causes: health-related environmental factors such as pollution, stress from sociocultural and physical environments and health behaviour such as smoking (Verheij et al., 1998). The drift hypothesis refers to the selection processes that result in a higher concentration of either ill people (direct selection) or in a spatial concentration of people more susceptible to illness (indirect selection) (Verheij, 1996). Result of this is that healthy people move away and the ill are left behind, or the other way around. Many studies observe that migrants cause selection effects when moving (Kibele & Janssen, 2013). Since migration is seen as health-selective, migrants tend to be healthier than stayers (except older people, see 2.2.3) (Boyle, 2004, Luy & Caselli, 2007). Migrants then contribute to the spatial concentration of people with favourable health outcomes, as is described in the drift hypothesis. This also means that migrants are expected to positively contribute to the life expectancy of a region they move to, while the region they leave behind would notice negative effects of their movement, as the 'unhealthy' people stay put.

2.2.2 Migration, health and deprivation

Following the breeder and drifter hypotheses and the health-selective nature of migration, a logical train of thought is that healthy migrants are more likely to move away from more deprived areas to less deprived areas, while unhealthy migrants are more likely to move to more deprived areas from less deprived areas (Norman et al., 2005). Healthy migrants could be attracted by labour market reasons (i.e. higher wages), while unhealthy migrants could be pushed out due to the increasing living costs in these more affluent areas. This theory suggests that migration could lead to a widening gap of health differences between regions. A framework regarding possible health deprivation and the widening gap between regions is provided by Darlington-Pollock & Norman (2020). Their framework accounts for both the effect of mobility on changing health gradients and to some extent for the interdependencies within the process of selective sorting. Concretely, people in the best health are expected to remain or transition between the most advantaged circumstances, while people in the worst health will remain or transition between the least advantaged circumstances. Consequently, this process can maintain, widen or constrain regional health differences.

2.2.3 Migration of the elderly

While migrants at working age are expected to be healthier than stayers, moves of migrants at older ages are however often related to unfavourable health conditions (Bentham, 1998, Boyle, 2004). Moves seem to happen for different reasons at old ages. Litwak and Longino (1987) identify three stages of moving for retirees. The first move is made around retiring age and is done when the migrants are in relatively good health. It is an amenity move (Zhang et al., 2013). The second move is made when people's health starts to decline and chronic diseases trouble the older people in everyday tasks such as grocery shopping (Litwak & Longino, 1987). Assistance might therefore be needed and family and friends can offer these, resulting in moves towards family and friends (Zhang et al., 2013). The third move is then made towards nursing homes and other institutions when the informal help from family and friends is no longer sufficient (Litwak & Longino, 1987, Zhang et al., 2013).

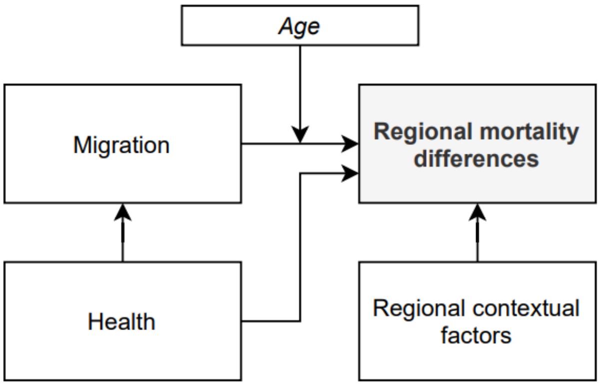
A large share of people of old age resides in nursing home institutions in the Netherlands, where mortality is higher than among the non-institutionalized people (Kibele & Janssen, 2013). Kibele & Janssen (2013) also observe that people of old age nearing death move more frequently than those with a longer remaining lifespan. As a result, the place of death might not represent the place the individual has spent most of its life, resulting in selection effects in the region where the individual died, while the region that was left behind 'avoids' the mortality. To illustrate this, this effect was found on neighbourhood level in Amsterdam by Jonker et al. (2013), where neighbourhoods with a nursing home had lower life expectancies than neighbourhoods without a nursing home. Thus, the selective migration of frail elderly people led to an increased mortality in a neighbourhood, whereas other neighbourhoods 'avoided' mortality numbers. This could also happen on a larger scale (e.g. municipalities), with the absence (or presence) of a nursing home institution.

Since people at old age have higher mortality risks, it is important to be aware of their role as migrants in regional mortality. Specifically, one should consider their possible selective migration, possibly causing misleading mortality numbers.

2.3 Conceptual model

Figure 1 shows the conceptual model based on the literature review. It is expected that regional mortality differences are impacted by migration, health and regional contextual factors. This can be a positive or a negative impact. Concretely, migrants can potentially cause an increase (positive) or decrease (negative) in life expectancy in a region, which contributes to potential regional life expectancy differences. Similarly, health can either have a positive or negative effect on regional mortality differences, for example by the clustering of healthy people in a region. Health is a predictor of regional mortality differences, but is also a mediator of migration, since migration is thought to be possibly driven by health reasons. Age functions as a moderator for migration. Migrants tend to be healthier than non-migrants at lower ages, but at older ages they are expected to be less healthy (see 2.2.3). Regional contextual factors are also considered to be a predictor of regional mortality differences. For example, the average prosperity in a region could function as a potential predictor of regional mortality differences.

Figure 1: Conceptual model



2.4 Hypotheses

In order to answer the research questions, several hypotheses are tested. The hypotheses are based on the theoretical framework, and the conceptual model (figure 1). Hypothesis 1 is tested in order to answer the first research question. Hypotheses 2 and 3 are tested to answer the second research question regarding migration. While hypothesis 4 is tested to answer the third research question, accounting for the regional context.

H1: Regional mortality differences are present in the Netherlands
H2: Regional differences in life expectancy are partially caused by selection effects
H3: Migrants have a higher life expectancy than non-migrants, except at old ages
H4: The average prosperity in a region has a positive impact on life expectancy in the region

3. Data & Methods

3.1 Setting

The total population of the Netherlands is studied for the period 2015-2019. For the regions, the 40 NUTS-3 – the hierarchical system of dividing up the economic territory of the EU - regions are studied. This is originally known as COROP regions in the Netherlands. All provinces are divided in multiple sub-regions, with the exception of Flevoland and Utrecht. The COROP regions are based on the nodal principle, for which commuter flows form the basis. This nodal principle entails that each COROP region has a central core with a surrounding catchment area. As a result the regions are roughly equal in area size, but not necessarily in population size (see table 2, 4.3). In some cases, regions do not fully follow the nodal principle as COROP regions do not cross provincial borders, this could create a slight bias in regional characteristics for regions that have a city hub on a provincial border, such as the city of Groningen. The borders of the COROP regions have remained unchanged during the study period of 2015-2019. See Appendix 1 for an overview of the distinguished COROP regions.

3.2 Measuring life expectancy

Individual and aggregated mortality data and population data were accessible through Statistics Netherlands for the period 2015-2019. Aggregated data were essential to calculate life expectancy. These datasets include sex, age (reshaped into age categories) and region. Demographic data in the Netherlands is maintained with a register-based system (Basisregistratie Personen). This entails that demographic statistics in the Netherlands are based on the digitised municipal population registers. These registers provide the personal details of every resident and non-resident (CBS, 2016). Personal information on events such as marriage, citizenship, data about death or data about the address are all included in this population register. Since 1980, the Netherlands has consistently scored well on the vital statistics performance index (VSPI) with scores of ≥0.85, placing them in the highest category in terms of data quality (Mikkelsen et al., 2015).

Datasets containing individual characteristics (e.g. COROP code) were aggregated on a regional level and ultimately these yearly datasets were combined to create one life expectancy for the period 2015-2019. This aggregated method is useful to take out randomness since some regions have low numbers of mortality (e.g. Delfzijl en Omgeving with 571 yearly deaths on average). Life expectancy at birth (e0) is used and referred to throughout this thesis, though life expectancy was calculated for all age categories (0, 1-4, 5-9, …, 95-99, 100+). Life expectancies were calculated for all COROP regions for the period 2015-2019, using life tables, the method taken from Preston, Heueveline & Guillot (2000). In this method values for ${}_na_x$ were taken from the national population in 2015, as found in the Human Mortality Database (HMD). Sticking to these values for ${}_na_x$ makes the life expectancy outcomes more consistent, since the datasets are eventually aggregated. Exception was made for the ${}_na_x$ values for the age categories 0 and 1-4, here the method of Coale and Demeny (1983), as found in Preston, Heuveline & Guillot (2000), was used as estimation of the time spent in the interval for those who are dying.

3.3 Analytical strategy

The analyses conducted to answer our research questions can be divided into two sections. First, a spatial analysis is done by mapping three perspectives of migration. Regional life expectancy outcomes are compared for the 'real observation' to answer the first research question. Then, life expectancy outcomes are compared for the different sub-populations (in-migrants, out-migrants and stayers), covering the second research question. Confidence intervals (Andreev & Shkolnikov, 2010) are used to prove significant differences between the sub-populations.

Second, the role of the regional context is analysed using several linear regressions, using regional life expectancy and regional prosperity (average in a region) of the 40 COROP regions. Additionally, regressions are conducted to further elaborate on findings of the three perspectives, meaning that the two sections of the analyses are not separate entities.

3.3.1 Operationalisations

Defining the subpopulations based on migration status

Aggregated analyses is used to prove disparities in life expectancy in the Netherlands and to seek possible explanations for these disparities in life expectancy. By looking into regional migration-flows possible selection effects could be revealed. It is important to define regional migrants in order to determine migrant-flows. In this thesis only internal migrants are relevant for migration-flows, dismissing international migrants for this part, though they are still present in the population data. The following definition is in place for internal migrants: an internal migrant is considered one when one lived in a different COROP region in the period 2015-2019, compared to the COROP region for the period 2005-2009. For example, when one lived in COROP region 22 in 2006, but lives in COROP region 38 in 2016, this person is considered an internal migrant. There has been a time span of 10+ years for a move to happen, allowing for a higher aggregate number of movers. By considering only those who move between COROP regions as internal migrants, moves over shorter distances are mostly filtered out. This is useful, because moves over shorter distances are more likely to happen purely for housing reasons, this is called residential mobility (Darlington-Pollock & Peters, 2020). When looking at a change of COROP code, it is implied that a mover has migrated over a larger distance, and thus had a significant reason to move, such as for labour-market reasons (Mulder & Malmberg, 2014). This group of migrants is our group of interest and making this distinction in migrants is useful to better assess the impact of internal migration on regional life expectancy differences

3.3.2 Three perspectives of mapping life expectancy

Cartographic analyses

To identify regional differences in life expectancy a cartographic analysis is done. To explore the effects of internal migration and regional life expectancies three perspectives are used (see also table 1).

First, the real observation is mapped by calculating a life expectancy for each of the 40 COROP regions in the Netherlands over the period 2015-2019, for both sexes.

Secondly, I artificially place back every person in the Netherlands to their place of residence 10 years prior, i.e. the period of 2005-2009. This means datasets with place of residence in 2005-2009 are also needed (see table 1). Persons that either were not born in 2005-2009 or did not reside in the Netherlands during that period are assigned with either the same COROP code as they have had in 2015-2019 (for the unborn) or are manually placed into the fictional COROP group 41 (for the international migrants). The aim of this second method is to generate a counterfactual scenario in which selective migration is eliminated in order to estimate the effect of this process on spatial mortality differences. By comparing the observed and "no-migration" scenarios, we can conclude the importance of migrants' self-selection on the interregional disparities in life expectancy. Practically, the regions that 'lose' life expectancy (i.e. have a lower life expectancy in the in the "no-migration" scenario than in reality), benefit from a net positive selection of migrants, as they either attract healthy migrants or lose unhealthy inhabitants. Reciprocally, the regions that show a 'gain' in life expectancy (i.e. have a higher life expectancy in the “no-migration” scenario than in reality) suffer from a net negative selection of migrants. They either attract unhealthy migrants or lose healthy inhabitants.

These effects are net in two ways. On the one hand, they are net because each migrant flow is composed of both healthy and unhealthy individuals (i.e. they have higher risk of death compared to the average of their age group), thus a migration flow to a certain region is deemed ‘positively selected’ if it is composed of more healthy migrants than unhealthy migrants, and vice versa. This first limitation, which is unavoidable with aggregate data, is not necessarily problematic, as the question I aim to answer is also of an aggregate nature. On the other hand, these effects are net because this ‘no-migration’ counterfactual scenario only considers the combined effect of in- and out-migrants, which can either compensate or reinforce each other. For instance, a region could theoretically experience both a positive selection of in-migrants (in-migrants are healthier than the residents in each age group), and a positive selection of out-migrants (out-migrants are healthier than the residents in each age group), thus leading to an uncertain effect that depends on the size of the in- and out-flows and the strength of the selection processes in each direction. This second limitation could be more problematic, but can be addressed by considering the two components independently, i.e. comparing the in- and out-migrants to the stayers in each region.

The third perspective of mapping life expectancy addresses the second limitation mentioned above. It is used to test negative and positive selection. Life expectancy of in-migrants and out-migrants are separately compared to the life expectancy of stayers. An overview the groups can be found in table 1. Positive selection happens when the in-migrants are healthier than the stayers, or when out-migrants are unhealthier than the stayers. Negative selection happens when in-migrants are unhealthier than stayers, or when out-migrants are healthier than stayers.

In-migrants and out-migrants are the same population, they are however separated by using the COROP code of 2015-2019 for in-migrants and the COROP code of 2005-2009 for out-migrants. For out-migrants this means that mortality rates of the first two age categories (0-1 & 1-4) are only present in the fictional COROP region of 41. To keep a consistent life expectancy in the life tables, the population average mortality rates are used for the first two age categories for all 40 COROP regions.

Since the size of the migrant population in a COROP region can be very small, it is useful to compute confidence intervals (Andreev & Shkolnikov, 2010), for the life expectancy outcomes. This allows for an effective comparison of life expectancy between the three groups (in-migrants, out-migrants and stayers), since it addresses the magnitude of the stochastic error in small populations (Andreev & Shkolnikov, 2010). When the outer bounds of the confidence intervals show no overlap, we can tell with certainty that the life expectancy outcomes of the sub-groups (point estimates) in a COROP region are different.

Table 1: the three perspectives of mapping life expectancy

		2015-2019		
		Region X	Region Y	Region Z
2005-2009	Region X	A	D	G
	Region Y	B	E	H
	Region Z	C	F	I

Perspective 1 (reality): groups = ABC vs DEF vs GHI
Perspective 2 (no-migration): groups = ADG vs BEH vs CFI
Perspective 3 (stayers vs in-migrants vs out-migrants): groups = A vs BC vs DG, E vs DF vs BH, I vs GI vs CF

Table 1 presents an overview of the comparison of the three sub-groups. Perspective 1 is the comparison of the regions (X,Y,Z) using the real population, meaning that the place of residence in 2005-2009 is not relevant, resulting in a comparison between groups ABC vs DEF vs GHI. In perspective 2, groups ADG, BEH and CFI are compared, as the place of residence in 2005-2009 is needed, while 2015-2019 is not relevant. Perspective 3 is a comparison between groups A vs BC vs DG, E vs DF vs BH, I vs GI vs CF. Group A, E and I represent the stayers in the regions, groups BC, DF and GI represent the in-migrants of the regions and groups DG, BH and CF represent the out-migrants of the regions.

3.3.3 Analyses of the regional context

The role of the regional context on potential regional life expectancy differences are tested by doing linear regressions with the life expectancy in a region as dependent variable (n=40). The average prosperity in a region is used as a regional contextual factor. This average is calculated by taking the average of all individuals in a region. The prosperity data is provided by Statistics Netherlands and is a combination of income and capital at the household level. The prosperity data is in percentiles. For other linear regressions, regional life expectancy differences between the sub-groups are taken as dependent variables (n=40), using the same explanatory variable in regional prosperity.

3.4 Ethical considerations

Some of the data used in this research is highly sensitive, as every registered individual in the Netherlands is present in the dataset. To protect the privacy of individuals in the Netherlands data is handled in a secure way. First, data is immediately pseudonymized after collection by Statistics Netherlands. This is done by changing citizen service number (BSN) into a ‘pseudokey’. Other possible identifications of personal characteristics are also immediately pseudonymized. Since the data is still not anonymous, data needs to be handled with care. The pseudonymized datasets are called microdata and are only available to third parties under very strict conditions. One of the measurements to ensure the safe use of the sensitive data is that employees have a duty of confidentiality and have signed a nondisclosure agreement. Another important measurement is that Statistics Netherlands only uses data for a scientific or statistic purpose and never for fiscal, administrative, control and judicial purposes. This is legally excluded. It is the responsibility of a researcher to ensure that none of the results presented can be traced back to any individual and that all results are therefore a result of aggregated data.

4. Results

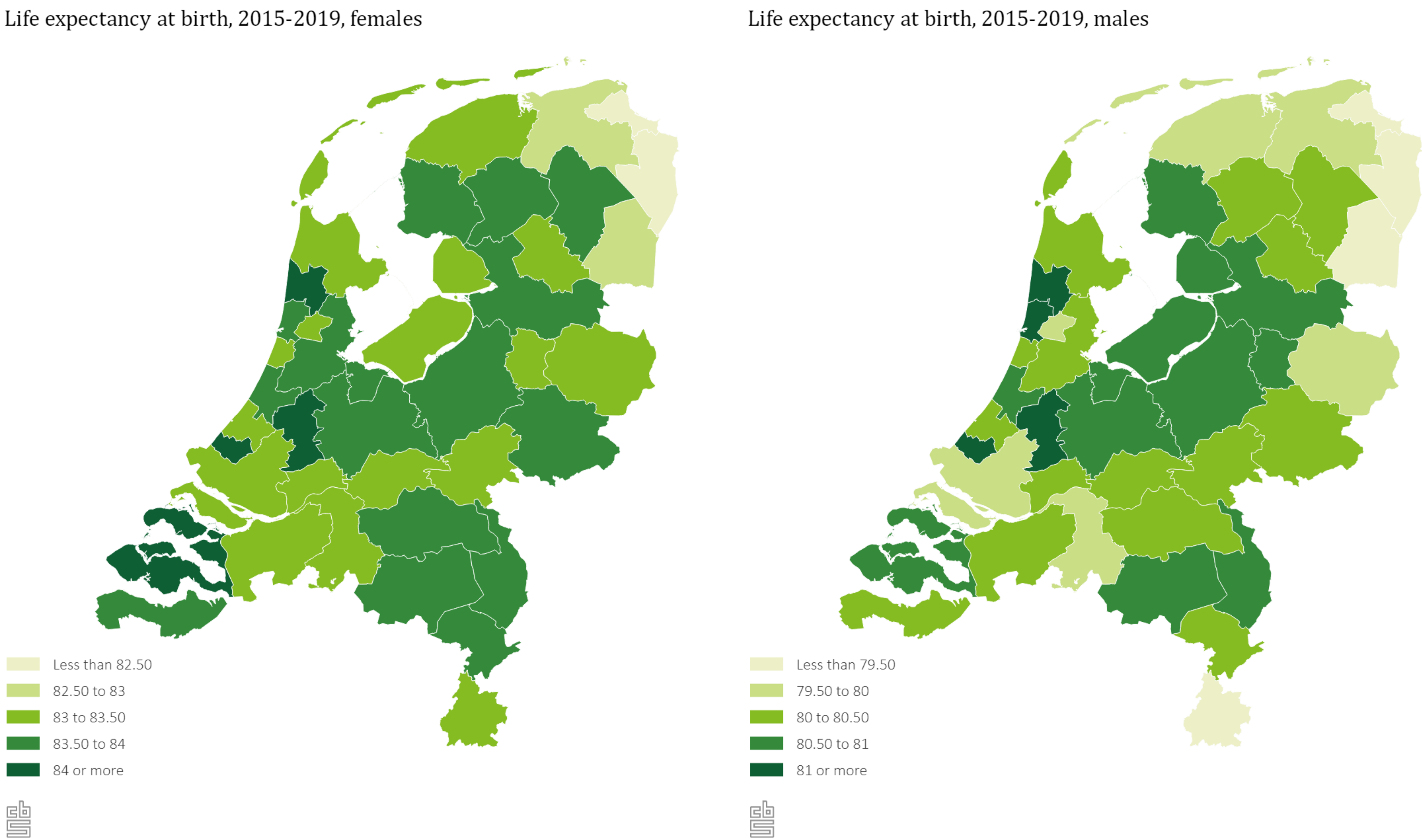
The three perspectives of mapping life expectancy (see table 1) are presented as follows. For the ‘real observation’ (figure 2), regional life expectancy at birth is classified into five coloured categories, with each category representing a gap of 0.5 years. For the ‘no-migration scenario’ (figure 3), the net loss or gain of life expectancy compared to the real situation is mapped, also identifying five coloured categories, ranging from less than -0.1 to 0.2 or more, with gaps of 0.1 years. The third perspective of comparing in-migrants, out-migrants and stayers is done through the use of confidence intervals (figure 4 & 5).

4.1 Life expectancy in the Netherlands: 2015-2019

During the period of 2015-2019 the calculated average life expectancy at birth (e0) for males was 80.31 for males and 83.52 for females. Life expectancy per COROP region for males and females are shown in figure 2. The maps for males and females show similar patterns, though the map for males includes more regions with life expectancy in the lowest categories. The disparities in life expectancy are therefore larger for males than for females. The lowest life expectancy for males is found in Oost-Groningen (78.67), while the highest life expectancy is found in Delft en Westland (81.36). For females the highest life expectancy is also found in Delft en Westland (84.66), and the lowest also in Oost-Groningen (82.11). Overall these maps demonstrate that disparities in life expectancy are present in the Netherlands, confirming hypothesis 1, as life expectancy varies across the COROP regions by more than 2.5 years for males (2.69) and females (2.55). Spatial patterns include that the four regions with the highest life expectancy are ‘clustered’ in Noord-Holland and Zuid-Holland, while the highest life expectancies for females can also mostly be found in these regions. The lowest life expectancies for males can be found in the border-regions, namely in Groningen (Oost-Groningen, Delfzijl en Omgeving and Overig Groningen), Zuid-Limburg and Zuidoost Drenthe. Life expectancy for females shows roughly the same pattern, although with smaller differences between them.

Even though these results are purely of a descriptive nature, they do confirm the existence of a problematic situation, namely the inequality of life expectancy at birth across different regions in the Netherlands. While there are countries with larger gaps in life expectancy, such as France, where a difference of 3.80 years was found for women between the *département* with the highest life expectancy and the lowest life expectancy (Bonnet & d’Albis, 2020), it is still a problematic phenomenon, especially since it is thought to be avoidable (van Raalte et al., 2019). Janssen et al. (2016) found that during the period of 1988-2009 the 40 COROP regions in the Netherlands showed smaller life expectancy differences, as the most disadvantaged regions showed the biggest increase in life expectancy, while the most advantaged regions showed the least increase in life expectancy. However, despite these differences becoming smaller, Janssen et al. (2016) mention that the differences in mortality between the economically advantaged and disadvantaged regions are likely to be persistent. Regions that showed some of the lowest life expectancy in 1988, such as Zuid Limburg & Oost-Groningen (Janssen et al., 2016) also show some of the lowest life expectancies for the period 2015-2019 (figure 2). Similarly, Delft en Westland showed the highest life expectancy from 1988 onwards, and still shows the highest life expectancy in the Netherlands for the period 2015-2019, for both males and females.

Figure 2: Mapping life expectancy at birth in the 40 COROP regions of the Netherlands, males and females, own calculations



4.2 Hypothetical 'no-migration' scenario

The second perspective of mapping life expectancy in the Netherlands is the 'no-migration' scenario (also see table 1). This generated counterfactual scenario presents the importance of migrants' self-selection, since selection effects are eliminated. Figure 3 shows these results. The maps (males and females) can be interpreted as follows: regions that are coloured red have a higher life expectancy in the hypothetical 'no-migration' scenario than in the real situation. While the regions that are coloured blue have a lower life expectancy in the hypothetical situation than in the real situation. It can then be concluded that the regions that are red suffer from a net negative selection of migrants. They either attract more unhealthy migrants than healthy ones or lose more healthy inhabitants than unhealthy ones. The opposite of this net selection applies to the blue regions.

Considering the 'no-migration' scenario in figure 3, it can be concluded that migrants do not play a notable role in existing life expectancy differences. The change in life expectancy ranges from -0.2 to 0.3, but most regions only show a change of -0.1 to 0.1. There is also a lack of certainty about the results considering the potential stochastic error that exists when studying small populations (Andreev & Shkolnikov, 2010). Thus, a situation without migration (i.e. the 'no-migration' scenario) does not create a notably different pattern of life expectancy outcomes, compared to the real situation. It is likely that the proportion of migrants to non-migrants is too low to be impactful on life expectancy outcomes.

Even for a region such as Utrecht, where the net-migration is high (i.e. there are more immigrants than emigrants for the region, see table 2), the scenario with no migrants did not result in a notably different result than the real observation. These findings mean that hypothesis 2 is mostly not supported.

Even though the effect on the total life expectancy outcome are more or less negligible, it does not mean that further analysis on migrants and life expectancy is not needed. The results just show that the impact of migration on the total life expectancy is small, not that there are no different life expectancy outcomes for migrants compared to non-migrants.

4.2.1 Spatial patterns of the 'no-migration' scenario

The change in life expectancy (figure 3) results in quite a diverse map for both males and females. For males a pattern of blue regions (lower life expectancy in the hypothetical situation without migration) is visible in the eastern border regions of the Netherlands. It contains regions such as Twente, Achterhoek, Noord-Limburg and Zuid-Limburg. Surprisingly, this pattern is reversed for females, as these regions form a pattern of red regions (higher life expectancy in the hypothetical situation). It suggests that the health of female and male migrants in these regions are opposite, as for males these regions have a net positive selection (i.e. they attract healthy male migrants), while these areas suffer from a net negative selection of females (i.e. they lose healthy migrants).

Regions such as Groot-Amsterdam and Utrecht are popular migrant destinations (see table 2). With the healthy-migrant theory in mind, it would be expected that these regions with a high positive net-migration suffer from a loss in life expectancy in the 'no-migration' scenario compared to the real situation, since a relative large share of their real population is moved out, and these migrants are expected to be healthy (i.e. healthy migrant theory). However, the spatial pattern (figure 3) does not confirm this line of thought. For example, Utrecht (red) and Groot-Amsterdam (blue) show opposite results. While an area that is known to lose migrants such as Delfzijl en Omgeving (see table 2) shows the same colour (red) as the popular migrant destination Utrecht.

Considering the mixed results, there are no unambiguous spatial patterns in the 'no-migration' scenario. We could speculate that different streams of migrants exist, i.e. wealthy migrants moving to urban areas in a COROP region and poorer migrants moving to rural areas in the same COROP region. In a way, these different streams of migrants would then ‘cancel’ out their selection effects.

The results of the ‘no-migration’ scenario can be a useful addition when comparing in-migrants, out-migrants and stayers. An additional analysis of this will be presented in section 4.3.1.

4.3 In-migrants, out-migrants and stayers

The size of the migrant population seems to be too small (table 2) to heavily weigh on regional life expectancy outcomes. However, health outcomes between different types of migrants and stayers might still be different. Additionally, by comparing in-migrants, out-migrants separately to stayers, the problem of only considering the combined effect of in- and out-migrants in the counterfactual scenario can be addressed. Table 2 presents an overview of average population per COROP region for the distinguished sub-groups.

Using confidence intervals, a method provided by Andreev & Shkolnikov (2010), a comparison between in-migrants, out-migrants, stayers and the real observation was done. Figure 4 & 5 show the results for males and females. In-migrants and out-migrants form the same population, but for in-migrants the place of residence for 2015-2019 was taken, while for out-migrants the place of residence 2005-2009 was taken. This means that the national average life expectancy of in-migrants and out-migrants is the same, but that on a COROP level differences exist.

Considering figure 4 & 5, it can be concluded that stayers are structurally healthier (i.e. have a higher life expectancy) than migrants for both males and females (using a 95% confidence interval). For females, life expectancy is significantly higher for all COROP regions for stayers, as there is no overlap in the outer bounds of the intervals. For males almost all COROP regions show a significantly higher life expectancy for stayers than for in-migrants and out-migrants, except for COROP region 2 & 25 (only out-migrants) and 32 & 33 (in-migrants and out-migrants). Out-migrants and in-migrants only show significantly differences in a few cases, namely 26 and 27 for females and 24 for males.

The largest gap between in-migrants and stayers can be found in Zuidwest-Overijssel (CR19) with 5.57 years for females and in Agglomeratie Leiden en Bollenstreek (CR22) with 4.94 years for males. The largest gap between out-migrants and stayers can be found in Delft en Westland (CR27) with 6.27 years for females and in Twente (CR20) with 4.73 years for males. The interpretation of these results is mixed. For example, the region Delft and Westland is the region with the highest life expectancy in the country. That the difference between female out-migrants and stayers is the largest seems to make sense, this area ‘loses’ unhealthy migrants. Delft and Westland is also one of only two regions for females where life expectancy of out-migrants is significantly lower than that of in-migrants. However, female in-migrants are also still unhealthier than stayers in this region, but they seem to be *less* unhealthy. The magnitude of the abovementioned gaps are large, but as discussed (see 4.2) seem to barely impact the overall life expectancy of a region.

Another specific case that draws the attention is that of Agglomeratie ‘s-Gravenhage (CR24) for males (figure 5). Here, out-migrants are significantly healthier (2.66 years) than in-migrants. And out-migrants do not have a significant different life expectancy than stayers (or the real population). It suggests an outflow of healthier male migrants in this region. In fact, the red colour (figure 3) of the region actually reflects this, as the red colour is an indication of a net loss of healthier people. They likely lose more healthy inhabitants than unhealthy ones.

Overall, the structural difference in life expectancy of in-migrants and out-migrants compared to stayers means that hypothesis 3 is not supported. The results are surprising in two ways. First, in many studies, migrants are considered healthier than non-migrants (see Boyle, 2004, Luy & Caselli, 2007), although moves at old age are more likely to be related to unfavourable health outcomes (see Bentham, 1998, Boyle, 2004). In section 4.3.2 a more detailed analysis on age-specific mortality rates is done to further elaborate on this theory.

Second, even in the regions that seem attractive for migrants (i.e. have good living conditions) a significant difference was found (e.g. Groot-Amsterdam). This refutes the theory of positive health selection in the Netherlands. While in the no-migration scenario the impact of migrants on regional life expectancy was found to be very small, they are also in fact unhealthier on average.

Table 2: Population of the sub-groups per COROP region, 5-year average 2015-2019, males and females

		Males				Females			
CR	**Region**	In-migrants	Out-migrants	Stayers	Real Obsv.	In-migrants	Out-migrants	Stayers	Real Obsv.
01	Oost-Groningen	11,195	7,531	59,360	70,556	11,127	8,038	59,849	70,975
02	Delfzijl en omgeving	2,950	3,468	20,319	23,268	2,918	3,531	20,437	23,355
03	Overig Groningen	39,540	29,603	157,918	197,459	40,312	31,089	157,904	198,215
04	Noord-Friesland	17,392	14,019	144,118	161,509	17,978	15,349	42,316	160,294
05	Zuidwest-Friesland	7,333	8,265	62,209	69,542	7,925	8,823	60,931	68,856
06	Zuidoost-Friesland	11,421	10,052	81,771	93,192	11,774	10,786	82,023	93,798
07	Noord-Drenthe	14,859	11,835	79,872	94,731	15,038	12,420	81,545	96,584
08	Zuidoost-Drenthe	9,316	8,173	74,376	83,692	9,527	8,921	74,907	84,434
09	Zuidwest-Drenthe	8,502	6,744	56,902	65,404	8,851	7,264	57,466	66,317
10	Noord-Overijssel	21,325	16,882	162,657	183,982	22,531	18,960	161,331	183,862
11	Zuidwest-Overijssel	10,389	8,446	66,525	76,913	10,676	9,086	66,974	77,650
12	Twente	26,917	18,917	289,021	315,937	25,581	18,887	286,359	311,940
13	Veluwe	43,450	30,328	293,697	337,148	45,160	31,825	298,383	343,543
14	Achterhoek	19,825	16,496	179,513	199,338	20,235	17,409	180,480	200,715
15	Arnhem/Nijmegen	50,491	29,824	310,274	360,765	53,344	33,539	318,356	371,701
16	Zuidwest-Gelderland	14,413	12,371	106,203	120,616	14,891	13,145	105,603	120,493
17	Utrecht	90,996	68,992	547,626	638,622	98,570	75,259	563,801	662,371
18	Kop van Noord-Holland	18,562	15,561	168,839	187,401	19,035	16,202	167,749	186,784
19	Alkmaar en omgeving	15,422	11,141	105,724	121,146	15,884	11,558	107,356	123,240
20	IJmond	13,187	8,886	84,137	97,325	13,114	9,120	86,162	99,276
21	Agglomeratie Haarlem	24,877	14,198	87,154	12,032	26,051	14,956	92,351	118,402
22	Zaanstreek	13,800	8,354	70,386	84,186	13,791	8,476	72,186	85,977
23	Groot-Amsterdam	144,844	64,533	522,446	667,290	145,979	69,049	539,137	685,115
24	Het Gooi en Vechtstreek	22,148	15,222	100,127	122,275	22,945	15,492	106,529	129,474
25	Agglomeratie Leiden en Bollenstreek	30,557	19,122	177,358	207,915	34,207	20,750	180,664	214,870
26	Agglomeratie 's-Gravenhage	83,026	37,859	339,221	422,247	84,936	39,473	352,142	437,077
27	Delft en Westland	29,269	18,046	87,951	117,220	23,316	15,195	87,430	110,746
28	Oost-Zuid-Holland	19,879	16,743	142,085	161,964	20,677	17,879	143,541	164,218
29	Groot-Rijnmond	90,004	47,681	612,321	702,325	91,340	48,210	633,697	725,037
30	Zuidoost-Zuid-Holland	30,351	15,532	163,821	194,173	31,161	16,080	167,411	198,572
31	Zeeuwsch-Vlaanderen	7,075	3,710	45,371	52,445	6,891	3,927	46,200	53,091
32	Overig Zeeland	13,766	10,519	123,355	137,121	14,607	11,209	124,787	139,395
33	West-Noord-Brabant	36,521	20,594	275,931	312,452	36,756	22,432	280,519	317,275
34	Midden-Noord-Brabant	31,296	19,602	205,143	236,439	31,126	20,324	206,442	237,568
35	Noordoost-Noord-Brabant	33,433	23,168	293,384	326,817	34,039	24,696	292,463	326,502
36	Zuidoost-Noord-Brabant	47,177	22,976	339,685	386,862	42,541	23,322	335,591	378,131
37	Noord-Limburg	13,549	9,919	127,888	141,437	13,817	10,870	126,069	139,886
38	Midden-Limburg	13,195	7,924	105,057	118,252	12,610	8,657	105,377	117,987
39	Zuid-Limburg	29,996	15,477	265,015	295,011	31,505	17,306	272,865	304,369
40	Flevoland	32,062	23,128	173,289	205,351	31,378	23,501	173,772	205,150

Figure 3: Mapping the no-migration scenario: the change of life expectancy at birth in the no-migration scenario compared to the 'real observations,' in the 40 COROP regions of the Netherlands, males and females, own calculations

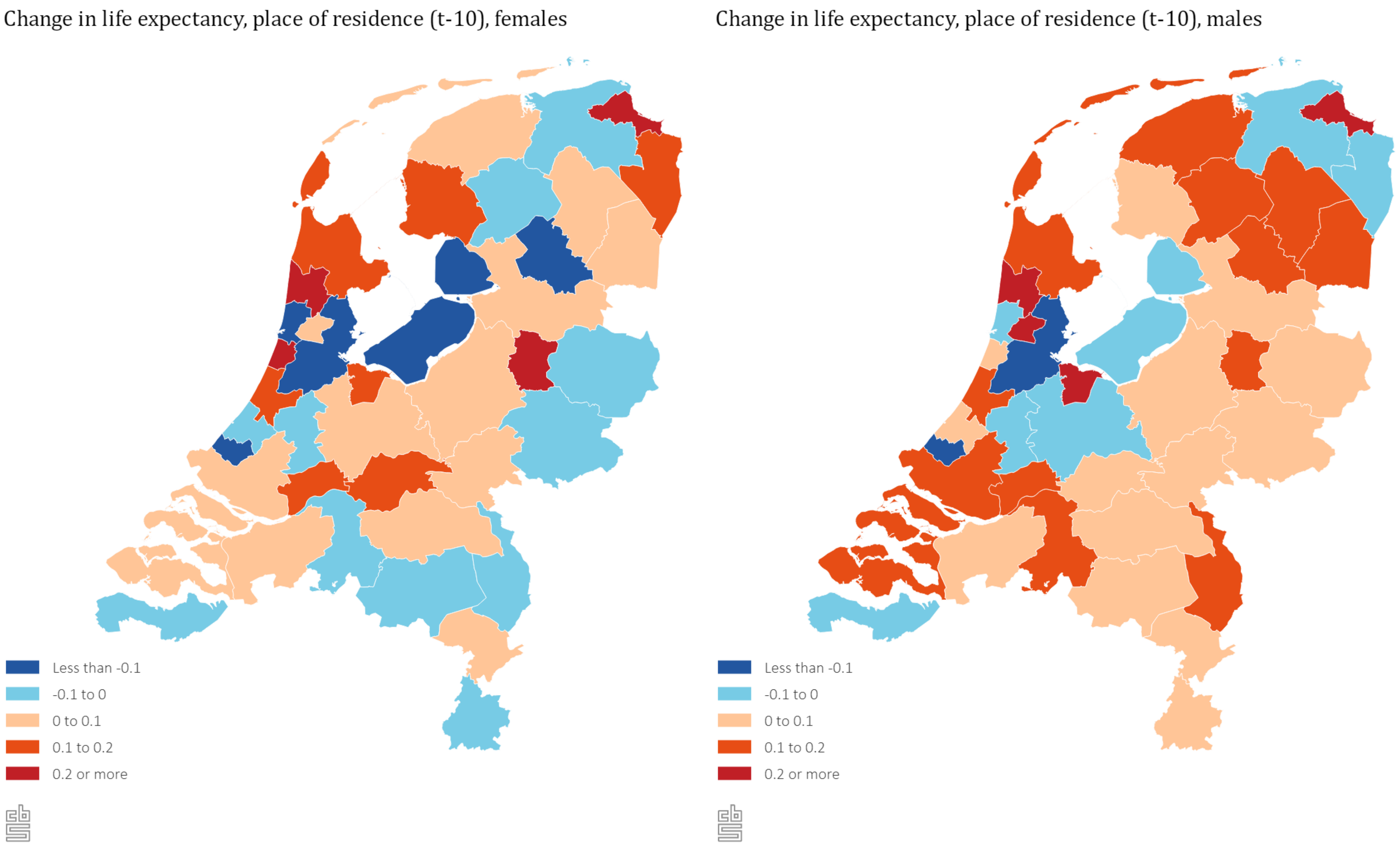


Figure 4: Confidence intervals of the four sub-groups, of life expectancy at birth (e0), the Netherlands, 40 COROP regions, females, own calculations

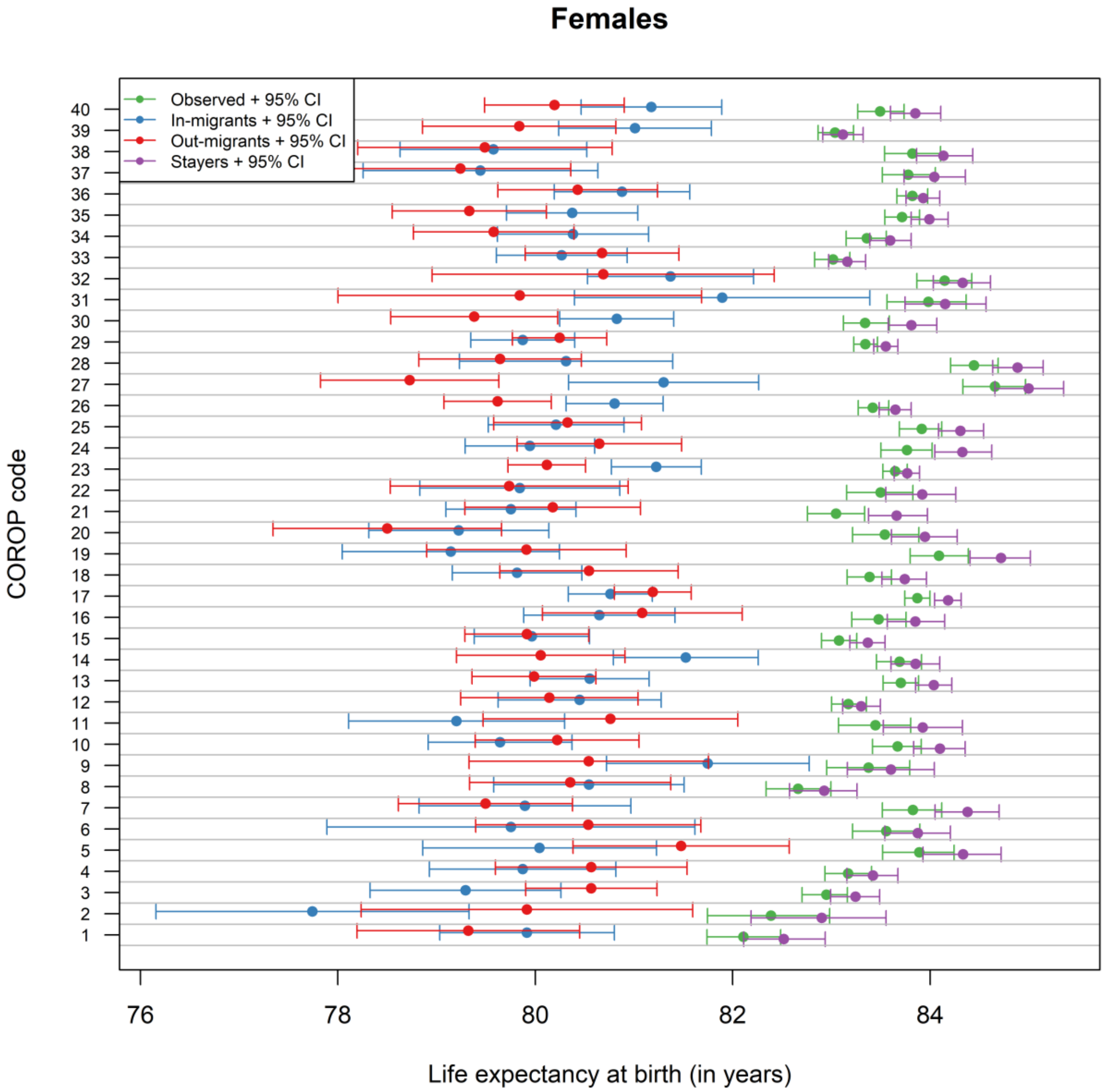
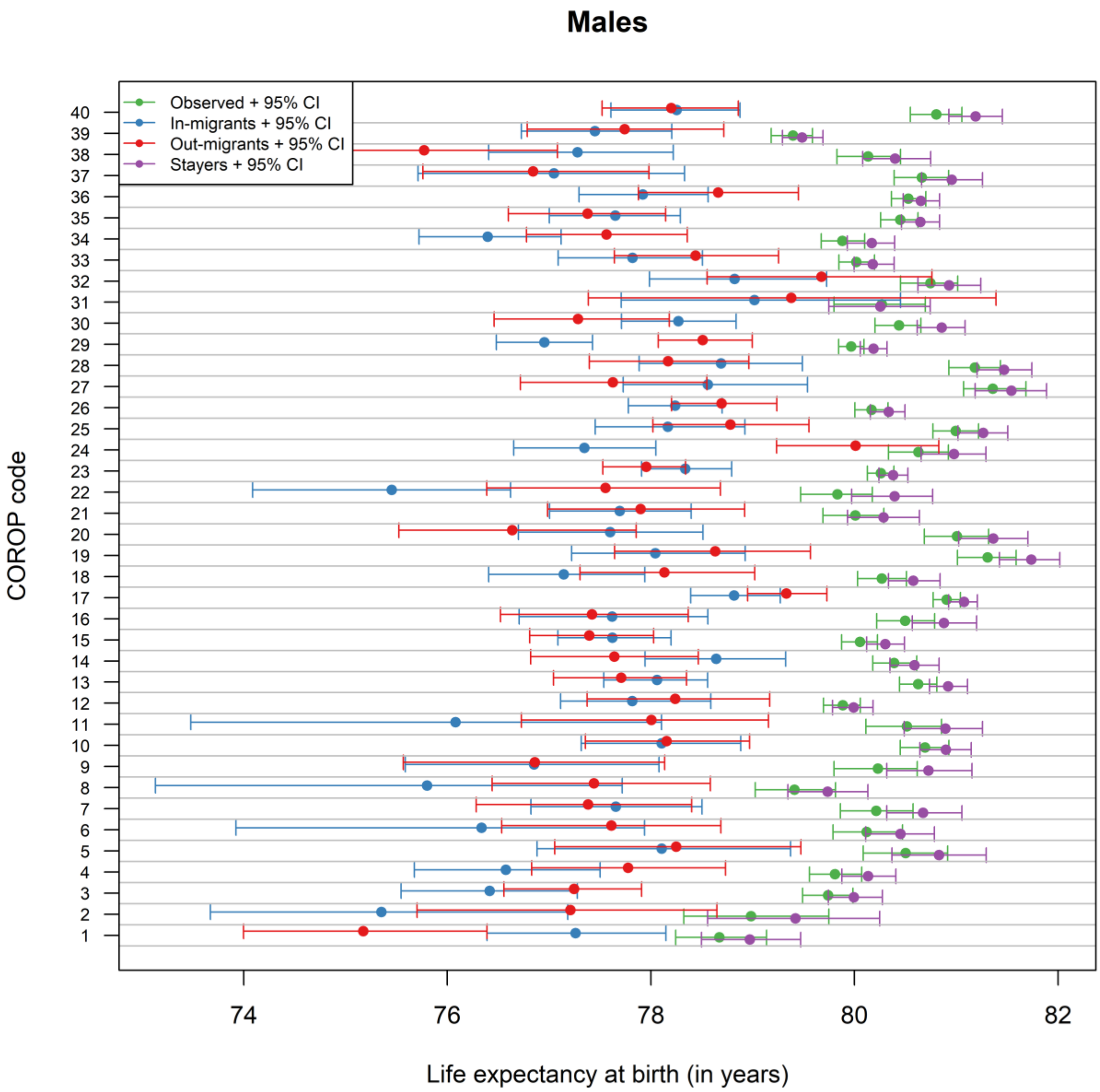


Figure 5: Confidence intervals of the four sub-groups, of life expectancy at birth (e0), the Netherlands, 40 COROP regions, males, own calculations



4.3.1 Spatial patterns: comparing the ‘no-migration’ scenario and the in-migrants, out-migrants and stayers

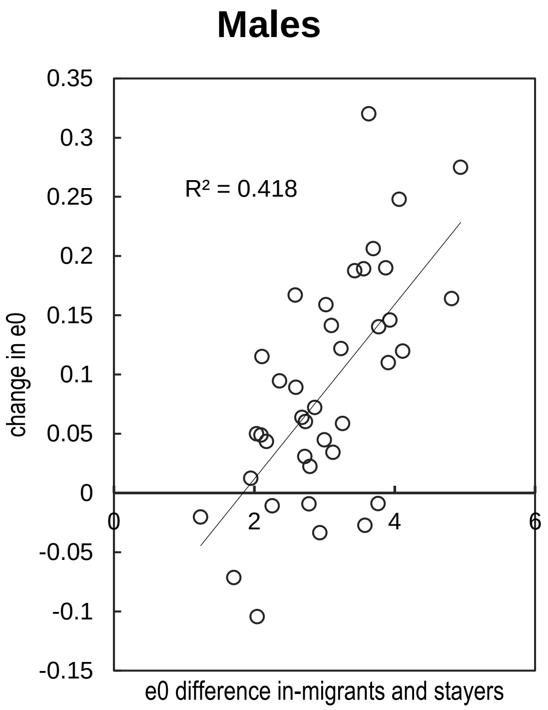
By comparing outcomes of the ‘no-migration’ scenario (figure 3) with the outcomes of the comparisons of in-migrants, out-migrants and stayers (figure 4 & 5), spatial patterns could possibly be explained. For example, by testing if a region coloured blue in the ‘no-migration’ scenario also attracts relative healthier (i.e. gap between stayers and in-migrants is smaller) migrants than regions coloured red in the ‘no-migration’ scenario, and vice versa. To test this, scatterplots were created, theorising that a region coloured blue (i.e. a lower life expectancy in the hypothetical scenario) attracts (relatively) healthier migrants, meaning that the gap between in-migrants and stayers would be smaller. When comparing out-migrants and stayers, the opposite effect is expected, meaning that a region coloured blue shows a bigger life expectancy gap between out-migrants and stayers.

Figure 6 shows the gap between in-migrants and stayers (x-axis) and the change in life expectancy in the ‘no-migration’ scenario (y-axis). For both males and females CR27 formed an outlier, and has been removed from the data. For males, it seems that the larger the gap between in-migrants and stayers, the larger the gap is in the ‘no-migration’ scenario. Thus, regions coloured blue (negative value on y-axis) seem to attract the relative healthier migrants than the regions coloured red (positive value on y-axis). For females, results are similar, although the R^2 is higher for males (0.418) than for females (0.2582). We can speculate that the observed patterns for males and females are caused by certain characteristics of the relative healthier in-migrants (e.g. brain-gain).

Additionally, figure 7 shows the gap between out-migrants and stayers (x-axis) and the change in life expectancy in the ‘no-migration’ scenario (y-axis). CR27 formed an outlier for these graphs as well, and has been removed. Blue areas (negative value on the y-axis) are scattered across the whole x-axis, meaning that blue areas (i.e. lower life expectancy in the hypothetical scenario) do not necessarily show the loss of relatively unhealthier out-migrants. Vice versa, are red areas (positive value on the y-axis) scattered across the whole x-axis, meaning that these red areas (i.e. higher life expectancy in the hypothetical scenario) do not unambiguously attract relatively unhealthier migrants. The R^2 is low for both males (0.0207) and females (0.0392).

Considering figure 6 and 7, it seems that migrant effects only apply for in-migrants and not for out-migrants, since figure 6 (in-migrants) shows somewhat of a linear relationship, while a linear relationship is almost totally absent in figure 7 (out-migrants).

Figure 6 Comparing the gap between in-migrants and stayers to the 'no-migration' scenario, for the 40 COROP regions, own calculations



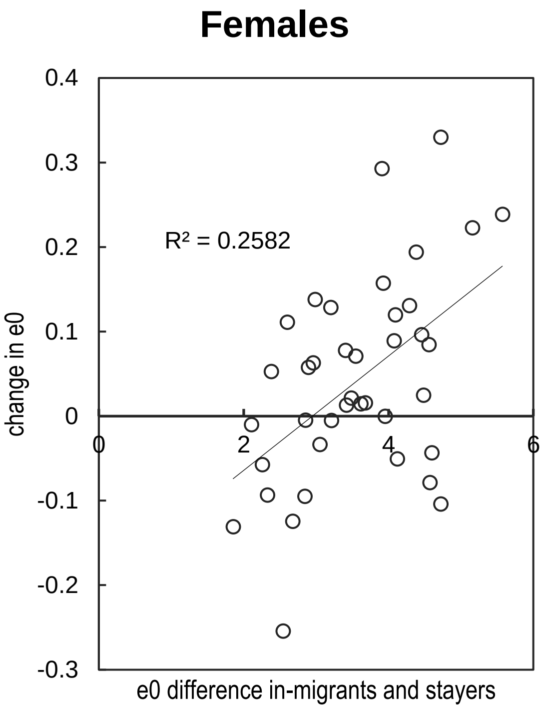
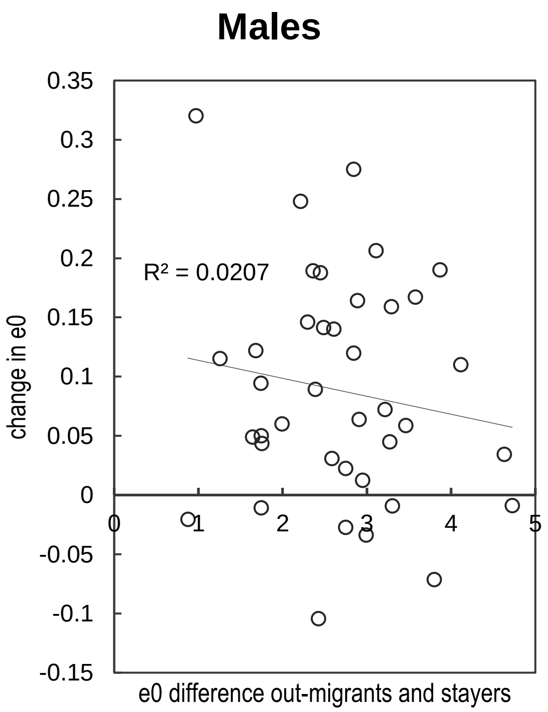
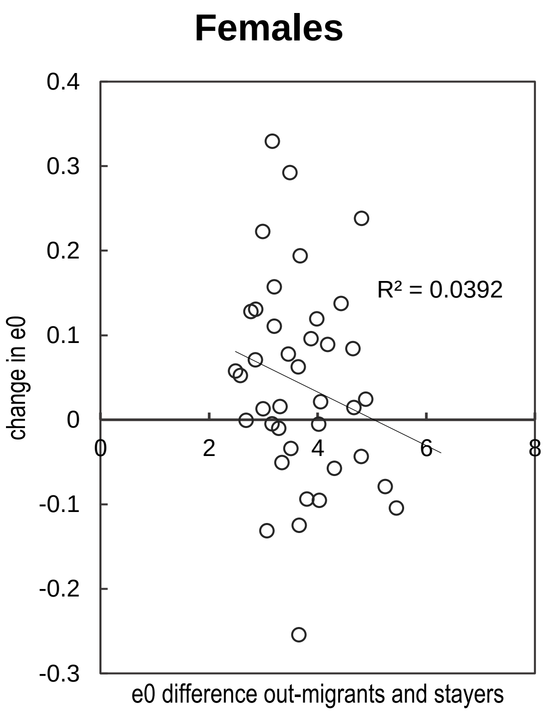


Figure 7 Comparing the gap between out-migrants and stayers to the 'no-migration' scenario, for the 40 COROP regions, own calculations



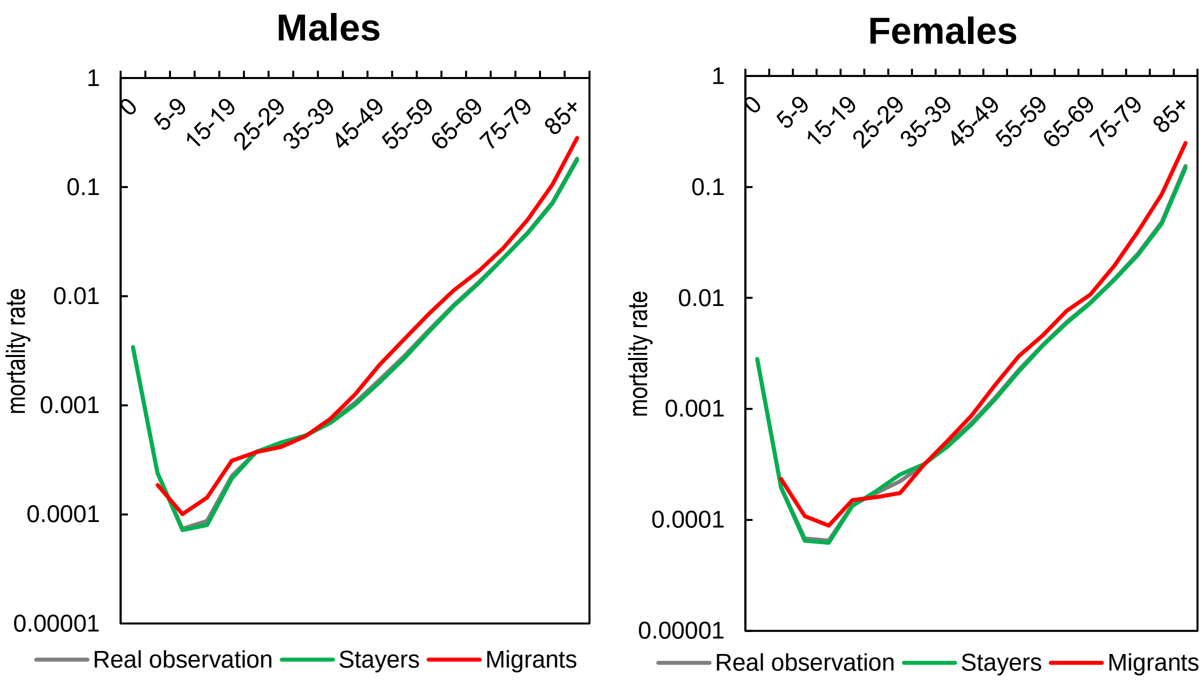


4.3.2 Age-specific mortality rates of stayers and migrants

Age-specific mortality rates for the Dutch population (2015-2019, aggregated) were calculated for the real observation, stayers and migrants (in- and out-migrants are the same population). The age-specific mortality rates are presented in figure 8 for males and females. Age-specific mortality rates could provide an insight on the structural difference in life expectancy, as migrants generally tend to be healthier than non-migrants, except at older ages (see also 2.2.1 and 2.2.3).

The results are somewhat in line with the abovementioned theory. Mortality rates are indeed higher for migrants at old age (60+) than for stayers, for both males and females. However, mortality rates are higher for migrants at almost all ages, except for ages 20 to 34 (males and females). This means that the structural lower life expectancy of migrants is caused by people outside the age groups of 20-34 years, since the migrants aged 20-34 prove to be (slightly) healthier than stayers. Thus it can be concluded that the results of the age-specific mortality rates are only partially in line with the theoretical expectations that migrants are healthier, except at older ages.

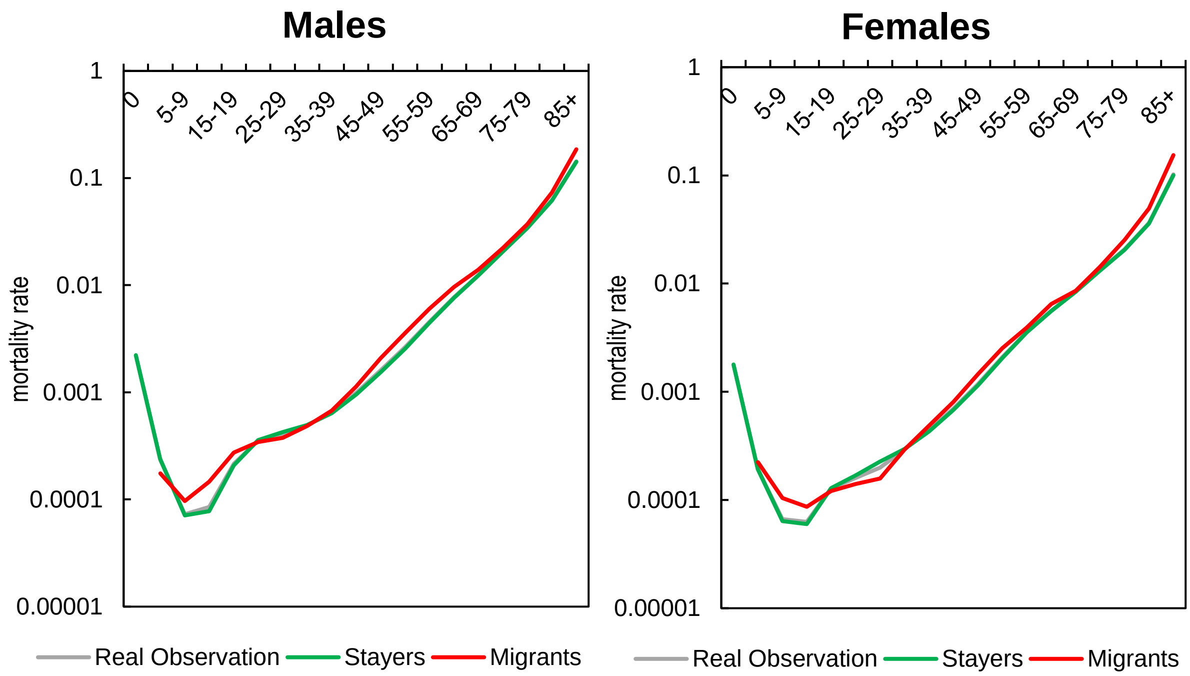
Figure 8: Age-specific mortality rates in the Netherlands (2015-2019), males and females, own calculations



Age-specific mortality rates were also calculated for the Dutch population (all three sub groups) while filtering everyone living in an institutional care home out. This allows to test whether migrant mortality rates are higher at old ages because they might be more likely to move to institutional care homes. Meaning that the proportion of elderly people living in institutional care homes of the migrant group is possibly higher than the proportion of the stayers.

Figure 9 shows the age-specific mortality rate of the three sub groups for males and females, while the population living in institutional care homes are filtered out. It can be concluded that although the overall mortality rates are much lower for all sub-groups, the proportion to one another is very similar to the age-specific mortality rates shown in figure 8. However, specifically for ages 70-79 the gap between stayers and migrants is much narrower when people living in institutional care homes are filtered out. Considering the three stages of moving of the elderly, the third move is made towards nursing homes and other institutions when the informal help from family and friends is no longer sufficient (Litwak & Longino, 1987). This is a likely explanation of the narrowing gap of mortality for migrants compared to stayers of the age 70-79. For migrants in this specific age group, moves towards institutional care homes partially explain the higher mortality rates in the 'real' situation, since mortality in institutional care homes is higher (Kibele & Janssen, 2013). For the other age categories, the possible bias formed by moves towards institutional care homes is not an explanatory variable of the structural lower life expectancy of migrants. Although it must be considered that only national mortality rates are studied in these two graphs, and that different patterns on COROP level could exist. However, seeing that migrants are significantly unhealthier for almost every COROP region in the Netherlands, results should not notably differ.

Figure 9: Age-specific mortality rates in the Netherlands (2015-2019), institutional care homes filtered out, own calculations



4.4 Regional analysis: prosperity

The outcomes in this section are twofold: First, a possible relationship between average regional prosperity and average life expectancy is tested, hereby trying to account for the regional contextual factors as a predictor of life expectancy differences. Second, the relationship between the average regional prosperity and the size of the gap in the life expectancy outcomes of stayers, in-migrants and out-migrants is reviewed (i.e. is the gap between stayers and in-migrants smaller in more affluent regions?).

Socioeconomic status is one of the best indicators of mortality (see also 2.1). On an individual scale it is expected that the higher the income of a person the higher the life expectancy outcomes would be. However, in this section, the regional character of income will be tested, by taking the average regional prosperity and the average regional life expectancy outcomes. Prosperity is a useful variable since it accounts for the welfare of a household. This eliminates possible skewed results one could get with income, since persons with a low income can still come from affluent households. The benefit of using the regional average of prosperity, is that the effect is measured as a regional contextual factor.

4.4.1 Descriptive statistics

Table 3 shows the descriptive statistics for the 40 COROP regions. The average prosperity in a region ranges from 47.29 to 62.52 for males, and from 45.96 to 60.71 for females. Prosperity is on average slightly higher (55.24) for males than for females (53.20). The average is not 50 as would be expected with percentiles, because individual prosperity is converted to averages for the regions and the regions differ in population size, meaning that for the regions a different average is shown. Similarly, life expectancy (real) does not equal the earlier found average life expectancy (see 4.1), because the average of the regions is taken instead of the average of the population. The average difference between in-migrants and stayers, and between out-migrants and stayers is higher for females (3.58, 3.76) than for males (3.04, 2.71).

Table 3: Descriptive statistics

Variable names	Observations	Mean	Standard Deviation	Min.	Max.
	Males				
COROP Code	40	--	--	1	40
Prosperity	40	55.23894	3.73603	47.28736	62.52029
Life Expectancy (Real)	40	80.28945	.5768693	78.67285	81.36024
Diff. in-migrants and stayers	40	3.036619	.8254359	1.236497	4.940887
Diff. out-migrants and stayers	40	2.708851	.9127501	.8733307	4.726381
	Females				
COROP Code	40	--	--	1	40
Prosperity	40	53.20314	3.718329	45.9568	60.70045
Life Expectancy (Real)	40	83.50735	.4977484	82.10798	84.65642
Diff. in-migrants and stayers	40	3.579915	.9007542	1.854521	5.572562
Diff. out-migrants and stayers	40	3.759498	.8526272	2.484665	6.270799

4.4.2 Testing regional prosperity and life expectancy

The coefficients of prosperity (percentiles) in table 4 show that there is a significant positive linear relationship ($p < 0.001$) between the average regional prosperity and the regional life expectancy at birth for both men and women, this confirms hypothesis 4. The effect is however, larger for men (0.104) than for women (0.087). The increase of the average prosperity in a region by 1 percentile would lead to an increase in life expectancy of 0.104 years for men and 0.087 years for women. This is a strong effect, suggesting that the average prosperity in a region is an important factor for life expectancy in regions. It can therefore be concluded that the average life expectancy in a region strongly benefits from a higher average regional prosperity. This is not a particularly surprising finding, as welfare is one of the best indicators of life expectancy. These results also prove that regional life expectancy differences for the real observation (figure 2) are partially explained by the average prosperity of a region. In light of the findings on selection effects, do these results highlight the importance of economic factors on life expectancy outcomes.

Table 4: linear regression of average life expectancy (real observation) and average prosperity percentiles, averages for 40 COROP regions, males and females

Life Expectancy (Real)	**Coef.**	**Std. Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
	Males					
Prosperity	.1038682	.0185337	5.60	0.000	.0663487	.1413877
_cons	74.55188	1.026062	72.66	0.000	72.47473	76.62904
	Females					
Prosperity	.086884	.01652	5.26	0.000	.053441	.120327
_cons	78.88484	.8810065	89.54	0.000	77.10134	80.66835

Table 5 and 6 show regressions of the difference in life expectancy of in-migrants and stayers (table 5) and of out-migrants and stayers (table 6), and prosperity. There is no significant linear relationship between prosperity and the size of the life expectancy gap between stayers and in-migrants, for both males and females. For men, there is also no significant result found between the life expectancy gap between out-migrants and stayers and prosperity. However, for women a significant ($p < 0.01$) positive linear relationship exists between the regional prosperity and the life expectancy gap between out-migrants and stayers, meaning that an increase of the average prosperity by 1 percentile in a region would lead to an increase of 0.098 years of the gap between out-migrants and stayers. Thus, the higher the average prosperity in a region, the bigger the life expectancy difference between female out-migrants and stayers will be.

The coefficients for females are larger for females in table 5 (0.060) and table 6 (0.098) than for males in table 5 (-0.018) and table 6 (0.010).The coefficient in table 6 is significant for females, while the p-value (0.122) is also quite low for females in table 5. Considering the size of the sample for the test, this could hint to a possible relationship between the difference in life expectancy between in-migrants and stayers and the average prosperity. These results suggest that moves by women could be driven by other reasons than for men, since the coefficient is either significant or close to significance for women. Though, the size of the coefficients are very small, suggesting a weak effect.

Table 5: linear regression of difference in life expectancy between in-migrants and stayers and the average prosperity percentiles, averages for the 40 COROP regions, males and females

Diff. in-migrants and stayers	**Coef.**	**Std. Err.**	**T**	**P>t**	**[95% Conf.**	**Interval]**
	Males					
Prosperity	-.0148486	.0357601	-0.42	0.680	-.087241	.0575439
_const	3.856839	1.979747	1.95	0.059	-.1509503	7.864628
	Females					
Prosperity	.0602011	.0380649	1.58	0.122	-.0168572	.1372594
_const	.3770267	2.029987	0.19	0.854	-3.732466	4.48652

Table 6: linear regression of difference in life expectancy between out-migrants and stayers and the average prosperity percentiles, averages for the 40 COROP regions, males and females

Diff. out-migrants and stayers	**Coef.**	**Std. Err.**	**T**	**P>t**	**[95% Conf.**	**Interval]**
	Males					
Prosperity	.0102897	.0395972	0.26	0.796	-.0698705	.09045
_cons	2.140457	2.192178	0.98	0.335	-2.297376	6.578289
	Females					
Prosperity	.0976642	.0336554	2.90	0.006	.0295325	.1657959
_cons	-1.436543	1.79483	-0.80	0.428	-5.069986	2.1969

5. Conclusion

5.1 Conclusion

The aim of this study was to identify spatial differences across the Netherlands and to examine the role of internal migration on these differences. Life expectancy differences were found to be existent across the 40 COROP regions in the Netherlands, confirming the first hypothesis. Life expectancy varies across the COROP regions by 2.5 years for males and females. By using a hypothetical 'no-migration' scenario, it was found that that migrants only play a very small role in existing life expectancy differences in the Netherlands, as the change in life expectancy in the 'no-migration' scenario compared to the real observation was of a more or less negligible size for most regions. Hereby, largely refuting the second hypothesis. Life expectancy of the subgroups of in-migrants and out-migrants were compared to the life expectancy of non-migrants. It was then found that non-migrants are significantly and structurally healthier – have a higher life expectancy – than in-migrants and out-migrants for almost every COROP region in the Netherlands. Age-specific mortality rates show that only migrants at young ages (20-34) show lower mortality rates than stayers, resulting in that the structural lower life expectancy is caused by people outside of these age groups. Only findings for people aged 20-34 and people of old age are in line with the third hypothesis and results found in other studies. Findings for the other age groups go against the third hypothesis and other studies done on migrants, as migrants are expected to be healthier than non-migrants at those ages. These findings are a useful addition to the 'no-migration' scenario as they highlight the absence of positive health selection in the Netherlands. Finally, several linear regressions were done to test the relationship between the average prosperity of a region and the life expectancy in that region. Findings show that there is a strong positive relationship between the average prosperity of a region and life expectancy. This confirms the fourth hypothesis. For female out-migrants a significant positive linear relationship was found between the average regional prosperity and the life expectancy gap between out-migrants and stayers.

5.2 Reflection

This study is a useful addition to existing literature mainly because of the findings on the life expectancy of migrants compared to non-migrants. The findings of the structural lower life expectancy of migrants were not in line with both hypothesis 3 and existing literature on the topic (Boyle, 2004, Luy & Caselli, 2007). It is likely that the substantial difference in life expectancy is caused by certain characteristics of the migrants compared to the non-migrants. For example, divorced people might be more likely to migrate while having worse health outcomes. Further research should focus on these individual characteristics. Since the gap between non-migrants and migrants was found to be very large in most regions, it could be hypothesised that the difference is caused by many 'small' variables and characteristics, rather than one main explanatory variable. Concretely, the characteristics of, for example, divorced people might only form a rather small percentage of the explanation.

Additionally, findings on selection effects turned out to be limiting in its explanation of regional life expectancy differences, and are therefore not in line with other studies on selection effects (Verheij, 1996, Verheij et al., 1998, Norman et al., 2005, Kibele & Janssen, 2013). Thus, this study has highlighted the importance of other variables as an explanation to existing life expectancy differences in the Netherlands. It should also be noted that results from this study only apply to the Netherlands, since migrant patterns and the characteristics of migrants could differ from country to country, making this topic rather complex for country comparisons, even when using similar methods. The hypothetical method to test selection effects, might result in more notable differences between the observed mortality and the hypothetical mortality when only considering people of old-age, such as is done in the study by Kibele & Janssen (2013), as moves of older people are more likely related to health.

Although selection effects seem to be have little effect, regional differences in life expectancy are existent in the Netherlands. This study highlights that this is problematic. While the average regional prosperity is a good indicator of life expectancy in a region, it is also likely that individual level characteristics form explanatory variables for life expectancy differences across the regions.

One of the limiting factors of the methodology is the definition of migrants in this study. A migrant status is determined on a comparison between points in time (i.e. 2005-2009 compared to 2015-2019). While the change of COROP is in place to filter out moves that are not of interest, such as moves within the daily activity space, it also means that moves made in between the periods of 2005-2009 and 2015-2019 are not considered. Practically, a person that has lived in the same COROP region during 2005-2009 as during 2015-2019 is considered a non-migrant, even if that person did actually move to a different COROP region between 2009 and 2015 and then returned to the original COROP region. Those possible moves are not tracked in our data. Ideally, one would use continuous migration data, while also holding the minimal moving distance in place. Additionally, not all short distance moves (i.e. within daily activity space) are filtered out, because a person can move only 5km and still see a change of COROP code. This means that the data is not perfect, but this is not necessarily problematic, since this type of distortion is not seen in large numbers.

A second limitation of this study is the absence of trends. There are no points in time to compare the obtained results to, meaning that developments of the results over time are not examined. This limitation is partly solved by looking into other studies on the topic including historical trends.

5.3 Recommendations

A useful addition to this study would be to conduct an analysis on the individual characteristics of migrants and non-migrants. In this thesis the macro processes have been studied using aggregated data. Studying micro processes on an individual level would make for a very useful addition. Individual survival analyses would be extremely suitable to gain more in-depth insights on the processes that drive regional morality differences. It would be especially useful to explore the individual characteristics of internal migrants in the Netherlands. This way, regional contextual factors can be turned into individual characteristics, which allows for better testing of the regional context. For example, multilevel regional characteristics are included, such as the degree of urbanity of a municipality, and the prosperity of a COROP region. Ideally, a multilevel survival analysis is conducted, to account for potential multi-level regional characteristics. This way, regional characteristics become a better predictor of regional mortality differences in the Netherlands. I recommend future research on this topic to include this individual level survival analysis. I have explored using this method and data is available and ready for this type of study, but the method has not been included due to time constraints for the master’s thesis.

In this study existing regional inequalities in life expectancy were presented, confirming the first hypothesis. It has become clear that selection effects barely influence these regional inequalities. Rather, they are likely to be caused by social inequality or an unfair distribution of resources, as the results of the linear regressions on prosperity and life expectancy suggest. Therefore, it is necessary that policy makers keep addressing the problem of spatial inequality of life expectancy. It is necessary for policy makers to address the underlying causes of mortality differences in the 40 COROP regions in the Netherlands, especially since mortality differences can be an indicator of underlying problems in a region. I recommend that regional life expectancy in the Netherlands will be tracked and that the aim of policy makers should be to minimise life expectancy differences over time.

References

Austin, P. C. (2017). A Tutorial on Multilevel Survival Analysis: Methods, Models and Applications. *Int Stat Rev*. 85(2), 185–203.

Andreev, E. & Shkolnikov, V. (2010). *Spreadsheet for calculation of confidence limits for any life table or healthy-life table quantity*. (MPIDR TECHNICAL REPORT 2010-005). Max Planck Institute for Demographic Research

Bentham, G. (1988). Migration and morbidity: implications for geographical studies of disease. *Social Science & Medicine*, 26(1), 49–54.

Bonnet, F. and d'Albis, H. (2020), Spatial Inequality in Mortality in France over the Past Two Centuries. *Population and Development Review*, 46: 145-168.

Boyle, P. (2004). Population geography: Migration and inequalities in mortality and morbidity. *Progress in Human Geography*, 28(6), 767–776.

Centraal Bureau voor de Statistiek. (2020). *Arbeidsdeelname; regionale indeling 2019*. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84703NED/table?ts=1610638118704 (Retrieved January 14, 2021).

Centraal Bureau voor de Statistiek. (2016). *Population register data, basis for the Netherlands' Population Statistics*. The Hague: Statistics Netherlands https://www.cbs.nl/-/media/imported/documents/2016/53/2015bt18-population-register-data.pdf?la=en-gb (Retrieved January 5, 2021).

Centraal Bureau voor de Statistiek. (2020). *Prognose periode-levensverwachting; geslacht en leeftijd, 2020-2070*. https://www.cbs.nl/nl-nl/cijfers/detail/84883NED (Retrieved January 4, 2021).

Coale, A. and Demeny, P. (1983). *Regional model life tables and stable populations*. New York: Academic Press.

Corsini, V. (2015). Inequalities in life expectancy by socioeconomic status in the EU. *European Journal of Public Health*, 25

Darlington-Pollock, F., & Norman, P. (2020). Establishing a framework of analysis for selective sorting and changing health gradients. *Population, Space and Place*

Darlington-Pollock, F., & Peters, K. (2020). Progress in the study of health inequalities and selective migration: Mobilising the new mobilities paradigm. *Progress in Human Geography*.

Gächter, M. & Theurl, E. (2011). Health Status Convergence at the Local Level: Empirical Evidence from Austria. *International Journal for Equity in Health*, 10(1), 34-47.

Gerry, C.J., Raskina Y. & Tsyplakova D. (2018). Convergence or Divergence? Life Expectancy Patterns in Post-communist Countries, 1959–2010. *Social Indicators Research*. 140(1), 309-332.

Janssen, F., van den Hende, A., de Beer, J. A. A. & van Wissen, L. J. G. (2016) Sigma and beta convergence in regional mortality: A case study of the Netherlands. *Demographic Research*, 35, 81-116.

Jasilionis, D., Meslé, F., Shkolnikov, V.M., Vallin, J., (2011). Recent Life Expectancy Divergence in Baltic Countries. *European Journal of Population / Revue européenne de Démographie*, 27(4), 403-431.

Jonker, M., Lenthe, F., Donkers, B., Congdon, P., Burdorf, A., & Mackenbach, J. (2013). The impact of nursing homes on small-area life expectancies. *Health & Place*, 19(1), 25–32.

Kibele, E. & Janssen, F. (2013). Distortion of regional old-age mortality due to late-life migration in the Netherlands? *Demographic Research*, 29(5), 105-132.

Kibele, E., Klüsener, S., & Scholz, R. D. (2015). Regional Mortality Disparities in Germany: Long-Term Dynamics and Possible Determinants. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 67(1 Supplement), 241-270.

Lerch, M., Oris, M. & Wanner, P. (2017). Periurbanization and the transformation of the urban mortality gradient in Switzerland. *Population*, 72(1), 93-122.

Litwak, E., & Longino, C. F., Jr. (1987). Migration Patterns Among the Elderly: A Developmental Perspective. *The Gerontologist*, 27(3), 266–272.

Luy, M. & Caselli, G. (2007). The impact of a migration-caused selection effect on regional mortality differences in Italy and Germany. *Genus,* 63, 33–64.

Meijer, M., Röhl, J., Bloomfield, K., & Grittner, U. (2012). Do neighborhoods affect individual mortality? a systematic review and meta-analysis of multilevel studies. *Social Science & Medicine*, 74(8), 1204–1212.

Mikkelsen, L., Phillips, D. E., AbouZahr, C., Setel, P. W., Savigny, D. de, Lozano, R., & Lopez, A. D. (2015). A global assessment of civil registration and vital statistics systems: Monitoring data quality and progress. *The Lancet*, 386(10001), 1395–1406.

Mulder, C.H. & Malmberg, G. (2014). Local ties and family migration. *Environment and Planning A*. 46(9), 2195-2211

Norman, P., Boyle, P., & Rees, P. (2005). Selective migration, health and deprivation: a longitudinal analysis. *Social Science and Medicine*, 60(12), 2755–2771.

Omran, A.R. (1998). The epidemiologic transition theory revisited thirty years later. *World Health Statistics Quarterly*. 51, 99–119.

Preston, S. H., Heuevline, P. & Guillot, M. (2001). *Demography: Measuring and Modeling Population Processes*. Malden, MA: Blackwell Publishers Inc.

Raalte, van, R. A. A., Klüsener S, Oksuzyan, A., & Grigoriev, P. (2020). Declining regional disparities in mortality in the context of persisting large inequalities in economic conditions: the case of Germany. *International Journal of Epidemiology*, 49(2), 486–496.

Rau, R., & Schmertmann, C. P. (2020). District-level life expectancy in Germany. *Deutsches Arzteblatt International*, 117(29-30), 493–499

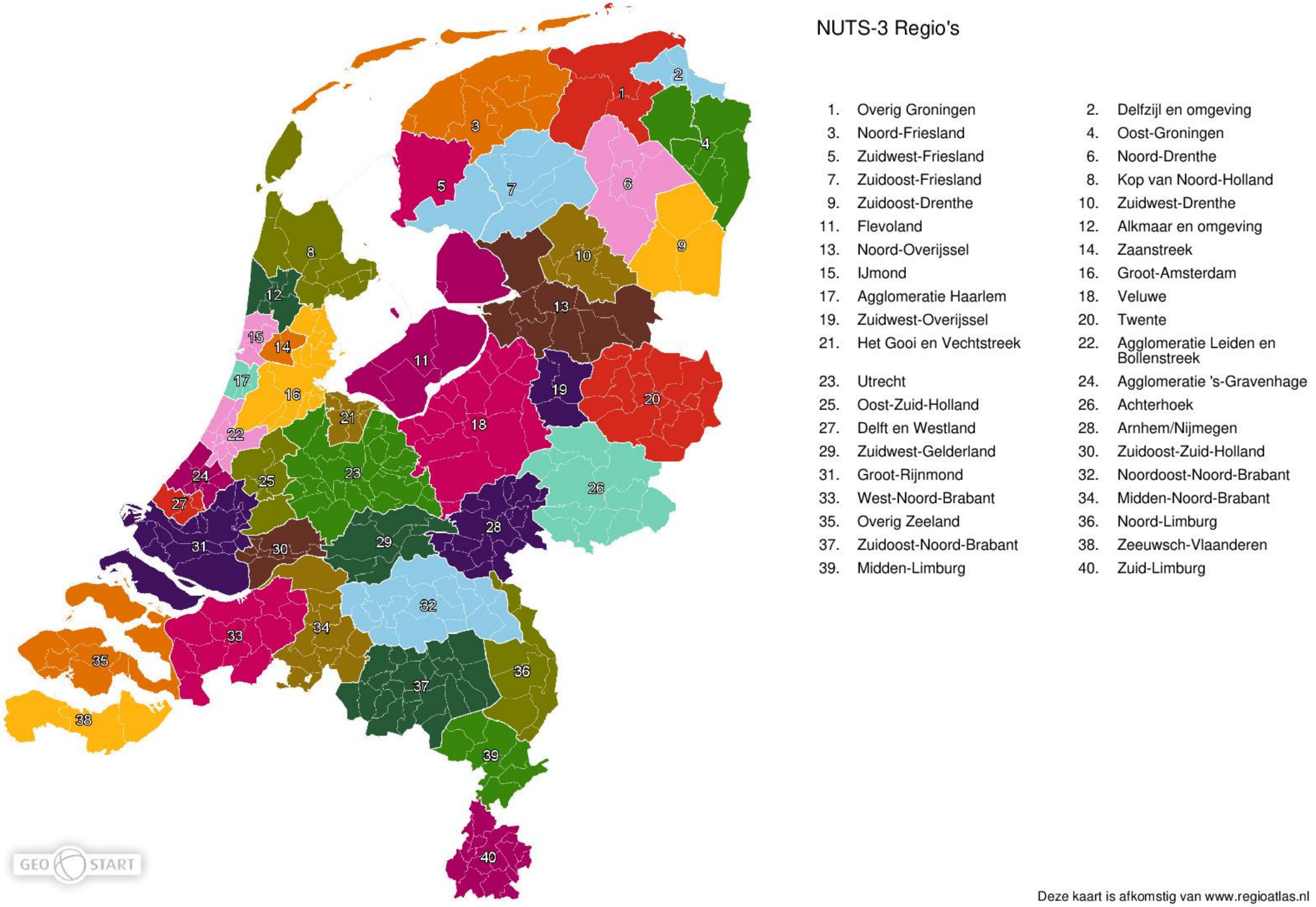
Vallin, J. & Meslé, F. (2005). Convergences and divergences: an analytical framework of national and sub-national trends in life expectancy. *Genus*, 61(1), 83-124.

Verheij, R. A. (1996). Explaining urban-rural variations in health: A review of interactions between individual and environment. *Social Science & Medicine*, 42(6), 923-935.

Verheij, R. A., de Bakker, D. H., Groenewegen, P. P., van de Mheen, H.D. & Mackenbach, J. P. (1998). Urban-rural variations in health in the Netherlands: does selective migration play a part? *Journal of Epidemiology & Community Health*, 52(9808), 487–493.

Zhang Y., Engelman M., & Agree E.M. (2013). Moving Considerations: A Longitudinal Analysis of Parent-Child Residential Proximity for Older Americans. *Research on Aging*, 35(6), 663–687.

Appendix 1: The 40 COROP regions of the Netherlands



Source: regioatlas.nl