Do Social Network Categories Affect Cognitive Functioning of the Elderly? Assessing Heterogeneities in SHARE.

Jay Kulkarni

April 05, 2025

Abstract

Previous studies have identified that mitigating social isolation among the elderly is a way to strengthen their cognitive function and reduce age-related degenerative diseases like dementia. With widespread ageing in Europe, it is important to understand the association between social networks of the elderly and their cognitive function. This thesis uses the 'Survey of Health, Ageing and Retirement in Europe' (SHARE), a large cross-sectional data with 70,000 participants across Europe and West Asia to examine this association. A novel clustering method is used to classify respondents into social network categories based on network size, geographical and emotional proximity, relationship type, and contact frequency. Both subjective (self-reported memory) and objective (verbal fluency) measures of cognitive function are analyzed. In agreement with the literature, larger network size is associated with better cognition. While network size remains important, richer characterization of the network using clustering helps us better gauge the quality of people's networks, and the clusters show greater magnitudes of association with cognitive function. Furthermore, considerable heterogeneity is observed in verbal fluency across the four regions, with the North and West generally showing better cognitive performance compared to the East and South. This suggests that regional histories and socio-cultural factors play a significant role in cognitive function, even when the quality of social network is the same. The thesis highlights the importance of considering the quality of social networks and the influence of regional contexts in understanding cognition in Europe.

Acknowledgements

I would like to thank Professor Arnstein Aassve for introducing me to demography and for his close involvement with this project. I also thank Professor Jari Saramäki for familiarizing me with social network clustering. I am grateful to my family for believing in me and supporting me throughout my journey in Bocconi University.

Data Acknowledgement

This thesis uses data from SHARE Wave 9 (DOI: 10.6103/SHARE.w9.900). See Börsch-Supan et al. (2013) for 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, VS 2020/0313, SHARE-EUCOV: GA N°101052589 and EUCOVII: GA N°101102412. Additional funding from the German Federal Ministry of Education and Research (01UW1301, 01UW1801, 01UW2202), 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, BSR12-04, R01_AG052527-02, R01_AG056329-02, R01_AG063944, HHSN271201300071C, RAG052527A) and from various national funding sources is gratefully acknowledged (see www.share-eric.eu).

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Ageing populations are a reality in Europe. Healthy ageing requires attempts to strengthen cognitive function of the elderly, which can mitigate the onset of degenerative diseases like dementia. Cognition is especially relevant because it predicts “careers, partnering, reproduction, and health more strongly at older ages” (Bratsberg et al., 2023). New research has shown promising ways to reduce the risk of cognitive decline and dementia (Livingston et al., 2024). These include a focus on improving cognitive function in the childhood and adolescence (Bratsberg et al., 2024) and mitigating social isolation in people of age 65 and above (Livingston et al., 2017).

As people age, their social networks see a transformation. Socioemotional selectivity theory (SST) posits that as people perceive their future time as more limited (elderly population), they prioritize emotionally meaningful goals, which contribute to their emotional well-being (Jiang & Fung, 2019). This prioritization may involve deepening close relationships and savoring the present, rather than seeking new connections- seeking a smaller but higher quality social network. Changes in social relations in older adults can also occur due to migration of children, family, and friends; a decrease in social network size due to death of network members; or due to declining health, which can hinder social engagement (Cudjoe et al., 2018).

Social networks are important for the elderly because they rely more on their family, children and friends to support them with daily activities and care in times of illness. Compared to the young, older people are more dependent on their social network to meet their emotional and physical needs (Bahramnezhad et al., 2017). Acceptance in their social network also plays a role in improving their quality of life (Huguet et al., 2008).

A strand of literature has established that social networks are as important for physical health as nutrition, excercise, and other lifestyle factors. Holt-Lunstad et al. (2015), in their meta-analysis find that both perceived and objective social isolation pose early mortality risk comparable to smoking, exposure to air pollution, and sedentary living. They report that social isolation, loneliness, and living alone increase likelihood of mortality by 29%, 26%, and 32% respectively. Nguyen et al. (2024) examine a combination of eight lifestyle factors on mortality, including social connections. They report a hazard ratio of 0.21 compared to the group that adopts neither of the eight lifestyle factors-suggesting relatively higher years until mortality .

Another strand of literature has examined the association between various characterizations of social network and cognitive function of individuals. Nie et al. (2021) used network size, contact frequency and social activity participation against cognitive function measures that included verbal memory and verbal fluency from a dataset of 6,691 participants in Eastern Europe (HAPIEE). They found that larger network sizes and social activity participation were positively associated with cognitive function, while no such effect was found for contact frequency. Kuiper et al. (2016) performed a meta-analysis and found that weak social networks based on network size and social activity participation predicted 1.08 higher odds of cognitive decline. In agreement, Cunha et al. (2024), through another meta-analysis reported that formal social participation is associated with reduced cognitive decline with an odds ratio of 0.78.

Others have examined heterogeneity in the association between social networks and cognition by gender and region. Lee et al. (2020) used a Korean dataset with 501 participants and concluded that the influence of social networks on cognitive function differs by gender. Wolfova et al. (2024) used the panel SHARE dataset with 66,670 respondents to examine whether cognitive decline itself, irrespective of social networks differed by gender. They did not find evidence of sex differences in cognitive decline among the elderly in Europe, but found heterogeneity across birth cohort and regions.

The literature has identified that stronger social networks in the elderly are an important way to reduce cognitive decline. There is an agreement on the positive association between social networks and cognitive function. Further, there have been attempts to better characterize social networks by using measures of contact frequency and social activity participation besides network size. This thesis aims to contribute to the literature by using a large dataset of about 70,000 participants across 28 countries in Europe and West Asia to examine the association between social networks and cognitive function. It builds on the literature that has tried to better understand the 'quality of social networks' using a novel clustering method. Parameters like network size, geographical and emotional proximity, relation, and contact frequency are used to classify survey respondents into social network categories. The association of belonging to a social network category on both subjective and objective measures of cognitive function are examined. Lastly, it assesses whether regional heterogeneities are significant for performance on cognitive function despite having the same quality of social network.

The rest of the thesis is structured as follows. Chapter 2 provides information about the data and specific variables used. Chapter 3 sheds light on the cognitive function module. Chapter 4 describes the clustering method and identifies what every social network category means. Chapter 5 presents a baseline stylized fact- the association between network size and cognitive function in Europe. Chapter 6 presents the findings on the association between social network categories and cognitive function, along with findings on regional heterogeneities. Chapter 7 concludes.

Chapter 2

Data

The empirical analysis makes use of individual level data from the 'Survey of Health, Ageing and Retirement in Europe' (SHARE). SHARE is a large bi-annual panel survey conducted across 28 countries in Europe. The data is publicly available and can be downloaded for free upon registering as a user. Each survey round is called a 'wave' in SHARE terminology. The survey for wave 1 was conducted in 2004-05. Subsequent waves have been roughly 2 years apart. This thesis uses data from wave 9, the latest survey round carried out in 2021-22, effectively using only a cross-section from SHARE.

The target population of this survey is all persons of age of 50 and above, who reside in one of the surveyed countries. For wave 9, this implies that the respondents were born in 1971 or earlier. The data collection was done using face-to-face interviews using a 'computer-assisted personal interviewing' tool. SHARE contains several thematic question modules, many of which require participants to take physical tests- this makes the in-person interviews necessary. The languages in which the survey was conducted were the main languages of the respective country. For instance, in Switzerland, the interviews were conducted in German, French, and Italian.

A highlight of SHARE is its large sample size of almost 70000 individuals spread across 28 countries. This provides power to the analysis that teases out spatial heterogeneities in cognitive function of the elderly. Since examining patterns by individual countries can get tedious, the countries are also grouped into regions according to the

M49 standard of the United Nations for Europe and West Asia[1]. This simple classification is based on geographic regions and their sub-regions. Figure 2.1 is a map that shows the reader how the 28 countries are decomposed into the 4 regions- North, South, East, and West. Table 2.1 adds further information about the sample size of each country and the region to which the country is assigned. The reader can confirm that each region has an adequately large sample size, ranging from 14150 to 20466 observations.

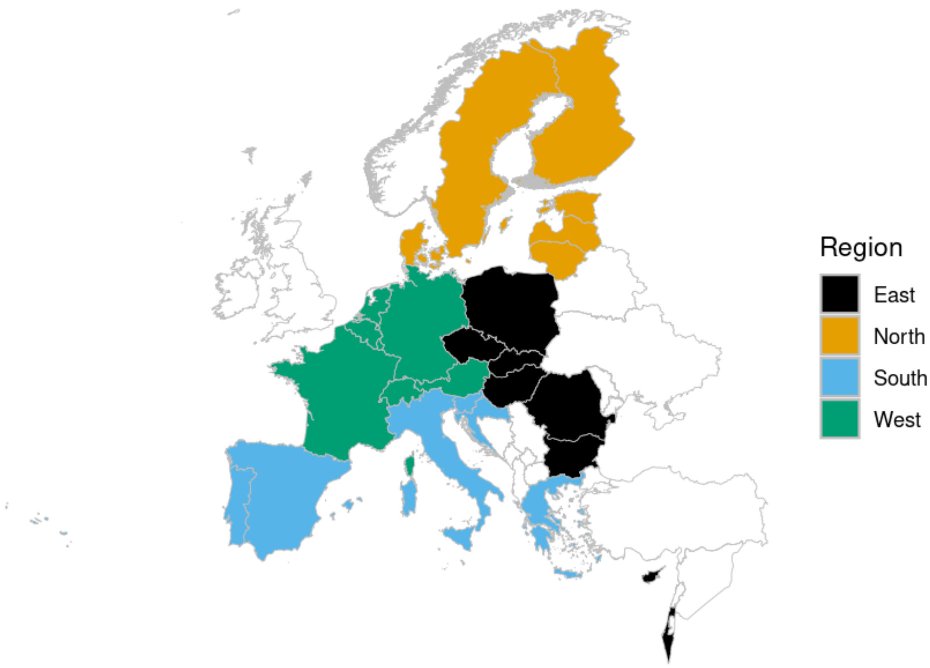

Figure 2.1: Map: Decomposition of SHARE's 28 countries into 4 regions

Table 2.1: Sample Size by Country and Region in SHARE Wave 9

North		South		East		West	
country	n	country	n	country	n	country	n
Denmark	2348	Croatia	4687	Bulgaria	831	Austria	3405
Estonia	4486	Greece	3112	Cyprus	736	Belgium	4497
Finland	1758	Italy	3623	Czech Republic	3319	France	2909
Latvia	1716	Malta	875	Hungary	1828	Germany	4467
Lithuania	1414	Portugal	1638	Israel	744	Luxembourg	810
Sweden	2428	Slovenia	4444	Poland	4828	Netherlands	2066
		Spain	2087	Romania	1492	Switzerland	1844
				Slovakia	1055		

[1]Cyprus and Israel are the only two countries from West Asia; the rest are from Europe.

Three modules of SHARE are relevant for this thesis- social networks (SN), cognitive function (CF), and generated variables (GV). The SN module records up-to seven people whom respondents consider confidants. For each person in the respondents' SN, detailed information about their characteristics is recorded. This includes *relation with respondent* (e.g. child/partner/non-kin), *gender*, *reported emotional closeness*, *contact frequency* (e.g. daily/monthly), and *geographical proximity*. Thus, besides having different social network sizes (zero to seven), individuals can belong different SN categories based on other characteristics of their SN constituents.

The cognitive function (CF) module contains few objective and subjective measures of respondents' cognitive functioning. This analysis uses one objective measure, *verbal fluency*, and one subjective measure, self-reported *memory*. Verbal fluency is a score between zero and hundred, based on the number of animals the respondent can name in a minute. The self-reported score of memory is on a five-point scale, where one is recoded to be poor and five is excellent. Other available measures from this module were not considered because they were binary variables with little variation.

The generated variables (GV) module provides us with other characteristics of the respondent that are useful control variables. These are *age* of the respondent, *years of education* received by the respondent, and their *gender* (sex at birth).

Chapter 3

Cognitive Function Module

This chapter describes the two variables used from the Cognitive Function (CF) module, which are the dependent variables in this thesis.

The verbal fluency test is a common neuropsychological test administered in research that measures cognitive function. The task is a simple one- to name as many animals as possible in one minute. It activates multiple cognitive processes in the respondents despite being simple (Sutin et al., 2019). The respondents engage with their verbal knowledge to produce examples of animals. More importantly, they have to keep track of examples already stated to avoid repetition. This involves monitoring and controlling their thoughts, which is key to cognition. By measuring respondents' executive function, the verbal fluency test provides information on how well they function cognitively (Sutin et al., 2019; Wolfova et al., 2024).

Verbal fluency is an objective measure of cognitive function because respondents' score is proportional to the animals they were able to name within the time limit. Figure 3.1 highlights that the bulk of repondents have a score between zero and fifty. In this range the distribution has a desired bell shape and benefits from a large sample size. Figure 3.2 highlights regional heterogeneity in the verbal fluency score. The South and East have relatively greater proportion of their samples with a low verbal fluency score (< 15). Consistent with the literature, North and West perform relatively better on this measure of cognition. They have a greater proportion of their sample with a high verbal fluency score (> 25).

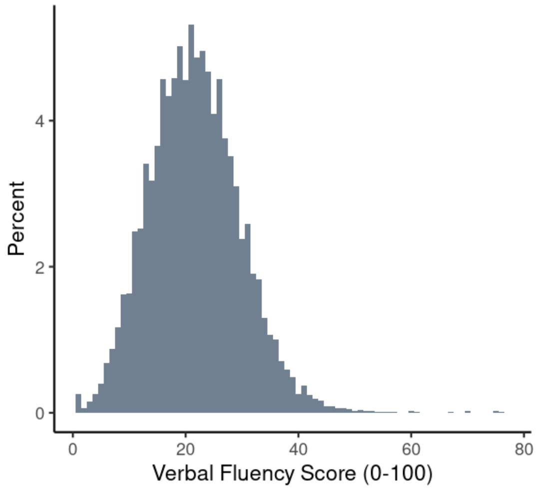

Figure 3.1: Verbal Fluency Distribution

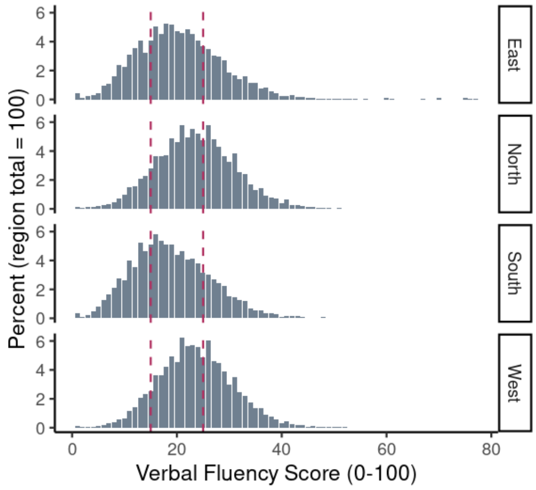

Figure 3.2: Verbal Fluency by Region

Memory is another typical measure of cognitive function. Here, the memory score is self-reported. It is a five-point score recoded such that one is poor and five is excellent memory. Having a self-reported measure has both advantages and limitations. Verdelho et al. (2011) have found self perceived memory complaints to be a strong predictor of Alzimer's disease, a type of dementia. This implies that people reporting poor memory may actually be capturing poor cognitive function, with the caveat that this may not scale up as the memory score improves. There may be respondents who over- or under-estimate their memory, and this may be region-specific. The limitation is that there is no way to test this. On the brighter side, the large sample size and a normal distribution of the memory score (Figure 3.3) are promising.

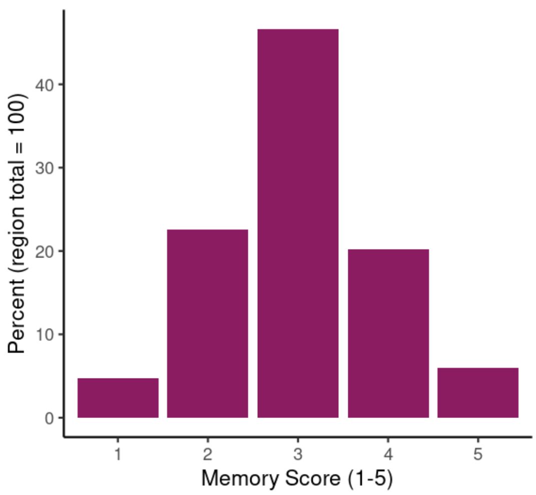

Figure 3.3: Memory Distribution

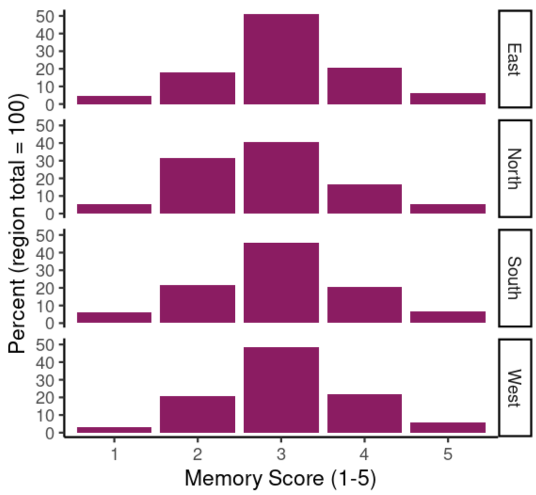

Figure 3.4: Memory by Region

Chapter 4

Social Network Analysis: Clustering

This chapter uses the social network (SN) module from SHARE to assign SN categories to each respondent and interprets them. The rich SN module provides several parameters that describe the social network of respondents. These are used as input variables for hierarchical clustering. This exercise classifies each respondent into a SN category. Finally, each SN category (cluster) is interpreted using summary statistics about respective clusters.

4.1 SN Parameter Description

Twenty parameters were used to classify respondents into their respective clusters. They are as follows. The total SN size is captured in the variable *n_size*. The number of people in respondents' SN can range from zero till seven. The number of kin and non-kin members of the SN are captured in *n_kin* and *n_nonkin*. The non-kin members are most likely friends of the respondents. The number of kins are subdivided into the number of children, parents, in-laws, and partner in the SN (*n_child, n_parent, n_inlaw, n_partner*). The gender decomposition of the SN is captured using the number of males and females (*n_male, n_female*). Note that this is the sex at birth of every person listed in respondents' SN. Further, the reported emotional closeness of every SN member is noted as 'very close' or 'average close'. The total number of very/average close members in the SN is captured in *n_close_very* and *n_close_avg*. The frequency with which respondents contact members from their SN is captured in four options: daily, in weeks, monthly, and rarely. The total number of SN members contacted with respective frequencies is captured in *n_daily, n_weeks, n_monthly, n_rarely*. The module also captures geographical proximity of respondents' SN members. The number of SN members who live in the same house/building are captured in *n_prx_location*. Those that live up-to 5 kilometers away are captured in *n_prx_city*. The number of SN members that live between 5 and 100 kilometers away are captured in *n_prx_district*. Those that live between 100 and 500 kilometers are captured in *n_prx_region*. Lastly, the number of SN members who live more than 500 kilometers away are captured in *n_prx_far*.

Each parameter can take values between zero and seven per respondent because seven is the maximum network size possible in the SHARE survey. However, distributions of parameters like *n_size* and *n_parent* are likely to be centered around different means. All parameter values were standardized before feeding them into the clustering algorithm so that their weights were fixed.

4.2 Methodology: Hierarchical Clustering

The seventy-thousand respondents have unique social networks based on twenty parameters/dimensions. This is the starting point of the hierarchical clustering- each observation is a separate cluster. Iteratively, closest observations (clusters) are combined into bigger clusters until a stopping criterion is reached or only one big cluster remains. We aim to simultaneously minimize the dissimilarity (maximize the similarity) between observations within each cluster, and maximize the distances (differences) between clusters. This bottom-up agglomerative method produces a series of nested clusters. The hierarchy is graphically represented in a dendogram (Figure 4.1) which is an inverted tree that describes the order in which observations are merged. The grey dotted line represents the stopping criterion. This was set ex-post, after trial and error, with two objectives: to maintain large distance between clusters while having a reasonable number of final clusters. Too few clusters might be combining meaningfully different observations, while too many clusters will be tedious to interpret. The clustering exercise here has resulted in ten clusters.

An important detail in clustering is how the similarity or difference of pairs of clusters is determined while doing the agglomeration. The distance measure used here is squared Euclidean distance. The general formula to calculate Euclidean distance is:

$$d_{ij} = \sqrt{\sum_{p=1}^{p}(x_{ip} - x_{jp})^2}$$

where x_{ip} and x_{jp} are scores on variable x_p for observation i and j respectively. In the p variable situation, the sum is of the squared distance between observations on all p variables.

A way to ensure that within-cluster similarity is maintained is to minimize the total within-cluster variance. Ward's minimum variance criterion or Ward's method does exactly this. At each stage, the pair of clusters that is joined is the pair whose agglomeration will result in the minimum increase in total within-cluster variance. After choosing the clustering method (Ward's) and the distance measure (squared Euclidean), the implementation of the hierarchical clustering was done using Python on a High-performance Computing Cluster. The algorithm is computationally heavy because it produces a proximity matrix at every step to join clusters according to Ward's method.

Table 4.1: Size of each SN Cluster

1	2	3	4	5	6	7	8	9	10
3475	3464	2474	5504	3200	3594	11893	19350	8661	7832

4.3 Interpretation of SN clusters

The outcome of the hierarchical clustering is that every respondent is assigned a SN cluster. It is necessary to interpret and give meaning to these clusters. Table 4.2 reports the mean of every parameter used in the clustering exercise. The clusters can be described as follows:

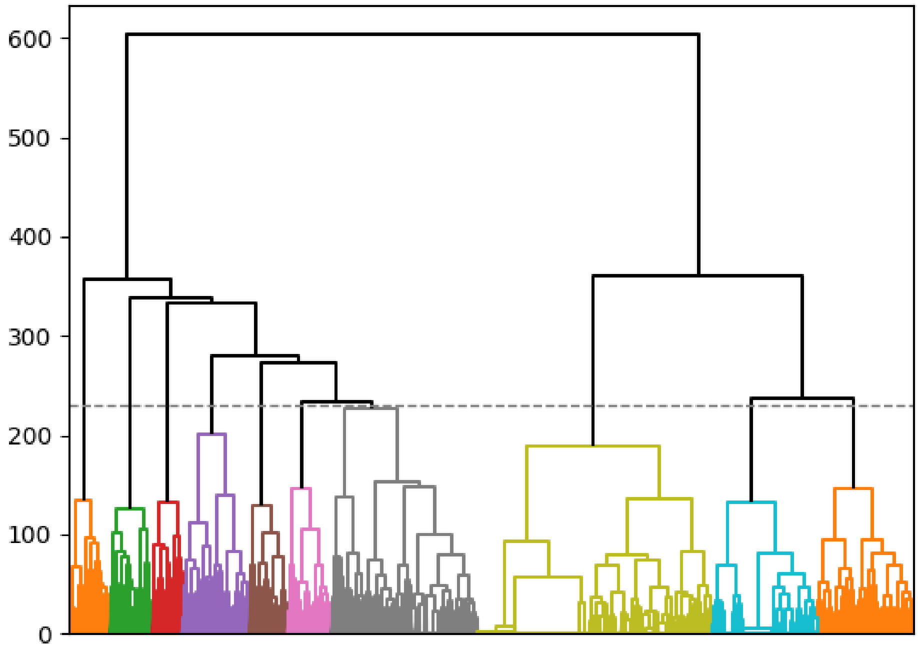

Figure 4.1: Dendogram from Hierarchical Clustering using Ward's method

- C1: Kin dominated SN including parent (parent, child, partner, friend)
- C2: Kin-dominated SN including in-laws
- C3: Large SN of family and friends, including not close people who are rarely contacted
- C4: Equal friends and family (2 each) within the district; includes not close people contacted in weeks/month
- C5: A child/spouse at home and two others (kin + friend) over 500 km away
- C6: A child/spouse at home and two others (kin + friend) 100-500 km away
- C7: Spouse/partner, two children and a friend living within the district
- C8: Only Spouse/cohabiting partner (very close, kin is partner, live in the same house)
- C9: Isolated respondents (smallest SN size of 0.9)
- C10: Children and a friend from same city/district

Out of these ten clusters, three need special mention because they are different on multiple dimensions. C9 is the base cluster that will be used for comparison in further analysis. This is the isolated cluster, with the smallest network size. Most clusters are better than C9 on some dimension. C8 is important because the largest number of participants are classified under it (about 19000). It can be characterized as the cluster with only spouse or cohabitating partner living under the same roof. Ex-ante too, we expect many elderly respondents to be living with their partner, and the partner being the sole source of interactions. Next is the unique cluster, C10, which can be characterised as children and a friend from the same city/district.

As can be seen from the dendogram, clusters 1 till 7 can only be differentiated with minute differences. C1, a family-dominated cluster is one of the best clusters because of large SN size (4 on average), diversity of relations (parent, partner, child, and friend), and both high emotional and geographical proximity. C2 is very similar to C1, with the exception of parent being replaced by in-law in the network. Cluster C3 is large, and consists of both friends and family, but the highlight is that it includes people who are rarely contacted. C4 is very similar to C3, with a difference in contact frequency- rarely contacted people are replaced by those contacted in weeks and months. C5 is characterized by children, spouse and a friend with minimum one person living over 500 kilometers away. C6 is similar to C5, with a difference in geographical proximity- it has minimum one person living in the region (100-500 km away). C7 is similar to both C5 and C6, again with a difference in geographical proximity- all persons in the network live within the district. Also note that C1 and C7 are very similar (see Table 4.2) with the difference that C7 has two children in the network, while C1 only has one child.

Figure 4.2 plots the decomposition of each cluster by region. Note that the Table 4.1 already shows great variation in the number of observations per cluster. The number of observations across the four regions have a slight variation too. It is therefore unsurprising to find that each cluster is not decomposed equally among the four regions. Clusters C3 and C4 are over-represented by the West, while clusters C8 and C9 are over-represented by the South and East. Underlying demographic changes are driving important differences in the regional decomposition of the social network clusters.

Table 4.2: Interpreting Social Network Clusters using Mean Parameter Values

	1	2	3	4	5	6	7	8	9	10
n_size	4.0	4.2	4.2	4.0	3.5	3.3	4.1	1.7	0.9	2.6
n_kin	3.3	3.7	2.6	2.3	2.6	2.6	3.3	1.6	0.7	1.7
n_nonkin	0.8	0.5	1.6	1.7	0.8	0.7	0.8	0.1	0.2	0.9
n_child	0.9	1.4	1.1	1.0	1.4	1.5	1.9	0.5	0.5	1.2
n_parent	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
n_inlaw	0.1	1.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
n_partner	0.7	0.5	0.5	0.6	0.5	0.5	0.9	1.0	0.0	0.0
n_male	1.1	1.1	1.2	1.3	0.8	0.7	1.0	0.5	0.1	0.3
n_female	2.0	1.7	1.9	1.8	1.3	1.2	1.3	0.7	0.2	1.0
n_close_very	3.6	3.7	2.9	2.5	3.0	3.0	3.9	1.5	0.8	2.2
n_close_avg	0.5	0.5	1.3	1.5	0.5	0.3	0.2	0.2	0.1	0.4
n_daily	0.5	1.0	0.3	0.4	0.6	0.6	0.9	0.2	0.3	0.8
n_weeks	1.4	2.1	1.6	2.1	2.0	2.0	2.2	0.3	0.3	1.7
n_monthly	0.1	0.1	0.5	0.8	0.2	0.2	0.0	0.0	0.0	0.0
n_rarely	0.0	0.0	1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0
n_prx_location	0.9	1.2	0.7	0.7	0.7	0.7	1.0	1.2	0.3	0.2
n_prx_city	0.7	1.2	1.0	1.2	0.6	0.5	1.0	0.2	0.3	1.2
n_prx_district	0.8	1.1	1.5	1.4	0.7	0.6	1.0	0.2	0.2	1.0
n_prx_region	0.2	0.2	0.5	0.3	0.2	1.5	0.1	0.0	0.0	0.0
n_prx_far	0.1	0.1	0.2	0.0	1.3	0.0	0.0	0.0	0.0	0.0

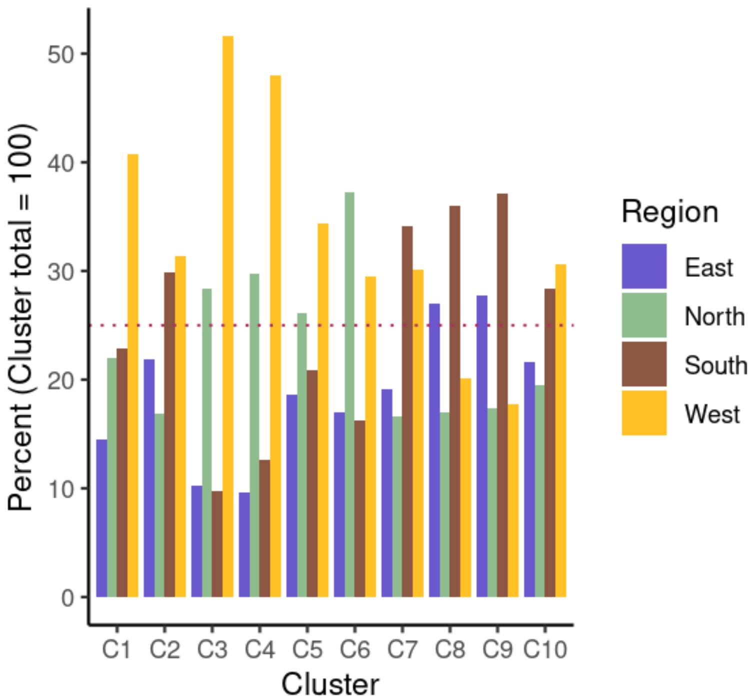

Figure 4.2: Cluster decomposition by Region

Chapter 5

Social Network Size Descriptives

This chapter uses just one parameter from the SN module, the network size of each respondent, and presents some descriptive findings.

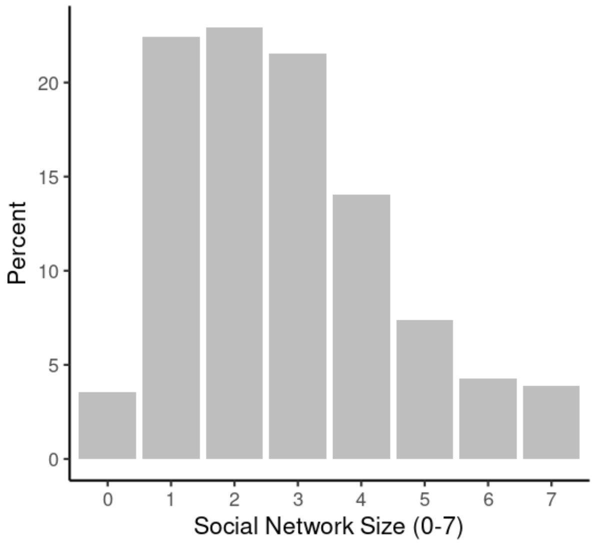

Figure 5.1: 81% of the respondents report an SN size between 1 and 4

Figure 5.2 shows the distribution of SN size by region. In the east, respondents report the highest frequency of having one person in their SN, and it monotonically decreases until seven. In the west, we see resemblance of a normal distribution. It is the region where respondents report the highest number of five, six, and seven people in their SN. In the north, the bulk of the SN size distribution is between one and three, decreasing thereafter. The south follows the same trend as north, but with higher levels throughout the distribution.

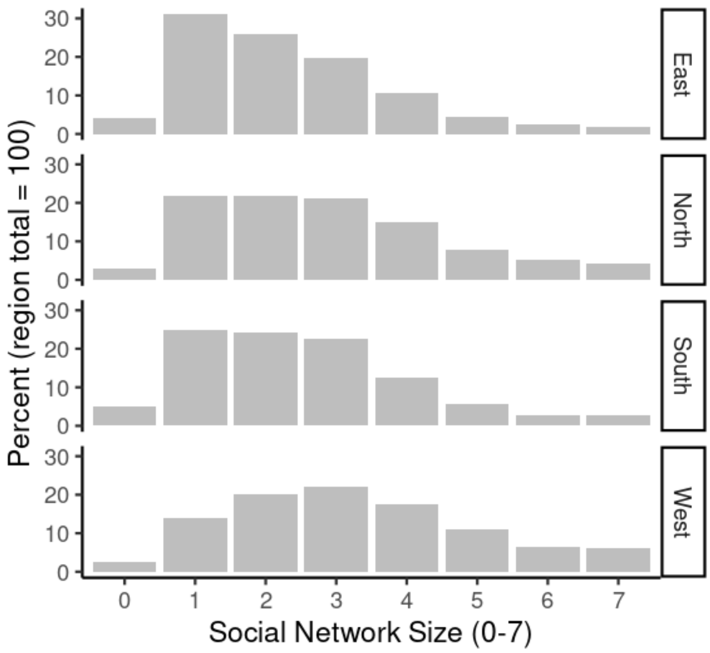

Figure 5.2: West has the most normal distribution, rest concentrated between 1 and 3

5.1 Does CF improve with larger SN size?

Now that we have seen the distribution of SN size among respondents in the pooled sample and by regions, it is important to understand whether larger network size is associated with better cognitive functioning. Simple OLS regressions were run with and without country fixed effects to test this association. The dependent variables were the two measures of cognitive function, memory and verbal fluency. The independent variable was network size (SN size). Age, education (years), and gender (female) were used as controls. It can be represented as follows:

$$CF_i = \beta_0 + \beta_1 SN\_size_i + \gamma_1 Age_i + \gamma_2 Education_i + \gamma_3 Female_i + u_i$$

where CF_i measures the cognitive function of individual i (memory or verbal fluency), β_0 is the constant, SN_size_i stands for the network size of individual i, and Age_i, $Education_i$ and $Female_i$ are the controls indicating age, education and gender of individial i. The main coefficient of interest is β_1, which reveals the impact of an additional person in the respondent's social network on the five-point memory score or the hundred-point verbal fluency score.

Table 5.1 presents the results of the above base specification for Memory in column (1) and for Verbal Fluency in column (3). Columns (2) and (4) add country fixed effects to the base specification. A unit increase in network size improves the memory score by 0.03 points, which is significant at the 1 percent level (column 1). A unit increase in network size improves the verbal fluency score by 0.73 points, which is also significant at the 1 percent level. Adding country fixed effects to the base specification reduces the magnitude of the association, but the trend remains the same- a unit increase in network size has a positive and significant increase in measures of cognitive function. Additionally, a unit increase in age has a negative and significant (1 percent level) association with cognitive function. Education has a positive and significant (1 percent level) association with cognitive function. Lastly, consistent with Wolfova et al. (2024), gender has no significant role to play in respondent's cognitive function.

Table 5.1: Memory and Verbal Fluency improve with Larger SN

Dependent Variables:	Memory		Verbal Fluency	
Model:	(1)	(2)	(3)	(4)
Variables				
Constant	4.316***		28.87***	
	(0.0290)		(0.2468)	
SN Size	0.0285***	0.0190***	0.7333***	0.5316***
	(0.0020)	(0.0042)	(0.0173)	(0.0478)
Age	-0.0245***	-0.0252***	-0.2229***	-0.2371***
	(0.0004)	(0.0012)	(0.0030)	(0.0099)
Yrs of Edu	0.0278***	0.0307***	0.4691***	0.3625***
	(0.0008)	(0.0035)	(0.0071)	(0.0291)
Female	-0.0080	0.0080	-0.1108*	-0.1413
	(0.0068)	(0.0120)	(0.0580)	(0.1189)
Fixed-effects				
Country		Yes		Yes
Fit statistics				
Observations	68,145	68,145	67,323	67,323
R^2	0.10132	0.16032	0.18566	0.26932
Within R^2		0.10889		0.16384

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Chapter 6

Findings

Chapter 5 concluded that a larger network size is associated with a better performance on both measures of cognitive function; memory and verbal fluency. Using SHARE's rich characterization of social networks, it is possible to assign social network categories to respondents- this was done in Chapter 4 using clustering. This chapter examines whether belonging to a different SN cluster affects cognitive function of the elderly differently. Moreover, we examine regional heterogeneities by assessing whether belonging to the same SN cluster in a different region changes the direction and magnitude of association with cognitive function.

6.1 Base Regressions

We test the association between SN clusters and CF by running the following regression:

$$CF_i = \beta_0 + \beta_1 Cluster_i + \gamma_1 Age_i + \gamma_2 Education_i + \gamma_3 Female_i + \gamma_4 Region + u_i$$

where CF_i measures the cognitive function of individual i (memory or verbal fluency), β_0 is the constant, $Cluster_i$ is a factor variable, indicating the cluster to which individual i belongs, with C9 being the reference cluster. Age_i, $Education_i$ and $Female_i$ are the controls indicating age, education and gender of individual i. $Region_i$ is also a factor variable indicating the respondent's region, with east being the reference region. The main coefficient of interest is β_1 for each cluster, which reveals the impact of belonging to the particular cluster on the five-point memory score or the hundred-point verbal fluency score.

Table A.1 has three columns with memory as the dependent variable, where column (1) excludes the region controls; column (2) excludes the SN clusters; and column (3) is the full specification. Here the coefficients on the controls are aligned with those in Table 5.1. Age is negatively and significantly associated with memory, Education (years) is positively and significantly associated with memory, and Gender has no effect. As most clusters are better than C9, we see a positive and significant association of all clusters with CF. What is more interesting is that the smallest positive coefficient is that on C3 (0.1039), which is multiple times higher than the coefficient on SN size in Table 5.1 (0.0190). This is the first key finding- the network size is not the only criterion that characterizes an individual's social network. By including emotional and geographic proximity, contact frequency, and relation with respondent, we have a more comprehensive understanding of the respondent's social network, which also show greater magnitudes of association with cognitive function (memory in this case).

Although the direction of association between all the SN clusters and memory is positive, the largest coefficient is double that of the smallest (0.1958 vs 0.1063), which warrants careful examination of Table A.1. Clusters C6, C7, and C1 have the largest positive coefficients. These clusters have an advantage by having large network sizes and a balance of kin and non-kin members in the network. Clusters C10, C8, and C3 have the smallest positive coefficients. Cluster C3, despite having a large network size is disadvantaged by having more than one person who is rarely contacted and has lower emotional proximity with the respondent. Clusters C8 and C10 are disadvantaged by having the second and third lowest network size respectively. When we look at regions, south is not significantly different from the east. North and West are negative and significant at the 1 percent level, indicating a worse performance on the five-point memory score. This is surprising, but the memory score is self-reported and it is possible that people from the North and West are systematically holding themselves to higher standards of memory.

Table A.2 has the hundred-point verbal fluency as the dependent variable, and the three columns are in similar order- column (1) excludes region controls; column (2) excludes SN clusters; and column (3) is the full specification. Again, coefficients on the controls are comparable with those in Table 5.1. Age and Education (years) are positively and negatively associated with verbal fluency respectively, while Gender has no effect. Compared to the coefficient of 0.53 on SN Size in Table 5.1 column (4), the coefficients on the clusters are much larger (1.33 to 2.78). This reiterates that the social network of respondents characterized along several dimensions shows higher magnitudes of association with the verbal fluency score.

In Table A.2, Clusters C7, C6, and C4 have the largest positive coefficients which are significant at the 1 percent level. The advantages of belonging to these clusters are a large network size, and a balance of family members and friends. Clusters C8, C10, and C5 have the smallest positive coefficients. Cluster C5 is disadvantaged in terms of geographical proximity- it has at least one person living more than 500 kilometers away. Clusters C8 and C10 are disadvantaged because of having the second and third smallest network sizes respectively. Interestingly, all regions are significantly different when compared to the east, with varying signs and magnitudes. North and West are positive and significant at the 1 percent level, indicating a better performance on the hundred-point verbal fluency score. The South has a significantly lower verbal fluency score compared to the East. The verbal fluency is an objective measure, and these regional differences are aligned with the literature.

6.2 Within-Cluster Regional Heterogeneity

This section examines whether having the same social network characterization, but being in a different region has differential effects on cognitive function. This is done by adding a Cluster-Region interaction to the base regression in Section 6.1. Those results are presented in Table B.1. For easier interpretation of the interaction terms, Table B.2 presents the marginal effect, where the interaction between cluster C9 and region East is the baseline. The coefficients and the standard errors from Table B.2 are visually represented in Figures 6.1 and 6.2.

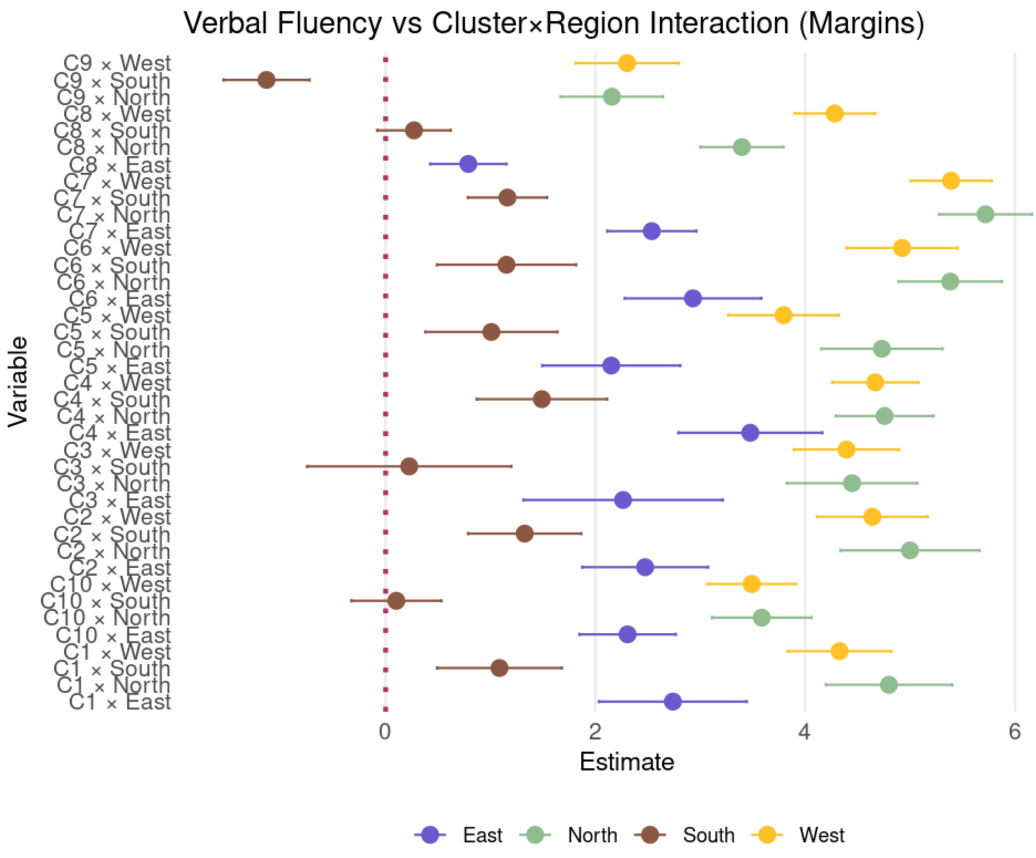

Figure 6.1: Cluster-Region interaction with Verbal Fluency as dependent variable

Figure 6.1, which plots the coefficients of the cluster-region interaction (margins, with C9 $\times$ East as reference) against verbal fluency, has a clear pattern. Within each cluster, North and West perform significantly higher on the verbal fluency score (controlling for age, education and gender), followed by East, and then South. Regional heterogeneities are a combination of the region's history, its culture, and institutions. It is possible that the most isolated cluster (C9) has better cognition in the North due to strong institutional elderly care compared to the South, where elderly care is family-centric. Thus, when the norm in the South is to rely on close family for elderly care, the lack of family members in one's network may be disproportionately disadvantageous. One narrative emerging from these regional heterogeneities is the advantage of state-led and institutional welfare models over family-based welfare models. Similarly, the East may perform worse in cognition compared to the West for all social network characterizations because of different histories of political transition pre- and post-communist era. Nie et al. (2021) suggest that rapid socio-cultural changes in the East may have had lasting effects on the social fabric of Eastern Europe, when compared to countries with stable political regimes in Western Europe. Thus, it is reasonable to speculate that many such historical and cultural reasons drive the differences in cognitive function despite having similar social networks. While policymakers can concretely address welfare models, they can only identify and acknowledge the meaningful socio-cultural differences in regions.

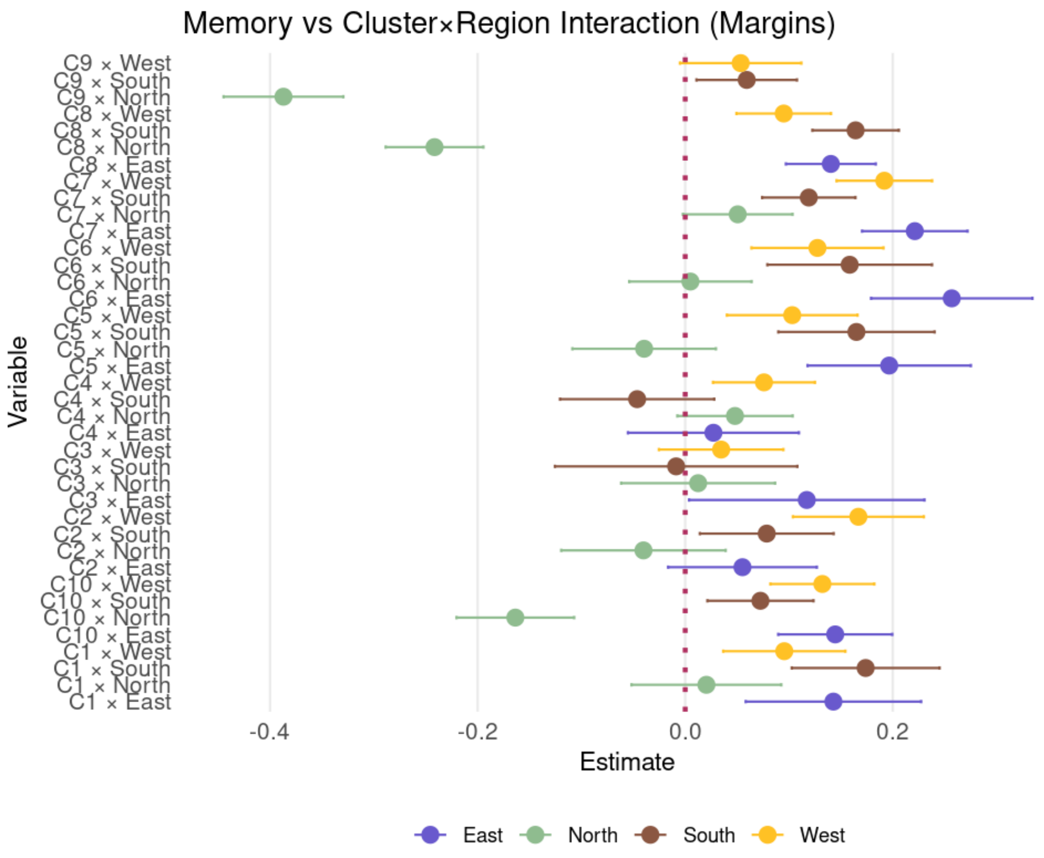

Figure 6.2: Cluster-Region interaction with Memory as dependent variable

Figure 6.2, with memory as the dependent variable again highlights that there are differences in cognitive function despite being characterized by the same social network. However, coefficients are not as distant from one another within-cluster, like in case of verbal fluency. There is also no monotonic pattern or ordering of region coefficients. We counter-intuitively see that for clusters C9, C8, and C10, the north performs significantly worse than the other regions. This could be attributed to the fact that memory is a self-reported measure, and region-specific benchmarks of memory may be leading respondents to over- or under-estimate their memory.

Chapter 7

Conclusion

The main findings of this thesis are as follows. Chapter 5 showed a significant positive association between respondents' social network size and their cognitive function. An additional member in respondents' social network increased the five-point memory score by 0.03 points and the hundred-point verbal fluency score by 0.73 points (Table 5.1). This can be treated as a baseline stylized fact for Europe given the large sample sizes for each country in the SHARE dataset. Chapter 6 added more nuance to this stylized fact. Using clusters based on a richer understanding of the respondents' social networks, by including parameters on geographical and emotional proximity, contact frequency, and relation with respondents, we saw greater magnitudes of association with cognitive function. The cluster coefficients in Table A.1 column (3) ranged between 0.10 and 0.20 for memory and Table A.2 column (3) ranged between 1.33 and 2.77 for verbal fluency- much larger compared to when only network size was used.

While clusters with smaller network sizes performed relatively worse on measures of cognitive function, other parameters were useful to distinguish the effect of clusters with similar network sizes. For instance, a cluster with geographically distant members in the network performed worse than a cluster with the same network size but closer geographical proximity. Similarly, another cluster that had members who were emotionally less close and were rarely contacted performed worse than a cluster with the same network size but better emotional proximity and contact frequency. This reiterates the need to measure 'quality of social networks' since better quality networks are associated with greater cognitive function, holding network size constant.

Considerable heterogeneity was observed within each cluster for the four regions when verbal fluency was the dependent variable. This heterogeneity was ordered - within each cluster, North and West had better cognition, followed by East, and then South. The pattern highlights that despite having the same social network characterization, cognition varies by region due to a combination of regional histories and socio-cultural factors. More concretely, isolated respondents in a region with family-led elderly care had worse cognition than isolated respondents in a region with state-led institutional elderly care. In regions where the norm is to rely on families for elderly care, people without family support are at higher risk of weak cognition. Similarly, regions exposed to rapid socio-cultural changes like Eastern Europe have worse cognition than regions with stable political regions, holding social network clusters constant.

A limitation of this thesis is that the effect of belonging to a social network category is not comparable between memory and verbal fluency. The scores could have been standardized for better comparability, but were deliberately left as is because of the belief that the subjective and objective scores were inherently capturing different things. Another limitation is that the memory and verbal fluency do not cover the entire breadth of cognitive function. One way to extend this research is to examine the same relationship with more objective measures of cognitive function. Another way to extend this research is by examining cognitive decline. It would be interesting to see whether social networks and their characterization remain relevant after accounting for baseline cognitive function in a panel dataset.

That said, the biggest strength of this thesis is its large sample across Europe, which makes its findings generalizable. The novel social network clustering showed that while network size remains important, richer characterization of the network helps us better gauge the quality of people's networks, and the clusters show greater magnitudes of association with cognitive functioning. Regions with institutional elderly care perform better than regions with family-led elderly care on cognition, especially for isolated respondents, who lack family support. Policy can aim to create institutional support in these regions for better wellbeing of the elderly.

References

Bahramnezhad, F., Chalik, R., Bastani, F., Taherpour, M., & Navab, E. (2017). The social network among the elderly and its relationship with quality of life. *Electronic physician*, *9*(5), 4306–4311. https://doi.org/10.19082/4306

Bergmann, M., Wagner, M., & Börsch-Supan, A. (2024). SHARE Wave 9 Methodology: From the SHARE Corona Survey 2 to the SHARE Main Wave 9 Interview. https://doi.org/10.6103/mv.w09

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

Bratsberg, B., Fjell, A. M., Rogeberg, O. J., Skirbekk, V. F., & Walhovd, K. B. (2024). Differences in cognitive function at 18 y of age explain the association between low education and early dementia risk. *Proceedings of the National Academy of Sciences*, *121*(41), e2412017121. https://doi.org/10.1073/pnas.2412017121

Bratsberg, B., Rogeberg, O., & Skirbekk, V. (2023). Chapter 22 - The relevance of cognition in the context of population ageing. In D. Bloom, A. Sousa-Poza, & U. Sunde (Eds.), *The Routledge Handbook of the Economics of Ageing* (pp. 396–409). Routledge. https://doi.org/10.4324/9781003150398-26

Cudjoe, T. K. M., Roth, D. L., Szanton, S. L., Wolff, J. L., Boyd, C. M., & Thorpe, J., Roland J. (2018). The Epidemiology of Social Isolation: National Health and Aging Trends Study. *The Journals of Gerontology: Series B*, *75*(1), 107–113. https://doi.org/10.1093/geronb/gby037

Cunha, C., Voss, G., Andrade, R., & Delerue-Matos, A. (2024). Is Formal Social Participation Associated with Cognitive Function in Middle-Aged and Older Adults? A

Systematic Review with Meta-Analysis of Longitudinal Studies. *Behavioral Sciences*, *14*(4), 262. https://doi.org/10.3390/bs14040262

Hoffman, J. I. (2015). Chapter 35 - Survival Analysis. In J. I. Hoffman (Ed.), *Biostatistics for Medical and Biomedical Practitioners* (pp. 621–643). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-802387-7.00035-4

Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and Social Isolation as Risk Factors for Mortality: A Meta-Analytic Review [PMID: 25910392]. *Perspectives on Psychological Science*, *10*(2), 227–237. https://doi.org/10.1177/1745691614568352

Huguet, N., Kaplan, M. S., & Feeny, D. (2008). Socioeconomic status and health-related quality of life among elderly people: Results from the Joint Canada/United States Survey of Health. *Social Science Medicine*, *66*(4), 803–810. https://doi.org/https://doi.org/10.1016/j.socscimed.2007.11.011

Jiang, D., & Fung, H. H. (2019). Chapter 6 - Social and Emotional Theories of Aging. In B. B. Baltes, C. W. Rudolph, & H. Zacher (Eds.), *Work Across the Lifespan* (pp. 135–153). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-812756-8.00006-2

Kuiper, J. S., Zuidersma, M., Zuidema, S. U., Burgerhof, J. G., Stolk, R. P., Oude Voshaar, R. C., & Smidt, N. (2016). Social relationships and cognitive decline: a systematic review and meta-analysis of longitudinal cohort studies. *International Journal of Epidemiology*, *45*(4), 1169–1206. https://doi.org/10.1093/ije/dyw089

Lee, S., Lee, S., Lee, E., Youm, Y., Cho, H. S., & Kim, W. J. (2020). Gender differences in social network of cognitive function among community-dwelling older adults. *Geriatrics & Gerontology International*, *20*(5), 467–473. https://doi.org/10.1111/ggi.13906

Livingston, G., Huntley, J., Liu, K. Y., Costafreda, S. G., Selbæk, G., Alladi, S., Ames, D., Banerjee, S., Burns, A., Brayne, C., Fox, N. C., Ferri, C. P., Gitlin, L. N., Howard, R., Kales, H. C., Kivimäki, M., Larson, E. B., Nakasujja, N., Rockwood, K., . . . Mukadam, N. (2024). Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. *The Lancet*, *404*(10452), 572–628. https://doi.org/10.1016/S0140-6736(24)01296-0

Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., Ballard, C., Banerjee, S., Burns, A., Cohen-Mansfield, J., Cooper, C., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Larson, E. B., Ritchie, K., Rockwood, K., Sampson, E. L., ... Mukadam, N. (2017). Dementia prevention, intervention, and care. *The Lancet*, *390*(10113), 2673–2734. https://doi.org/https://doi.org/10.1016/S0140-6736(17)31363-6

Nguyen, X.-M. T., Li, Y., Wang, D. D., Whitbourne, S. B., Houghton, S. C., Hu, F. B., Willett, W. C., Sun, Y. V., Djousse, L., Gaziano, J. M., Cho, K., & Wilson, P. W. (2024). Impact of 8 lifestyle factors on mortality and life expectancy among United States veterans: The Million Veteran Program. *The American Journal of Clinical Nutrition*, *119*(1), 127–135. https://doi.org/https://doi.org/10.1016/j.ajcnut.2023.10.032

Nie, Y., Richards, M., Kubinova, R., Titarenko, A., Malyutina, S., Kozela, M., Pajak, A., Bobak, M., & Ruiz, M. (2021). Social networks and cognitive function in older adults: findings from the HAPIEE study. *BMC Geriatrics*, *21*(1). https://doi.org/10.1186/s12877-021-02531-0

SHARE-ERIC. (2024). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 9. Release version: 9.0.0 [Data set]. https://doi.org/10.6103/SHARE.w9.900

Sutin, A. R., Stephan, Y., & Terracciano, A. (2019). Verbal fluency and risk of dementia. *International Journal of Geriatric Psychiatry*, *34*(6), 863–867. https://doi.org/https://doi.org/10.1002/gps.5081

Verdelho, A., Madureira, S., Moleiro, C., Santos, C. O., Ferro, J. M., Erkinjuntti, T., Poggesi, A., Pantoni, L., Fazekas, F., Scheltens, P., Waldemar, G., Wallin, A., & Inzitari, D. (2011). Self-Perceived Memory Complaints Predict Progression to Alzheimer Disease. The LADIS Study [PMID: 21841255]. *Journal of Alzheimer's Disease*, *27*(3), 491–498. https://doi.org/10.3233/JAD-2011-110494

Wolfova, K., Frycova, B., Seblova, D., Tom, S., Skirbekk, V. F., & Brennan Kearns, P. (2024). Sex differences in cognitive decline among middle-aged and older adults: a cohort study in Europe. *Age and Ageing*, *53*(4), afae078. https://doi.org/10.1093/ageing/afae078

Appendix A

Base Regression Results

Table A.1: Memory as Dependent Variable

Dependent Variable:		Memory	
Model:	(1)	(2)	(3)
Variables			
Constant	4.224^{***}	4.364^{***}	4.217^{***}
	(0.0321)	(0.0295)	(0.0326)
Cluster 1	0.1461^{***}		0.1588^{***}
	(0.0187)		(0.0187)
Cluster 2	0.1227^{***}		0.1215^{***}
	(0.0180)		(0.0179)
Cluster 3	0.0793^{***}		0.1039^{***}
	(0.0204)		(0.0205)
Cluster 4	0.0927^{***}		0.1219^{***}
	(0.0156)		(0.0158)
Cluster 5	0.1413^{***}		0.1582^{***}
	(0.0185)		(0.0185)
Cluster 6	0.1540^{***}		0.1958^{***}
	(0.0178)		(0.0178)
Cluster 7	0.1917^{***}		0.1895^{***}

	(0.0129)		(0.0129)
Cluster 8	0.1165***		0.1138***
	(0.0120)		(0.0120)
Cluster 10	0.1022***		0.1063***
	(0.0141)		(0.0140)
Age	-0.0239***	-0.0241***	-0.0234***
	(0.0004)	(0.0004)	(0.0004)
Yrs of Edu	0.0284***	0.0324***	0.0312***
	(0.0008)	(0.0009)	(0.0009)
Female	0.0058	0.0058	0.0076
	(0.0070)	(0.0068)	(0.0070)
Region North		-0.2334***	-0.2440***
		(0.0104)	(0.0105)
Region South		-0.0116	-0.0174*
		(0.0096)	(0.0096)
Region West		-0.0170*	-0.0315***
		(0.0095)	(0.0098)
Fit statistics			
Observations	68,145	68,145	68,145
R^2	0.10194	0.10810	0.11145
Adjusted R^2	0.10178	0.10802	0.11125

IID standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.2: Verbal Fluency as Dependent Variable

Dependent Variable:	Verbal Fluency		
Model:	(1)	(2)	(3)
Variables			
Constant	28.59***	31.36***	29.57***
	(0.2738)	(0.2482)	(0.2738)
Cluster 1	2.974***		2.268***
	(0.1587)		(0.1562)
Cluster 2	2.810***		2.471***
	(0.1529)		(0.1499)
Cluster 3	3.206***		2.007***
	(0.1736)		(0.1717)
Cluster 4	3.638***		2.492***
	(0.1328)		(0.1320)
Cluster 5	2.606***		1.980***
	(0.1576)		(0.1548)
Cluster 6	3.607***		2.765***
	(0.1511)		(0.1489)
Cluster 7	2.988***		2.782***
	(0.1102)		(0.1082)
Cluster 8	1.297***		1.331***
	(0.1033)		(0.1011)
Cluster 10	1.856***		1.465***
	(0.1200)		(0.1178)
Age	-0.2209***	-0.2360***	-0.2304***
	(0.0032)	(0.0030)	(0.0031)
Yrs of Edu	0.4693***	0.4112***	0.3937***
	(0.0072)	(0.0072)	(0.0072)
Female	0.0680	0.1137**	0.0996*
	(0.0596)	(0.0570)	(0.0583)
Region North		2.691***	2.445***
		(0.0870)	(0.0876)
Region South		-1.091***	-1.185***
		(0.0805)	(0.0801)
Region West		2.789***	2.439***
		(0.0801)	(0.0815)
Fit statistics			
Observations	67,323	67,323	67,323
R^2	0.18217	0.20646	0.21690
Adjusted R^2	0.18203	0.20639	0.21673

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Appendix B

Regression Results with Cluster-Region Interaction

Table B.1: Including Cluster-Region interaction

Dependent Variables:	mem	verbal
Model:	(1)	(2)
Variables		
Constant	4.208***	29.69***
	(0.0355)	(0.2986)
Cluster 1	0.1428***	2.738***
	(0.0431)	(0.3597)
Cluster 2	0.0552	2.474***
	(0.0365)	(0.3061)
Cluster 3	0.1171**	2.264***
	(0.0578)	(0.4847)
Cluster 4	0.0272	3.477***
	(0.0420)	(0.3504)
Cluster 5	0.1965***	2.151***
	(0.0401)	(0.3358)
Cluster 6	0.2568***	2.930***

	(0.0396)	(0.3319)
Cluster 7	0.2212***	2.537***
	(0.0259)	(0.2171)
Cluster 8	0.1403***	0.7887***
	(0.0221)	(0.1857)
Cluster 10	0.1445***	2.307***
	(0.0280)	(0.2347)
Region North	-0.3871***	2.157***
	(0.0294)	(0.2488)
Region South	0.0593**	-1.136***
	(0.0247)	(0.2094)
Region West	0.0534*	2.302***
	(0.0298)	(0.2517)
Age	-0.0235***	-0.2310***
	(0.0004)	(0.0031)
Yrs of Edu	0.0311***	0.3926***
	(0.0009)	(0.0072)
Female	0.0047	0.0844
	(0.0070)	(0.0583)
Cluster 1 $\times$ Region North	0.2647***	-0.0965
	(0.0580)	(0.4845)
Cluster 2 $\times$ Region North	0.2917***	0.3685
	(0.0562)	(0.4711)
Cluster 3 $\times$ Region North	0.2824***	0.0265
	(0.0705)	(0.5906)
Cluster 4 $\times$ Region North	0.4078***	-0.8767**
	(0.0525)	(0.4394)
Cluster 5 $\times$ Region North	0.1510***	0.4233
	(0.0552)	(0.4624)
Cluster 6 $\times$ Region North	0.1353***	0.2954
	(0.0516)	(0.4327)
Cluster 7 $\times$ Region North	0.2164***	1.025***

	(0.0398)	(0.3341)
Cluster 8 × Region North	0.0053	0.4509
	(0.0353)	(0.2973)
Cluster 10 × Region North	0.0789*	-0.8784**
	(0.0426)	(0.3576)
Cluster 1 × Region South	-0.0282	-0.5167
	(0.0554)	(0.4629)
Cluster 2 × Region South	-0.0359	-0.0123
	(0.0485)	(0.4071)
Cluster 3 × Region South	-0.1851**	-0.9028
	(0.0827)	(0.6909)
Cluster 4 × Region South	-0.1328**	-0.8515*
	(0.0560)	(0.4680)
Cluster 5 × Region South	-0.0908*	-0.0070
	(0.0550)	(0.4603)
Cluster 6 × Region South	-0.1576***	-0.6423
	(0.0561)	(0.4700)
Cluster 7 × Region South	-0.1614***	-0.2402
	(0.0336)	(0.2826)
Cluster 8 × Region South	-0.0354	0.6174**
	(0.0294)	(0.2484)
Cluster 10 × Region South	-0.1313***	-1.069***
	(0.0374)	(0.3147)
Cluster 1 × Region West	-0.1008*	-0.7130
	(0.0541)	(0.4525)
Cluster 2 × Region West	0.0582	-0.1357
	(0.0509)	(0.4266)
Cluster 3 × Region West	-0.1359**	-0.1714
	(0.0670)	(0.5619)
Cluster 4 × Region West	-0.0047	-1.111***
	(0.0510)	(0.4266)
Cluster 5 × Region West	-0.1469***	-0.6594

	(0.0534)	(0.4475)
Cluster 6 × Region West	-0.1828***	-0.3084
	(0.0533)	(0.4465)
Cluster 7 × Region West	-0.0828**	0.5524*
	(0.0379)	(0.3182)
Cluster 8 × Region West	-0.0988***	1.191***
	(0.0351)	(0.2957)
Cluster 10 × Region West	-0.0658	-1.119***
	(0.0407)	(0.3414)
Fit statistics		
Observations	68,145	67,323
R^2	0.11644	0.21875
Adjusted R^2	0.11590	0.21826

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table B.2: Marginal Effect: Cluster-Region interaction

Dependent Variables:	mem	verbal
Model:	(1)	(2)
Variables		
Constant	4.208***	29.69***
	(0.0355)	(0.2986)
Age	-0.0235***	-0.2310***
	(0.0004)	(0.0031)
Yrs of Edu	0.0311***	0.3926***
	(0.0009)	(0.0072)
Female	0.0047	0.0844
	(0.0070)	(0.0583)
Cluster 1 × Region North	0.0204	4.799***
	(0.0368)	(0.3075)

Cluster 2 × Region North	-0.0402	5.000***
	(0.0404)	(0.3380)
Cluster 3 × Region North	0.0125	4.447***
	(0.0378)	(0.3167)
Cluster 4 × Region North	0.0480*	4.757***
	(0.0283)	(0.2380)
Cluster 5 × Region North	-0.0396	4.731***
	(0.0353)	(0.2955)
Cluster 6 × Region North	0.0050	5.383***
	(0.0301)	(0.2516)
Cluster 7 × Region North	0.0506*	5.720***
	(0.0269)	(0.2256)
Cluster 8 × Region North	-0.2415***	3.397***
	(0.0240)	(0.2019)
Cluster 9 × Region North	-0.3871***	2.157***
	(0.0294)	(0.2488)
Cluster 10 × Region North	-0.1636***	3.585***
	(0.0289)	(0.2426)
Cluster 1 × Region South	0.1739***	1.086***
	(0.0363)	(0.3033)
Cluster 2 × Region South	0.0785**	1.326***
	(0.0329)	(0.2758)
Cluster 3 × Region South	-0.0087	0.2249
	(0.0596)	(0.4965)
Cluster 4 × Region South	-0.0463	1.489***
	(0.0379)	(0.3170)
Cluster 5 × Region South	0.1650***	1.008***
	(0.0384)	(0.3213)
Cluster 6 × Region South	0.1585***	1.152***
	(0.0405)	(0.3389)
Cluster 7 × Region South	0.1191***	1.161***
	(0.0229)	(0.1926)

Cluster 8 × Region South	0.1642***	0.2701
	(0.0213)	(0.1787)
Cluster 9 × Region South	0.0593**	-1.136***
	(0.0247)	(0.2094)
Cluster 10 × Region South	0.0725***	0.1017
	(0.0261)	(0.2190)
Cluster 1 × Region West	0.0955***	4.328***
	(0.0300)	(0.2513)
Cluster 2 × Region West	0.1669***	4.640***
	(0.0322)	(0.2696)
Cluster 3 × Region West	0.0346	4.395***
	(0.0306)	(0.2564)
Cluster 4 × Region West	0.0759***	4.668***
	(0.0250)	(0.2099)
Cluster 5 × Region West	0.1031***	3.794***
	(0.0321)	(0.2689)
Cluster 6 × Region West	0.1274***	4.924***
	(0.0324)	(0.2717)
Cluster 7 × Region West	0.1919***	5.392***
	(0.0235)	(0.1971)
Cluster 8 × Region West	0.0949***	4.282***
	(0.0232)	(0.1957)
Cluster 9 × Region West	0.0534*	2.302***
	(0.0298)	(0.2517)
Cluster 10 × Region West	0.1321***	3.490***
	(0.0255)	(0.2143)
Cluster 1 × Region East	0.1428***	2.738***
	(0.0431)	(0.3597)
Cluster 2 × Region East	0.0552	2.474***
	(0.0365)	(0.3061)
Cluster 3 × Region East	0.1171**	2.264***
	(0.0578)	(0.4847)

Cluster 4 × Region East	0.0272	3.477***
	(0.0420)	(0.3504)
Cluster 5 × Region East	0.1965***	2.151***
	(0.0401)	(0.3358)
Cluster 6 × Region East	0.2568***	2.930***
	(0.0396)	(0.3319)
Cluster 7 × Region East	0.2212***	2.537***
	(0.0259)	(0.2171)
Cluster 8 × Region East	0.1403***	0.7887***
	(0.0221)	(0.1857)
Cluster 10 × Region East	0.1445***	2.307***
	(0.0280)	(0.2347)
Fit statistics		
Observations	68,145	67,323
R^2	0.11644	0.21875
Adjusted R^2	0.11590	0.21826

IID standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*