"I can't, I have to work"

A study on the employment type of social contacts and feeling socially integrated

Bachelor thesis

Sociology

Subject: Assessing the benefits of leisure time on health, social integration and life satisfaction

By:

Renske Kleefstra

Date: 8-06-2022

S4132319

r.a.kleefstra@student.rug.nl

Supervisor: Francesca Giardini

Abstract

This study investigates the effects of the diversity of the employment types of a close contact network on the feeling of being socially integrated and investigates if there is a difference between men and women for this effect. This is interesting and important as the width of close contacts networks are decreasing, whereas their importance for support and social integration for people's well-being is big. It was expected that a more diverse network would lead to different schedules and thus different amounts of and moments of leisure time within the network. Therefore, the ego would experience less social interaction and more rejection, wherethrough the ego would feel less socially integrated. It was also expected that this effect would be weaker for women, as they already have less leisure time and already feel less well socially integrated than men. Lastly, it was expected that having more part-time contacts would be better for feeling socially integrated than having more full-time contacts. This study has been performed by using secondary data from the LISS data panel and analysed with a regression analysis in SPSS. Results have shown no evidence for any of the hypothesis. However, there has been found support for the theory that women feel less socially integrated than men.

Contents

1. Introduction

The basis of sociology is the idea of the social context shaping constraints and opportunities for the behaviour of and relations between people. The effects of environmental, social, economical, individual characteristics and dynamics between people and their social environment have therefore been widely studied. The context shapes the way in which people behave, feel and make choices.

The social network, leisure time and work are examples of the social context of people that form constraints and opportunities for social interaction and forming and maintaining relationships with people (Stalker, 2008). Forming and maintaining social relationships and feeling that you belong somewhere are crucial needs for people and studies have shown that the social ties of people affect their physical and mental health, and that social integration is beneficial for people's health (Seeman, 1996). Feeling socially integrated, which include the needs of social relationships and belonging, is an important factor in this relation and is much needed. When someone does not have many social relationships, is being excluded by people or feels like they do not belong with a certain group, they can start to feel lonely, which has consequences for their mental and physical health. A lack of attention for the need to form and maintain social relationships will thus undermine people's health (Baumeister & Leary, 1995; Cassie, Miller-Cribbs & Smith, 2020; Chang, Wray & Lin, 2014; Doolaard, Lelieveld, Noordewier, van Beest & van Dijk, 2020; Kuczynski, Halvorson, Slater & Kanter, 2021; Seeman, 1996). Studying social integration by looking at social relationships is therefore relevant for explaining people's well-being and for developing health policies. This is especially important as American studies showed that the number of close contacts of people are dropping, partly due to the increasing social media use (Hampton et al., 2011; Smith-Lovin, Brashears & McPherson, 2008). Given that the societies of America and the Netherlands are both considered as western societies, I assume that this development will also be found in the Netherlands. Thereby, forming close ties occurs less on social media, whilst people spend more time online (Neves, 2015). Meanwhile close contacts are very important for support in many ways, especially life saving support and support during hard times (Hampton et al., 2011; Smith-Lovin et al., 2008) .

Therefore, this study will investigate the close contact networks of people and how they affect the feeling of being socially integrated. An individual characteristic of these social relationships that will be investigated in this study are the employment types of the social relations. The employment types of the social relations namely affect the amount of leisure time that they have left to potentially spend on social interaction with their social relations. The amount of social interaction people have with their social relations will affect their feelings of social integration. When there is not much social interaction between them and

their social relations, they can start to feel lonely or excluded and thus socially isolated (Cassie et al., 2020; Kuczynski et al., 2021). In this way the employment types of social relations can affect the feelings of social integration of people and therefore indirect affects people's health.

This research will focus on the five closest contacts within the social networks of people. First of all, because the data provides small networks with close contacts of the respondent. But second of all, the famous sociologist Granovetter underlined the importance of strong ties, or otherwise said, close contacts. These strong ties are very important for support, protection uncertainty reduction, trust and comfort (Granovetter, 1973; Hampton et al., 2011), which are indicators for social integration.

I will study the impact of the type of employment of the five closest contacts of a person on the feeling of being socially integrated. Therefore, the research question is:

What is the effect of network diversity (of the five closest contacts mentioned in the survey) in terms of employment on social integration of Dutch people? Is there a difference between men and women?"

The employment types of the five closest contacts in the social network will be taken into account as a proxy of the time spent at work, such as full-time and part-time. I will focus on the closest contacts of Dutch people and will research if the link between the employment types of the social network and social integration feelings will be different between men and women. For this I will use data from the LISS data panel of 2020 and a regression analysis (Centerdata, 2021). I will also look into the mechanisms that lie beneath the effect of the employment of the social network and social integration. Theories of leisure time and social interaction, time-use and the homophily principle will be discussed.

2. Theoretical framework

Studying human behaviour can be done in various ways. Next to psychological or social psychological approaches that study individual behaviour, the sociological approach studies individual and group behaviour and the relation and places this behaviour in a context. A much used concept in sociology is the study of social networks. Social networks have been studied in sociology for decades (Davis, Chung & Hossain, 2005; Scott, 1988). The social network approach takes the idea of placing behaviour in a context by looking at the interaction with the individual and the context a little step further. It comes from the idea that individuals are connected and interaction between individuals will have individual outcomes. A social network can for example, have certain norms that an individual in that network will adopt. Therefore, the individual will adjust his or her behaviour to that norm. It also can be the other way around; individual attributes can have an effect on the social relationships that the individual has. This approach is thus conceptually and methodologically useful when looking at how the social context affects individuals an how individual attributes relate to people's social context (Robins, 2015). Also, Granovetter (1973) touched on the importance of studying social networks, as a social network has influence on individual feelings, behaviour and experiences, outside of the individual's control. Given the social and relational nature of the main concept of this study (social integration), looking at the social network is very useful.

Studying social networks can be done on different levels. Firstly, there can be looked at the whole network. This can be a network of a classroom, family, friends, colleagues, at almost every possible connection that a person can have. Secondly, the dyadic relation, the relationship between two people, can be studied. Thirdly, the dynamics within a network can be studied, but another way is looking at egocentric networks. This kind of study focuses on ego-centred networks. These are networks of a given individual called the ego. The people with whom the ego has relationships with and form the network with are called ties or contacts. Ego networks study the network from the perspective of the ego (Davis et al., 2005). In this study I will look at the influence of the network on a given individual of that network. I will focus on the feelings of the ego, in contrary to other methods where also the feelings of other actors within a network are studied. Looking at an egocentric network is the best way to study this.

So, in this study there will be looked at the individual judgement of the social network and social life, to determine their social integration, and the individual characteristics of the people the individual is closely connected to. These characteristics are good to analyse, here the employment types of the ties, as it can explain the social structure and the assessments of quality of the relationships. Looking at the social networks of people is a way to assess the

social integration of people, as it shows how connected a person is with other people. For studying social integration, social networks are quite crucial to look at as social integration regards the embeddedness in a social structure and social networks give information about this embeddedness and people's experience of being connected (Granovetter, 1973; Robins, 2015).

Another characteristic that will be looked at in this study is on the group level. There will be looked at network diversity. Network diversity is the degree in which the network of an individual differs on certain aspects. It is how different or similar people are from each other within a network. Research on network diversity has mainly been done in terms of difference in social characteristics such as race, gender and religion (Mcpherson, Smith-Lovin & Cook, 2001; Son & Lin, 2012). This research will focus on network diversity in terms of employment. This means, the difference in the types of employment of the five closest contacts of the network. The type of employment in this case is whether a contact works full-time, part-time, or not at all.

Social integration

Networks are thus a good way to look at the degree of social integration of people. But to understand this better, a good definition of this concept is needed. Social integration is a wide concept. It has been known as "the extent to which they [people] have social ties or social connections" (Seeman, 1996), or as the converse of social isolation. It is commonly measured by the amount of social support and the amount social ties people have, or the size of someone's social network (Seeman, 1996). Another way of looking into the social integration of people in social networks is looking at transitivity: the likelihood that a friend of your friends are friends (Louch, 2000).

However, in this research I will define social integration as the degree of which a person feels socially isolated or, otherwise put, the absence or presence of feelings associated with social isolation. This definition can be called subjective social integration or subjective social isolation. Social isolation is a situation wherein a person lacks interaction and meaningful relationships with other people (Cassie et al., 2020; Teater, Chonody & Davis, 2021). Social isolation brings up negative feelings such as, loneliness, anxiety, stress and less life satisfaction. Reversely, more social interaction, especially with close contacts, or a good social network delivers a smaller risk for developing mental health issues, feelings of loneliness and is related to increased well-being. Loneliness is the most important feeling that is often associated with social isolation, even though these two do not always go hand in hand (Cassie et al., 2020; Kuczynski et al., 2021; Teater et al., 2021; Tonković, Cepić & Puzek, 2021). This, once again, shows the importance of social networks for social integration. Since this research specifically focuses on the five closest contacts of a person, as the questions

asked to the respondents for five of their closest contacts. As a close contact indicates that the contact has a good relationship with the ego, I presume that the ego enjoys interaction with their close contacts. The quality of the interactions when they occur are therefore assumed perceived as positive by the ego.

Based on the above, I state that people who have more social interaction with their close contacts feel better socially integrated and thus less socially isolated. Therefore, they feel less lonely and are more satisfied with their social contacts.

Homophily

A well-known principle within theories of network diversity is the concept of homophily. The homophily principle is well know for the expression: 'birds of a feather flock together'. Which is a metaphor for the idea that similar people tend to form groups with each other. Homophily states that people who have similar attributes are likely to have more contact with each other, compared to people who have less similar attributes. Additionally, people tend to conform to their network as the heterophilic effect makes people who have ties that are different from them change to correspond more with their ties and therefore get similar attributes. This will result in more social interaction, as the homophily principle states that similar people have more social interaction with each other (Lozares, Verd, Cruz & Barranco, 2014; Mcpherson et al., 2001).

In this research I will start from the homophily principle. That is, in this research, egos with contacts that have similar employment types are more likely to have the same employment type as their ties. Since the similarity of employment of the network indicates that the heterophilic effect has occurred. This means that the ego has adapted to their network and thus is likely to adapt to the attributes of others. Also, when a network is similar there is a more obvious dominant attribute (employment type) to correspond to than when a network is more diverse. When the network of the ego is more diverse, it is more difficult to assess to which attribute the ego must correspond to as there is not a dominant employment type. Therefore, when the employment types of the network are more diverse, there will be less ties that the ego has similar employment types with. Applying this, when all of your friends work full-time, you are likely to go work full-time as well. However, it could be that you already worked full-time, just as your friends, before you started the relationships. Otherwise said, it could be that these attributes are selected for the relationships. But an influence process (heterophilic effect) could also occur. When you work part-time and your friends full-time, you are very likely to start working full-time as well. So, homophily can work as a selection process, similar people get together, but also in a influential way, adapting

to the network. Which way the homophily works is hard to determine and in this study no statements about this can be made.

Employment diversity of the network and feeling socially integrated

Now that the homophily principle and the reasoning of the idea that social interaction with close contacts will lead to feeling better socially integrated is clear, the question remains how the network diversity affects the social integration of the ego. Whereas large and diverse networks are mostly associated with less loneliness (Kuczynski et al., 2021), I hypothesize the opposite for employment diversity. I stated earlier, based on the homophily effect, that when an ego has a diverse network in terms of employment types, the ego is more likely to be less similar to their network. Therefore, I assume that when a network is diverse in terms of employment, the time-use of the ties are different from the ego. As the one works part-time, therefore spends less time on working and has more leisure time, the other works full-time, has less leisure time and spends more time on working. The differences in time-use will cause less social interaction between the ego and the network and this will lead to negative feelings of social integration. Studies namely showed that leisure time and social interaction with the closest social contacts has the biggest influence on mental health and feelings of loneliness (Chang et al., 2014; Kuczynski et al., 2021; Tonković et al., 2021). This means that a network who has more leisure time will lead to more social interaction between the ego and its ties. This will enhance the feeling of being socially integrated. As part-time employees have less time spend on work and thus more leisure time than full-time employees, therefore the first hypothesis will be:

H1: More part-time working contacts, will lead to feeling better socially integrated, compared to more full-time working contacts.

Additionally, during the process of planning social interaction with the ties, egos who have diverse networks in terms of employment will face more rejection because the leisure time is very different. For instance, when the ego has leisure time to spend with a close tie as it works part-time, the close tie may not have the time as it works full-time. When the ego calls the tie to ask if they want to meet up, the tie rejects the offer. "I'm sorry I can't, I have to work" is what the ego will hear a lot. This will make the ego feel less connected to the tie as I will elaborate.

People have a very strong need and desire to belong, be connected and to form and maintain social relationships. For this, a feeling of relatedness is not enough, frequent interaction with the same strong relationships is very important for meeting these needs. Neglections of these needs cause people to feel socially excluded and causes them to feel

negative feelings such as unhappiness and loneliness. Even the slightest sign of rejection or social exclusion can bring up these feelings. Frequent rejection of social interaction will therefore make people feel less satisfied with their social life (Baumeister & Leary, 1995; Buckley, Winkel & Leary, 2004; Doolaard et al., 2020). When an attempt to engage in social interaction is rejected by a tie, there will be negative feelings experienced by the ego. Because the ties of the respondents are close contacts, frequent interaction is important in giving the respondents a sense of belonging and prevent them to feel unhappy. A diverse network in regard to employment type, with whom social interaction is more difficult to realise and arrange, will have less social interaction, experience more rejection for social interaction, which will result in people being less happy about their social life and feel less well socially integrated. Therefore, the second hypothesis states as follows:

H2: As a person's network is more divers in terms of employment types, the ego will feel less socially integrated.

Gender

A less diverse network in terms of employment will thus be better for social integration then a very diverse network. I expect that the feeling of being socially integrated will be less influenced by the diversity of employment of the network for women, compared to men. Women spend more of their leisure time on housework and caring activities than men. This means that women have less actual leisure time than men as housework and caring activities are still forms of labour. In the literature there has been said women have less 'pure' leisure time, or women have a 'second shift'. Not only the quantity of the leisure time is less, also the quality of leisure time is lower because of the multitasking women need to do and the presence emotions during care (Firestone & Shelton, 1988; Stalker, 2008; Tomaskovic-Devey, 1993; Warren, 2010; Yerkes, Roeters & Baxter, 2020).

Women spend more of their time on housework because they supposed to do that according to the social role and norms that stick to women in society. Women are supposed to clean and care for the children (Firestone & Shelton, 1988; Yerkes et al., 2020). This is a very traditional perspective on the role of women and the norms differ between social cultural contexts between countries. In western Europe there seems to be a trend towards more gender equality. Among which The Netherlands is one of the best gender equal societies. The equality index in the Netherlands has grown since 2010. Therefore, the Netherlands is the fifth most gender equal country of Europe (European Institute for Gender Equality, 2021). In more equal countries, the difference in quality and quantity of leisure between the two genders is smaller because the traditional norms are less strong and men do more housework and care than in less equal countries (Yerkes et al., 2020). This points to an equal effect of network diversity in terms of employment on feeling socially integrated for

men and women. However, Dutch women are working more. In spite of this development, women still spend more time on householding and caring activities than men and spend less time on social activities than men (*Gender Equality Index | 2021 |*, 2021; European Institute for Gender Equality 2021). This indicates that women work more but still do more other activities, which are not leisure time, than men. Thus, the main argument why women will feel less socially integrated regardless of their network is that they have less leisure time then men.

So, it does not matter if a woman works full- or part-time, she will still have less leisure time left to spend with her close contacts than a man would have with the same employment type. The housework takes up a part of the time that she also could have spend on social interaction. This way, women will feel less socially integrated, regardless of their network, than men. Therefore, my second hypothesis is as follows:

H3: The diversity of employment types of the network has less effect on the feeling of being socially integrated for women then for men.

Age, income and domestic situation

It has been shown that the social status of a person has influence on the degree of social interaction (Berger, Cohen & Zelditch, 1972). Therefore, in this research I will incorporate three other factors, besides work, that also have an influence on how socially integrated a person feels.

First the age of a person is a social status that gives a representation of the amount of time spend on work, how likely it is that they have children, and how healthy they are. This will also influence they degree of homophily and the amount of time that can be spent on social interaction. It has namely been shown that the degree of homophily within a network differs between age groups (Berger et al., 1972; Lozares et al., 2014). Also, a study about loneliness in Europe showed that age was an important factor with comparing countries in loneliness. In Nordic countries, a younger age was related to more loneliness compared to other countries. This while loneliness is a well-known problem for elderly people (Tonković et al., 2021). It has thus been shown that age has different influences on concepts that are important in this study. Therefore, there will be controlled for age in the analysis.

Secondly income is a social status that also represents the amount of time spend on work and the possibility to spend money to participate in social activities. To go drink something witch someone for example. It has been proven that people with low incomes are more socially isolated and experience a lower sense of belonging than people with high income (Stewart et al., 2009). People with low income withdraw themselves from social interaction, which can result in social isolation. Moreover, poverty can provoke anxiety about

their social contacts (Eckhard, 2018). Income thus has an effect on loneliness and is therefore incorporated as control variable.

Lastly, the domestic situation can make a person spend more time at home with family or a spouse, which will cause less social interaction with other people that may be important to them. Especially women who live with children and or are married have less leisure time to spend on their network, as they have more domestic labour to do. However, people who live alone do feel lonelier. Often people with whom a person live together are close and important ties. When living alone, such a close tie is missed, as social interaction with close ties on a daily basis is negatively correlated with loneliness (Firestone & Shelton, 1988; Tonković et al., 2021). The domestic situation thus has an effect on leisure time, but also on loneliness and is therefore incorporated as control variable.

The research model is visualized in figure 1, based the research question and the theory.
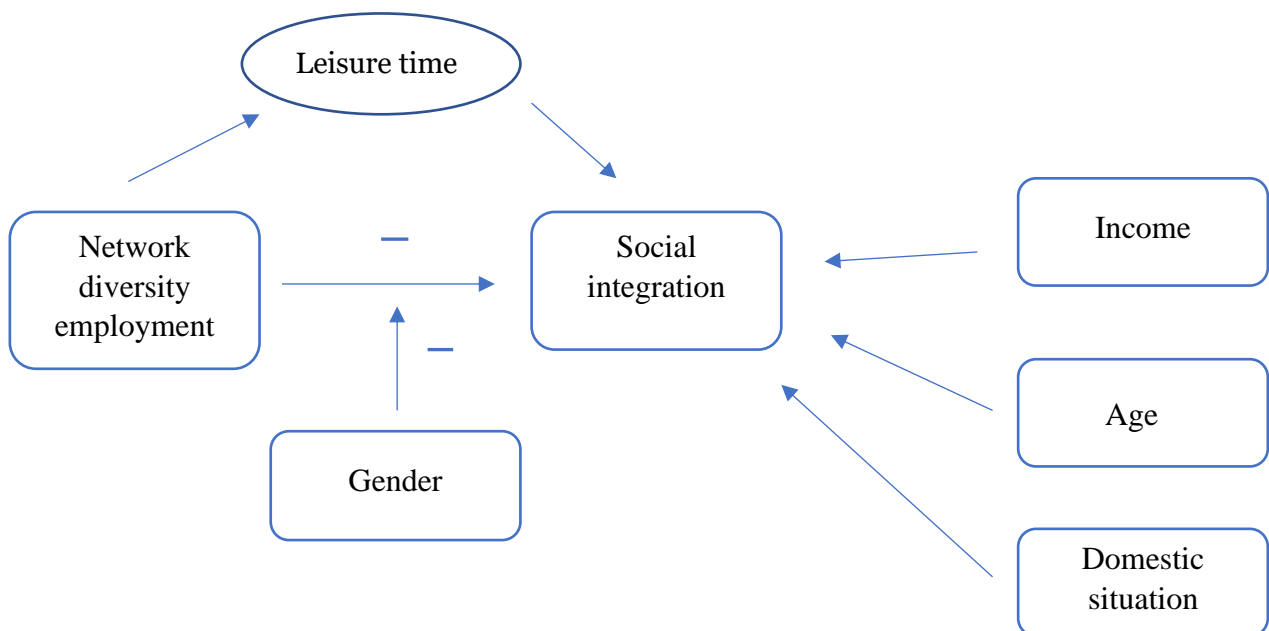

Figure 1 Research model with hypotheses with in the circle theoretical mechanisms and in the rectangles the variables that will be statistically tested

3. Methods

For this study I used secondary quantitative data from the LISS data panel. This is a panel that consists of five thousand households and around the seven thousand five hundred people who complete online questionnaires once a month in exchange for a payment. When a respondent does not have equipment to complete an online questionnaire, it is borrowed from the panel. The people in the sample are first approached with a letter. When there is no response, people are reached by telephone and lastly visited at their houses (*About the Panel | LISS Panel Data*, n.d.). I used wave thirteen from the *Core study on Social integration and leisure*. This study is longitudinal and the data was collected in October and November 2020 among Dutch panel members of 16 years and older. 6680 people have been selected for this specific study. Eventually 5970 responses were collected, which means that there were 710 non-responses. From the responses 87 were incomplete and 5883 complete (Centerdata, 2021). A dataset which only contained the variables that were needed for this study was made. Other variables were deleted. Respondents with missing data were deleted, therefore the final sample size is N=4544. The data was saved and analysed in SPSS (26th edition). For an extensive elaboration of the dataset and the items and operationalisations see appendix 1.

3.1   Variables and operationalizations

3.1.1   *Social integration*

I will discuss which questions of the LISS data I have used and how I constructed them into variables. Firstly, I used seven questions to construct the dependent variable social integration. These seven items contain two questions. The first question is: "How satisfied are you with your social contacts?" respondents could rate their satisfaction with a number between 0-10, where a high number indicates better satisfaction. The other six question were: "I have a sense of emptiness around me ", "There are enough people I can count on in case of a misfortune", "I know a lot of people that I can fully rely on", "There are enough people to whom I feel closely connected", "I miss having people around me" and "I often feel deserted". These six questions have the answer categories 1=Yes, 2=More or less and 3=No. In table 1 an overview of the items and their questions and operationalizations before the scale construct can be seen. For the item cs20m283, I only recoded the answer category -9 into missing. I mirrored cs20m285, cs20m268 and cs20m287, so that for every item a high score represents a positive feeling. Then, for constructing the eventual variable, I made a sum of all the seven items which I called Social integration. This variable runs from 6-28, where a higher score represents a better social integration feeling.

Table 1 Items and operationalisations of the items for the variable Social Integration

Variable	Question	Answer categories	Scale	Operationalisations
cs20m283	How satisfied are you with your social contacts?	0-10 0= not at all satisfied 10= completely satisfied -9 I don't know	Ordinal	-9 → missing
cs20m284	I have a sense of emptiness around me	1=Yes 2= More or less 3= No	Ordinal	
cs20m285	There are enough people I can count on in case of a misfortune	1=Yes 2= More or less 3= No	Ordinal	Mirrored 1=No 2= More or less 3= Yes
cs20m286	I know a lot of people that I can fully rely on	1=Yes 2= More or less 3= No	Ordinal	Mirrored 1=No 2= More or less 3= Yes
cs20m287	There are enough people to whom I feel closely connected	1=Yes 2= More or less 3= No	Ordinal	Mirrored 1=No 2= More or less 3= Yes
cs20m288	I miss having people around me	1=Yes 2= More or less 3= No	Ordinal	
cs20m289	I often feel deserted	1=Yes 2= More or less 3= No		

3.1.2 *Network diversity in terms of employment*

For the construction of the variable that represents the diversity of the close contacts in terms of employment I used the question: "*Does [person 1/2/3/4/5] work full-time, part-time, or not at all?*". With the answer categories: 1=full-time, 2=part-time and 3=not at all and -9=I don't know. There is an item for every close contact that the respondent has assigned for this question. These are the items: cs20m327 (person 1), cs20m338 (person 2), cs20m349 (person 3), cs20m360 (person 4), cs20m371 (person 5). Only answers where the respondents have assigned five persons are considered in this question.

I recoded the answer category -9 into missing for all items, as this is answer is not useful. Then I mirrored the answer categories 1-3 for all items, so that a higher score represents more working. Next, I made two dummies for every item. One dummy for every items where 1=full-time and 0=part-time and not at all (fullt1, fullt2, fullt3, fullt4 and fullt5). The second dummy for each item is coded into 1=part-time and 0=full-time and not at all (partt1, partt2, partt3, partt4 and partt5). Then, I made a mean score variables of the full-time dummies, called propfull, and the part-time dummies, called proppart. This way, proportions of the employment types of the contacts are represented. The rest proportion is the proportion of contacts that do not work. I also made centred versions of propfull and proppart called Cpropfull and Cproppart, so that they can be used for the moderation analysis.

3.1.3 *Gender*

For the gender variable I used the variable called geslacht from the dataset. This variable represents the gender of the respondent. For this variable 1=woman and 2=man. I recoded this dummy into another dummy called Gender, where 1=woman and 0=man. Then I made interaction variables between Gender and the centred network diversity variables Cpropfull and Cproppart. These interaction variables are called GenXCpropfull and GenXCproppart.

3.1.4 *Control variables*

For the control variable age, I used the variable Age from the dataset. This variable represents the age of the respondent. No operationalisations have been done on this variable. For income I used the variable nettoink_f. This variable represents the monthly net income

of the respondent. No further operationalisations took place for this variable. Lastly, for the domestic situation I used the variable called woonvorm. This variable has five answer categories: 1=single, 2=(un)married co-habitation, without child(ren), 3=(un)married co-habitation with child(ren), 4=single, with child(ren) and 5=other. Firstly, I recoded the answer category five into missing, as it is not a useful answer. Next, I made three dummies. The first dummy represents that the respondent lives in co-habitation without children, called Cowithout. Here 2= (un)married co-habitation, without child(ren) is recoded in to a 1 and the rest of the categories are coded into a 0. The second dummy represents co-habitation with children called Cowith. Now 1=(un)married co-habitation with child(ren) and the rest of the categories score a 0. The third dummy represents single habitation with children called Singlewith. Here 1=single, with children. When a respondent scores a zero on all dummies, then the respondent lives alone without children.

3.2   Analysis plan

The hypotheses will be tested by a regression analysis. This will be done in a few steps. First the control variables (age, income and domestic situation) will, together with the dependent variable social integration, form the first model. Here the effects of the control variables on social integration will be researched. Next, the network diversity dummies will be added to the first model. This will form model 2 and will test what influence the network diversity in terms of employment has on social integration, controlled for age, income and domestic situation. This model tests the first and second hypothesis. Next, the gender variable will be added to the second model, which will form model 3. Here the influence of gender on social integration will be studies, as well if the adding of gender to the model changes something on the effect of network diversity. Lastly, the third hypothesis will be tested in the last model, model 4. In model four the interaction variables will be added to really test the moderating effect of gender on the effect between network diversity in terms of employment and social integration.

4. Results

4.1   Descriptive and bivariate statistics

4.1.1   Descriptive statistics

In table 1 the descriptive statistics of all the variables that will be included in the regression analysis are shown. Here can be seen that on average people feel quite well socially integrated, as the mean is 19,3. Most of the close contacts of the respondent do not work (38,48%) next 33,68% work full time and 27,8% work part time. In table 1 there can also be seen that the majority of the respondents are female and that the average age of the respondents is 51,16 years old. The average income of the respondents is €1809,75 per month. However, the standard deviation is very big (s=2526,15). This indicates that the distribution of the incomes of the respondents is very skewed. Some people earn much more than the average and some people earn much less. Mostly, the respondents live in co-habitation without children (37,6%). The second most domestic situation of respondents is single with children (31,3%). Furthermore 25% of the respondents lives alone without children and 6,2% lives in co-habitation with children.

Table 2 Descriptive statistics of all the variables included in the analysis

Variable	Mean or proportion (standard deviation)[a]	Minimum	Maximum
Social integration	19,30 (2,37)	8	28
Network diversity of employment	33,68% Full-time	0	1
	27,84% Part-time		
	38,48% Does not work		
Gender	43,20% Male		
	56,80% Female	0	1
Age	51,15 (18,44)	16	103
Income	1809,70 (2526,15)	0	146.652
Domestic situation	25,0% Single, without children	0	1
	37,60% Co-habitation, without children		
	31,20% Co-habitation, with children		
	6,20% Single, with children		

N=4541

4.1.2   Bivariate statistics

In table 2, the coherency measurements of the variables can be seen.

Table 2 Coherency measurements of the variables incorporated in the analysis

	1.	2.	3.	4.	5.	6.	7.
1. Proportion part-time	-	**-0,452	-0,005[a]	*-0,036[a]	**-0,201[a]	**0,106[d]	**0,033[c]
2. Proportion full-time		-	*0,038[a]	-0,009[a]	**0,322[a]	0,010[d]	*0,038[c]
3. Social integration			-	0,027[a]	**0,125[a]	**-0,053[d]	**0,013[c]
4. Income				-	**0,103[a]	**-0,167[d]	**0,005[c]
5. Age					-	**-0,085[d]	**0,179[c]
6. Gender							**0,084[b]
7. Domestic situation							

**significant with $p<0,01$ *significant with $p<0,05$; a Pearson correlation, b Cramer's V; c correlation with ANOVA F-test; d correlation with paired t-test

In table 2 can be seen that the correlation between the proportion part-time working contacts of the respondents and social integration is very weak, negative and not significant (*r*=-0,005). There seems to be no relation to the proportion part-time working contacts and social integration. In increase above the average proportion part-time working contacts is not associated with a higher or lower feeling of social integration. The relationship between the proportion full time working contacts and social integration is however bigger and significant with an alpha of 0,05 (*r*=0,038). In contrary to the relationship between the part-time proportion and social integration, the relationship between the full-time proportion and social integration is positive, but still weak. Though, it can be stated that a higher proportion of full-time working contacts is related to feeling better socially integrated, but this relationship is not very strong.

The relationship between gender and social integration is significant, but weak (*Cramer's V*=-0,053). Being female is related to feeling less socially integrated, however the relationship is not very strong. There is a big difference between the relationships between gender and the employment type proportions. The relationships between the proportion part-time working contacts and gender are much stronger than the relationship between the proportion full-time working contacts and gender. The relationship between gender and the proportion part-time working contacts is positive, significant and moderately strong (*Cramer's V*=0,106). Being female is related to a higher proportion of part-time working contacts.

The relationships with regards to the hypotheses are not very strong. The same goes for the relationships between the control variables and social integration. However, the relationship between age and social integration is relatively strong (*r*=0,125). This relationship is positive and significant. An older age is related to feeling better socially integrated.

The correlations that stand out the most are the correlations between the employment type proportions and age. Both are moderately strong. For the proportion part-time working contacts a negative relation applies (r=-0,201). A higher proportion part-time working contacts is associated with an older age. This can have something to do with the probability that the contacts are around the same age as the respondents. When the contacts are older pensions and health play a role, wherethrough contacts will work less, and the proportion part-time working contacts is therefore higher for older respondents. This means more leisure time to potentially spend on social interaction.

4.1.3   *Assumptions of linear regression*

A few sidenotes on the tested models most be made. The assumptions of linear regression have been tested. The assumption of independency has not been violated, but as the data is secondary the assumption of independency can not be made with a hundred precent certainty. The assumption of linearity and homoscedasticity are not violated. The assumption of normality is violated. The dependent variable is left skewed distributed. After exploring possible interventions, the conclusion has been taken that the variable social integrated is probably skew distributed in the population by itself. No other variables could explain the skew distribution. This could affect the results negatively. Therefore, the tests will be done more strictly. An alpha of *p*=0,01 will be used.

4.1.4   *Multicollinearity and outliers*

No problematic multicollinearity has been detected. There have been detected three influential points or outliers. These have been excluded from the analysis. Two of the three

were extremes regarding income. The analysis has not changed very much after deleting these outliers. Therefore, only the last analysis is presented in this study and in appendix 2. More details about the outliers and assumptions can be found in appendix 3.

4.2   Hypothesis testing

For the hypothesis testing a linear regression has been done according to the analysis plan. The results are displayed in table 3. An alpha of p=0,01 will be used to determine significance.

In table 3 model 1, the effects of the control variables on social integration have been tested. What first stands out is the really small effect of income (b=1,813E-5; p=0,190). This effect is not significant but indicates when earning €1000 more per month, people will feel 0,0813 more socially integrated, which is still a small effect. The effect of age is, positive and significant (b=0,016; p<0,001), however still small. When being 10 years older, you will feel 0,16 more socially integrated. This means that the older people get, the more socially integrated they will feel. However, this is a small effect and slow process, as it takes 10 years to feel a little bit more socially integrated. On the scale of social integration these effects are neglectable. The effects of the domestic situations are however bigger. Being single with children will lead to feeling b=0,488 more socially integrated, compared to living alone. This effect is significant (p<0,001) and means almost half a point increase of social integration. The effect of living in co-habitation without children is comparable (b=0,591; p<0,001). Living in co-habitation without children will lead to 0,0591 on the 6-28 scale of social integration, compared to living alone. This far, not living alone seems to lead to feeling better socially integrated. This also applies to the last category of domestic situation, however this effect is much smaller and not significant (b=0,256; p=0,190). The effect implies that living in co-habitation with children will cause a 0,256 increase in feeling socially integrated, compared to living alone. However, this effect is not significant.

Table 3 Parameters od the regression analysis with as dependent variable social integration

	Model 1		Model 2		Model 3		Model 4		VIF
	b (SE)	p	b (SE)	p	b (SE)	p	b (SE)	p	
1. Social Integration (constant)	18,065 (0,134)	<0,001	18,066 (0,138)	<0,001	18,216 (0,149)	<0,001	18,220 (0,149)	<0,00 1	
2. Age	0,016 (0,002)	<0,001	0,016 (0,002)	<0,001	0,015 (0,002)	<0,001	0,015 (0,002)	<0,00 1	1,339
3. Income	1,813E -5 (0,000)	0,190	1,898E-5 (0,000)	0,171	1,344E-5 (0,000)	0,338	1,294E-5 (0,000)	0,357	1,721
4. Single with children	0,488 (0,097)	`<0,001	0,480 (0,098)	<0,001	0,467 (0,098)	<0,001	0,465 (0,098)	<0,00 1	1,206
5. Co-habitation with children	0,256 (0,158)	0,190	0,254 (0,159)	0,122	0,251 (0,159)	0,114	0,254 (0,159)	0,111	1,589
6. Co-habitation without children	0,591 (0,090)	<0,001	0,591 (0,090)	<0,001	0,577 (0,090)	<0,001	0,575 (0,090)	<0,00 1	1,045
7. Centred proportion part-time			0,172 (0,120)	0,183	0,210 (0,130)	0,107	0,204 (0,190)	0,282	2,758
8. Centred proportion full-time			0,097 (0,120)	0,417	0,120 (0,120)	0,320	0,218 (0,167)	0,192	2,648
9. Gender					-0,180 (0,072)	0,012	-0,189 (0,073)	0,009	1,083
10. Gender X Centred proportion part-time							-0,006 (0,260)	0,192	2,876
11. Gen X Centred proportion full-time							-0,193 (0,232)	0,405	2,755
R2 adjusted	0,025		0,025		0,026		0,026		
F Change (p)	24,272 (<0,001)		0,928 (0,395)		6,257 (0,012)		0,424 (0,654)		

N=4541

In model 2 of table 3, the main effects, centred proportion part-time working contacts and the centred proportion full-time working contacts, are added. Here the first and second hypotheses will be tested: "More part-time working contacts, will lead to feeling better socially integrated, compared to more full-time working contacts." And: "As a person's network is more diverse in terms of employment types, the ego will feel less well socially integrated."

Both effects of the proportions part-time and full-time working contacts are positive, but not significant. The effect of the proportion part-time working contacts is the biggest of the two (b=0,172; p=0,183). As the close contact network contains of five people and the effect is based on 1% increase, an increase of 1 person will mean a 20% increase in proportion. Therefore, the effect implies that a 20% increase in the proportion part-time working contacts above the average will lead to a 3,44 (0,172*20) point increase of social integration, compared to the proportion not working contacts and controlled for age, income and domestic situation. On the 6-28 scale of social integration, this is quite a big effect. However, the effect is not significant. The effect of the proportion full-time working contacts on social integration is also positive (b=0,097; p=0,320). The effect implies that a 20% increase proportion full-time working contacts above the average, leads to an increase of 1,94 (0,097*20) points on the scale of social integration, compared to the proportion not working contacts and controlled for age, income and domestic situation. This too, is quite a big effect but not significant. The effects are in line with the first hypothesis, as more part-time contacts leads to a bigger increase in feeling socially integrated than the proportion full-time contacts. However, the effects are not significant. The effects are not in line with the second, as an increase in the proportion of both employment types will lead to an increase in feeling socially integrated.

In model 3 of table 3, the effect of gender has been added. The effect of gender on social integration is negative and on the edge of being significant with an alpha of p=0,01 and is significant with an alpha of p=0,05 (b=-0,180; p=0,012). Scoring higher on gender, being a female, leads to a decrease of 0,180 points on the social integration scale, compared to being male and controlled for age, income and domestic situation. Women feel less socially integrated than men. This is in line with the theory. However, this effect is not significant. The effects of the proportion part-time working contacts and the proportion full-time working contacts have increased (Proportion part-time: b=0,210; p=0,107; Proportion full-time: b=0,120; p=0,320). Adding gender to the model had a strengthening effect on the employment type proportion variables.

Lastly the complete model is visible in table 3 as model 4. Here the interaction variables between gender and the employment type proportion variables have been added. Here, there will be tested if there is a difference between men and women for the effect of the

employment type proportions on social integration. Specifically, the hypothesis: "The diversity of employment types of the network has less effect on the feeling of being socially integrated for women then for men", will be tested. The interaction variables are both negative and not significant. The negative interaction effect between gender and the proportion part-time working contacts (GenXCproppart: b=-0,006; p=0,192) indicates that the effect of the proportion part-time working contacts on social integration will decrease with 0,006 given that the respondent is a woman and that the proportion part-time working contacts is average (Cproppart=0). On the 6-28 scale of social integration and looking at the effect size of the variable Cproppart this is a very small effect and is not significant. The interaction effect is stronger between gender and the proportion full-time working contacts (GenXCpropfull: b=-0,193; p= 0,405). The effect of the proportion full-time working contacts decreases with 0,193 given that the respondent is a woman and that the proportion full-time working contacts is average (Cpropfull=0). Compared to men, the effect of the proportion full-time working contacts on social integration is much weaker for women. Given the effect size Cpropfull, this is a fairly big effect. The effect will would more than double its size. The effects are in line with the theory. However, this effect is not significant. The effect of gender has increased and is now significant (b=-0,189); p=0,009). Females score 0,189 lower on social integration than men. This is in line with the theory. The effect sizes of the proportion variables have become bigger and are now more or less equally strong (Cproppart: b=0,204; p=0,282; Cpropfull; b=0,218; p=0,192). This is not in line with the first and second hypotheses.

4.2.2 Model evaluation

In table 3 can be seen that the first model explains 2,5% of the variance within social integration ($r^2adjusted$=0,025). This is not much, therefore the first model is not a very good model for predicting social integration. Then, the second model has the same amount of explained variance as the first model. Model 2 also explains 2,5% of the variance within social integration ($r^2adjusted$=0,025). Model 2 does not add any explained variance. This can also be seen in the F change which is very small and not significant (*Fchange*=0,928; p=0,395). Model 2 is thus not a very good model for predicting social integration.
In model 3 the proportion explained variance has gone up compared to the second model, but not significantly (*$r^2adjusted$*=0,026; *Fchange*=6,257; p=0,012). Still, the explained variance is very small. Model 3 also does not a very good job in predicting social integration, but there seems to be a small improvement in comparison to model 2, even though it is not significant. Lastly, compared to model 3, model 4 did not add any explained variance as it stayed the same. There was also no significant F change (*$r^2adjusted$*=0,026; *Fchange*=0,424; p=0,654). Therefore, model 3 can be seen as the best model for predicting social integration.

5. Conclusion & Discussion

The number of close contacts in networks are decreasing, which has negative consequences for people's social integration. Meanwhile, feeling socially integrated is important for people's physical and mental health. To feel socially integrated, interaction with and support from close contacts are important. To make this possible, leisure time is needed to spend on social interaction with each other. Therefore, in this study the effect of the employment types of close contact on feeling socially integrated has been studied with the question:"*What is the effect of network diversity (of the five closest contacts mentioned in the survey) in terms of employment on social integration of Dutch people? Is there a difference between men and women?*" Not much research has been done on the relationship between the diversity of employment types of close contacts on social integration. Therefore, the results of this study contribute to the state of art on this topic and only gives more reason for future research as some results of the analysis were interesting but not significant.

Results of a regression analysis have shown that the proportions part-time and full-time working contacts did not have an effect on feeling socially integrated. No support for the hypothesis that a more diverse network in terms of employment will lead to feeling less socially integrated has been found. This is not in line with the theory as it was expected that a more diverse network means that the close contacts have different work schedules and leisure time, wherethrough it is more difficult to make plans with each other. This way people will experience more rejection and less social interaction. However, the effects were in the direction of the hypothesis that a network with more part-time working contacts is better for feeling socially integrated than a network with more full-time contacts. This is in line with the theory that when the network has more leisure time, the ego will feel socially integrated because there is more time for social interaction. Further research is recommended to confirm this theory.

The results did find support for the theory, that women averagely feel less socially integrated then men, because women have less leisure time than men. This supports the theory that women have less time to invest in social interaction and other activities that enhance the feeling of being socially integrated. It is important to investigate this further and direct more attention to the social well-being of women in policies. Policy could be directed to unburden women from their 'second shift', with free day-care or encouraging men to pick op housework with campaigns or longer parental leave for men.

It was therefore also expected from the theory that, regardless of the employment type of the woman and her close contacts, she would still feel less social integrated than a man would. Therefore, the effect of network diversity on feeling socially integrated was expected to be weaker for women than for men. However, there was no evidence found for the third

hypothesis that there is a difference between men and women with regards to the effect of the diversity in employment types of the network on feeling socially integrated. The interaction effects where however interesting, as the interaction between gender and the proportion full-time working contacts was much bigger than the interaction with the proportion full-time working contacts. This would mean that women with a high proportion of full-time working contacts feel much less socially integrated than women with a high proportion of part-time working contacts. The theory about leisure time and the opportunity of social interaction would fit here. Perhaps the amount of social interaction is more important than the quality for women, wherethrough a network with not a lot of leisure time left to spend on social interaction would have a greater negative impact on women than on men. This effect is interesting for further research.

This study did not find support for an effect of network diversity in terms of employment on social integration. However, a previous study of Olson, Pizzagalli and Rosso (2021) has found that a less diverse network in terms of social roles (employment is also a part of that role) is associated with social anhedonia. Which is a state of social withdrawal and depression (Olson et al., 2021). So, there is some indication for a relationship between network diversity and feeling socially integrated. Further research would add more explanation on this.

The results of this study are also in line with the findings of Seeman (1996). He found that the positive effects of support from social ties had greater benefits for men than for women. The theory behind this was comparable to the theory in this study wat the reasoning that women are more likely to "bear the costs of caring" than men. However, another reasoning that was different from the theory behind this study is that Seeman (1996) reasoned that women are more prone to negative social interactions with ties than men. Therefore, women will experience negative emotions about their social ties quicker than men would, which leads to more negative health outcomes.

A few limitations of this study must be addressed. First, in the theory and in the data, it was not considered if the close contacts were colleagues of each other. Social interaction on the work floor is therefore not considered. This could mean that people who work full-time would compensate for their lack of social interaction during leisure time, with social interaction on work. This potential mechanism would cause smaller effects and is therefore interesting for future research.

Secondly, this study used secondary data. This brings limitations for measuring concepts in the way they were meant to for this study, because the data was not collected with this study in mind. This erodes the validity and reliability of this study. The concepts would have been measured differently and more relevant information could have been collected. For this study it would have been great to have information about the employment type of the

ego also, to make better assessments. On top of that, the main concept, social integration, is a hard to measure concept. Now, social integration has been measured by asking the respondents how satisfied they are with their social contacts and how often they feel certain negative emotions. This way, the concept has been measured merely in a subjective way. I would recommend defining and measure social integration more objectively, by looking at the amount of contact and number of close contacts, perhaps in addition to the more subjective definition.

Another potential limitation of this study is the small network size that has been chosen to measure. This way, there is less of a chance to find an effect because a small amount of people need to cause this effect. Additionally, the number of close contacts that were mentioned gradually decreased with the person. This meant that a lot of people mentioned one person, but less people mentioned a second contact and so on. Only the people who mentioned five close contacts were incorporated in this analysis. Therefore, quite some data has been lost. A possible explanation for this could be a method bias. A study of Galesic and Bosnjak (2009) has shown that questions at the end of the questionnaire had higher non-response rates, were answered shorter, faster and of lower quality, compared to when the same questions were asked earlier in the questionnaire. This is a consequence of a long questionnaire, wherethrough the burden for the respondent increases towards the end of the questionnaire (Galesic & Bosnjak, 2009). These questions about the close contacts were almost at the end of the questionnaire.

Also, the contacts who did not work were a part of the analysed dataset but were not really analysed in this study. The comparison was mainly made between part-time and full-time working contacts. Adding not working contacts would give greater contrast on the amount of leisure time the network has and could bring other, perhaps significant, results.

A more statistical limitation of this study is the fact that the normality assumption was violated. Therefore, there has been chosen for a smaller alpha, which made it harder to find significant results. Though, it must be noted that the dependent variable, social integration, is highly likely to be skewed in the population. Reliable analyses are therefore difficult to obtain.

What does strengthen this study is that the sample size is big, which compensates for the statistical problems of this study. Furthermore, the procedures are described in a detailed way, wherethrough this research is replicable. Though, it is recommended to investigate this topic in a different way. That is, measuring social integration and network diversity in a more objective way, as in not merely questionnaires. This means, measuring the quantity and quality and characteristics of the ties. This could be in combination with questionnaires and social network visualizations. This way more characteristics could be measured and network diversity will be handled in a broader way, wherethrough more information about the effect

of diversity will become known. Network diversity can be studied better in this way, as with statistical variables it is more difficult than measuring it with special network software as UCINET. Additionally, no other statistical problems about skew distributions etcetera will occur. Therefore, I consult to do further research about this topic with social network analysis, with for example UCINET.

Literature

About the Panel | LISS Panel Data. (z.d.). LISS PANEL. Retrieved March 21, 2022, from https://www.lissdata.nl/about-panel

Baumeister, R. F., & Leary, M. R. (1995). The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation. In *Psychological Bulletin, 117*(3), 497-529.

Berger, J., Cohen, B. P., & Zelditch, M. (1972). Status Characteristics and Social Interaction. In *American Sociological Review, 37*(3), 2241-255. Retrieved from https://www.jstor.org/stable/2093465

Buckley, K. E., Winkel, R. E., & Leary, M. R. (2004). Reactions to acceptance and rejection: Effects of level and sequence of relational evaluation. *Journal of Experimental Social Psychology*, *40*(1), 14–28. https://doi.org/10.1016/S0022-1031(03)00064-7

Cassie, K. M., Miller-Cribbs, J., & Smith, A. (2020). An exploratory study of factors associated with social isolation and loneliness in a community sample. *Social Work in Health Care*, *59*(7), 485–498. https://doi.org/10.1080/00981389.2020.1795780

Chang, P. J., Wray, L., & Lin, Y. (2014). Social relationships, leisure activity, and health in older adults. *Health Psychology*, *33*(6), 516–523. https://doi.org/10.1037/hea0000051

Centerdata. (2021, March). Social Integration and Leisure LISS Core Study (wave 13) [Dataset and codebook] Centerdata.

Chung, K. K., Hossain, L., & Davis, J. (2005, November). Exploring sociocentric and egocentric approaches for social network analysis. In *Proceedings of the 2nd international conference on knowledge management in Asia Pacific* (pp. 1-8).

Doolaard, F. T., Lelieveld, G. J., Noordewier, M. K., van Beest, I., & van Dijk, E. (2020). Get out or stay out: How the social exclusion process affects actors, but not targets.

Journal of Experimental Social Psychology, *88*.
https://doi.org/10.1016/j.jesp.2019.103946

Eckhard, J. (2018). Does Poverty Increase the Risk of Social Isolation? Insights Based on Panel Data from Germany. *Sociological Quarterly*, *59*(2), 338–359. https://doi.org/10.1080/00380253.2018.1436943

European Institute for Gender Equality. (2021). *Gender Equality Index 2020 NETHERLANDS*.

Firestone J, & Shelton B. (1988). An Estimation of the Effects of Women's Work on Available Leisure Time. *Journal of Famliy Issues*, *9*(4), 478-495.

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, *73*(2), 349–360. https://doi.org/10.1093/poq/nfp031

Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, *78*(6), 1360-1380.

Hampton, K. N., Sessions, L. F., & Her, E. J. (2011). Core networks, social isolation, and new media: How internet and mobile phone use is related to network size and diversity. *Information Communication and Society*, *14*(1), 130–155. https://doi.org/10.1080/1369118X.2010.513417

Kuczynski, A. M., Halvorson, M. A., Slater, L. R., & Kanter, J. W. (2021). The effect of social interaction quantity and quality on depressed mood and loneliness: A daily diary study. *Journal of Social and Personal Relationships*. *39*(3), 734-756. https://doi.org/10.1177/02654075211045717

Louch, H. (2000). Personal network integration: transitivity and homophily in strong-tie relations. In *Social Networks, 22*. 45-64

Lozares, C., Verd, J. M., Cruz, I., & Barranco, O. (2014). Homophily and heterophily in personal networks. From mutual acquaintance to relationship intensity. *Quality and Quantity*, *48*(5), 2657–2670. https://doi.org/10.1007/s11135-013-9915-4

Mcpherson, M., Smith-Lovin, L., & Cook, J. M. (2001). *BIRDS OF A FEATHER: Homophily in Social Networks*. *27*, 415-444. Retrieved from www.annualreviews.org

Neves, B. B. (2015). Does the Internet matter for strong ties? Bonding social capital, Internet use, and age-based inequality. *International Review of Sociology*, *25*(3), 415–433. https://doi.org/10.1080/03906701.2015.1050307

Olson, E. A., Pizzagalli, D. A., & Rosso, I. M. (2021). Social Anhedonia is Associated with Low Social Network Diversity in Trauma-Exposed Adults. *Journal of Traumatic Stress*, *34*(1), 241–247. https://doi.org/10.1002/jts.22577

Robins, G. (2015). Doing Social Network Research: Network-based Research Design for Social Scientists. In *Doing Social Network Research: Network-based Research Design for Social Scientists*. https://doi.org/10.4135/9781473916753

Scott, J. (1988). TREND REPORT SOCIAL NETWORK ANALYSIS. In *Source: Sociology, 22*(1), 109-127. Retrieved May 23, 2022, from https://www.jstor.org/stable/42854660

Seeman, T. E. (1996). Social Ties and Health: The Benefits of Social Integration. In *Science Inc. Ann Epidemiol, 6*(5), 442-451.

Smith-Lovin, L., Brashears, M., & McPherson, M. (2008). The Ties that Bind are Fraying. *Contexts*, *7*(3), 32–36. https://doi.org/10.1525/ctx.2008.7.3.32

Son, J., & Lin, N. (2012). Network diversity, contact diversity, and status attainment. *Social Networks*, *34*(4), 601–613. https://doi.org/10.1016/j.socnet.2012.06.006

Stalker, G. J. (2008). Measuring diversity in daily social contact: The contribution of social context, work and leisure on the opportunity for engagement. *Social Indicators Research*, *86*(2), 275–295. https://doi.org/10.1007/s11205-007-9115-5

Stewart, M. J., Makwarimba, E., Reutter, L. I., Veenstra, G., Raphael, D., & Love, R. (2009). Poverty, sense of belonging and experiences of social isolation. *Journal of Poverty*, *13*(2), 173–195. https://doi.org/10.1080/10875540902841762

Teater, B., Chonody, J. M., & Davis, N. (2021). Risk and Protective Factors of Loneliness among Older Adults: The Significance of Social Isolation and Quality and Type of Contact. *Social Work in Public Health*, *36*(2), 128–141. https://doi.org/10.1080/19371918.2020.1866140

Tomaskovic-Devey, B. (1993). Women, Men and Time: Gender Differences in Paid Work, Housework and Leisure (Book). *Social Forces*, *71*(4), 1094–1095.

Tonković, Ž., Cepić, D., & Puzek, I. (2021). Loneliness and Social Networks in Europe: ISSP Data from 13 European Countries. *Revija Za Sociologiju*, *51*(3), 381–407. https://doi.org/10.5613/rzs.51.3.3

Warren, T. (2010). Work time. leisure time. on women's temporal and economic well-being in Europe. *Community, Work and Family*, *13*(4), 365–392. https://doi.org/10.1080/13668801003765713

Yerkes, M. A., Roeters, A., & Baxter, J. (2020). Gender differences in the quality of leisure: a cross-national comparison. *Community, Work and Family*, *23*(4), 367–384. https://doi.org/10.1080/13668803.2018.1528968

Appendixes

Appendix 1

Exploration of original items, operationalisations of scale constructions and frequencies of eventual variables

1.   Social integration

For constructing the variable social integration, I used 7 questions. In table 1 a schematic overview of the used items and their operationalisations can be seen. The operationalisations mentioned below are already performed before the descriptive analysis later in this appendix. Because the only operationalisations are mirroring and recoding a category to missing, this will cause no troubles. It will only change the interpretation of the statistics.

Syntax

RECODE cs20m285 cs20m286 cs20m287 (1=3) (3=1).

EXECUTE.

RECODE cs20m283 (-9=SYSMIS).

EXECUTE.

Table 1 Original items of social integration and performed operationalisations

Variable	Question	Answer categories	Scale	Operationalisations
cs20m283	How satisfied are you with your social contacts?	0-10 0= not at all satisfied 10= completely satisfied -9 I don't know	Ordinal	-9 → missing
cs20m284	I have a sense of emptiness around me	1=Yes 2= More or less 3= No	Ordinal	
cs20m285	There are enough people I can count on in case of a misfortune	1=Yes 2= More or less 3= No	Ordinal	Mirrored 1=No 2= More or less 3= Yes
cs20m286	I know a lot of people that I can fully rely on	1=Yes 2= More or less 3= No	Ordinal	Mirrored 1=No 2= More or less 3= Yes
cs20m287	There are enough people to whom I feel closely connected	1=Yes 2= More or less 3= No	Ordinal	Mirrored 1=No 2= More or less 3= Yes
cs20m288	I miss having people around me	1=Yes 2= More or less 3= No	Ordinal	
cs20m289	I often feel deserted	1=Yes 2= More or less 3= No		

In table 2 the descriptive statistics of the original items are shown.

Syntax

FREQUENCIES VARIABLES=cs20m283 cs20m284 cs20m285 cs20m286 cs20m287 cs20m288 cs20m289

 /STATISTICS=STDDEV MEAN MEDIAN

 /BARCHART FREQ

 /ORDER=ANALYSIS.

Table 2 Descriptive statistics of the items of social integration

Statistics

		How satisfied are you with your social contacts?	I have a sense of emptiness around me	There are enough people I can count on in case of a misfortune	I know a lot of people that I can fully rely on	There are enough people to whom I feel closely connected	I miss having people around me	I often feel deserted
N	Valid	4544	4544	4544	4544	4544	4544	4544
	Missing	0	0	0	0	0	0	0
Mean		7,19	2,71	2,74	2,57	2,68	2,55	2,83
Median		7,00	3,00	3,00	3,00	3,00	3,00	3,00
Std. Deviation		1,746	,553	,517	,629	,562	,684	,461

The respondents seem to be quite satisfied with their social contacts as they averagely give it a 7,12. Also, the respondents do not seem to feel a sense of emptiness very much as the median is 3 (=no) and the mean is 2,71. Additionally, the respondents on average feel like they have enough people to count on in case of misfortune and rely on (mean=2,74 and mean=2,57). They also feel quite connected and do not feel deserted or miss people around them very much (means= 2,68; 2,54; 2,83).

1.2   Frequency distributions of the items of social integration

In table 3 till 10 frequency tables of the items are shown. In figure 1 till 7 the frequency distributions of the items are visualised in a bar chart.

Syntax:

FREQUENCIES VARIABLES=cs20m283 cs20m284 cs20m285 cs20m286 cs20m287 cs20m288 cs20m289

 /STATISTICS=STDDEV MEAN MEDIAN

 /BARCHART FREQ

 /ORDER=ANALYSIS.

Table 3 Frequency table of item: How satisfied are you with your social contacts?

How satisfied are you with your social contacts?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Not at all satisfied	28	,6	,6	,6
	1	21	,5	,5	1,1
	2	60	1,3	1,3	2,4
	3	93	2,0	2,0	4,4
	4	114	2,5	2,5	7,0
	5	288	6,3	6,3	13,3
	6	492	10,8	10,8	24,1
	7	1280	28,2	28,2	52,3
	8	1406	30,9	30,9	83,2
	9	407	9,0	9,0	92,2
	Completely satisfied	355	7,8	7,8	100,0
	Total	4544	100,0	100,0	
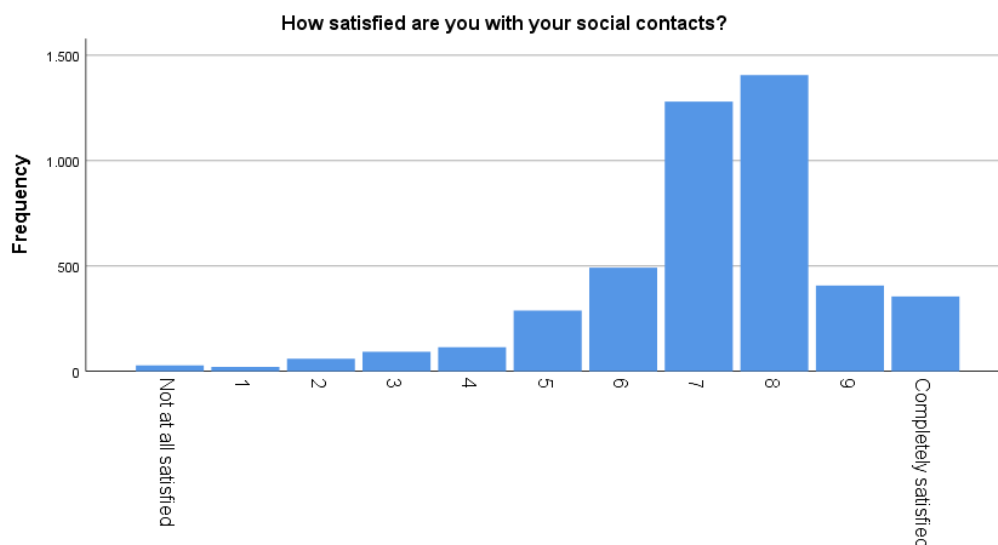
Figure 1 Frequency distribution of the item: How satisfied are you with your social contacts?

First the distribution of the item of satisfaction of the social contacts looks left skewed in figure 1. The numbers 7 and 8 are the most common given answers. People almost never were completely not satisfied with there social contacts, but the other extreme answers towards completely satisfied are more given than the bottom extreme answers.

Table 4 Frequency table of the item: I have a sense of emptiness around me

I have a sense of emptiness around me

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	226	5,0	5,0	5,0
	More or less	864	19,0	19,0	24,0
	No	3454	76,0	76,0	100,0
	Total	4544	100,0	100,0	
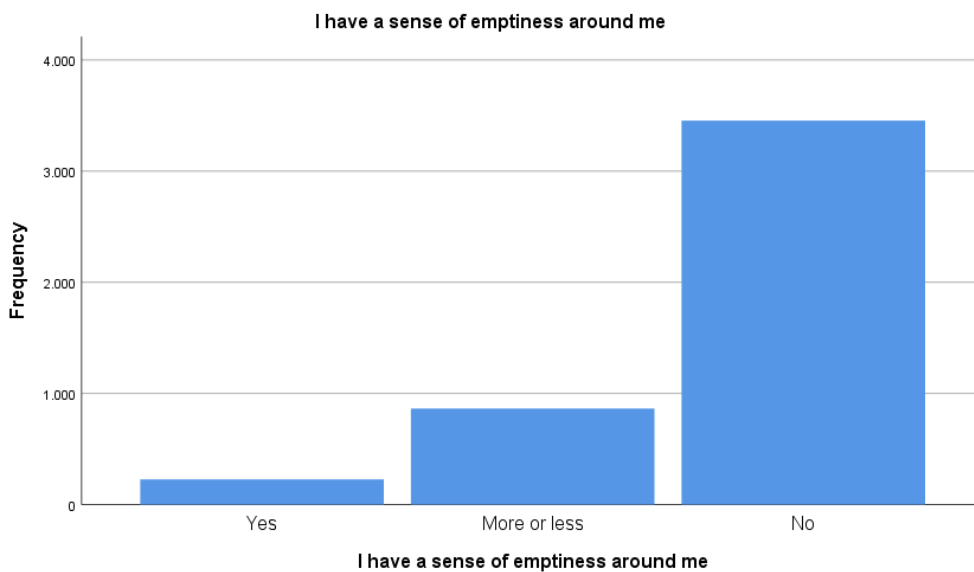
Figure 2 Frequency distribution of the item: I have a sense of emptiness around me

In table 4 and figure 2 the item "I have a sense of emptiness around me" looks rightly skewed. Most of the respondents do not feel a sense of emptiness around them.

Table 5 Frequency table of the item: There are enough people I can count on in case of a misfortune

There are enough people I can count on in case of a misfortune

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	167	3,7	3,7	3,7
	More or less	859	18,9	18,9	22,6
	No	3518	77,4	77,4	100,0
	Total	4544	100,0	100,0	
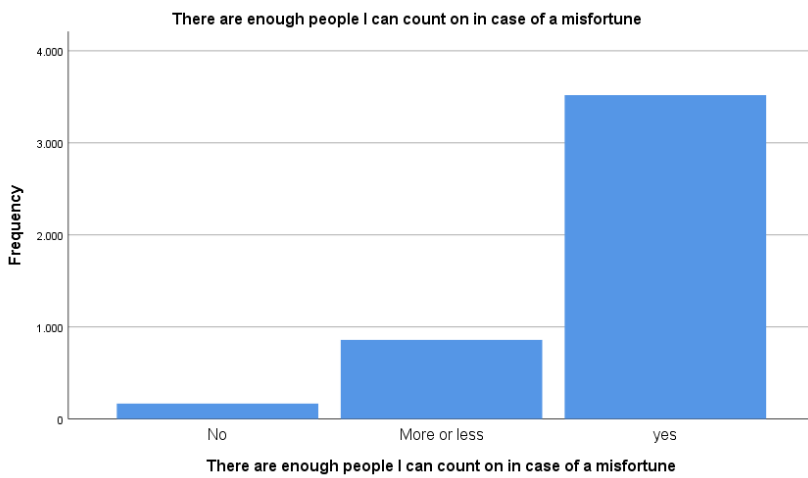
Figure 3 Frequency distribution of the item: There are enough people I can count on in case of a misfortune

In table 5 and figure 3 the frequencies of the item "There are enough people I can count on in case of a misfortune" are shown. This item is heavily rightly skewed, the majority of the respondents feel like they have enough people to count on in case of a misfortune.

Table 6 Frequency table of the item: I know a lot of people that I can fully rely on

I know a lot of people that I can fully rely on

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	341	7,5	7,5	7,5
	More or less	1266	27,9	27,9	35,4
	Yes	2937	64,6	64,6	100,0
	Total	4544	100,0	100,0	
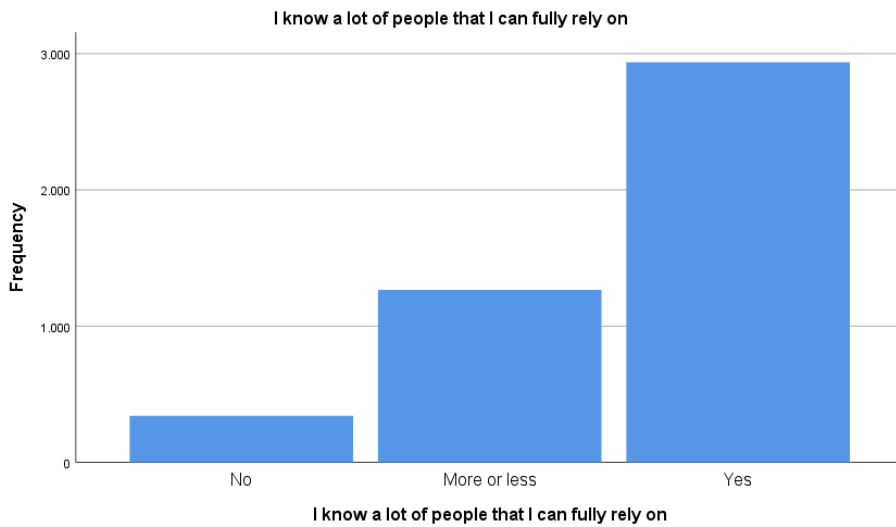
Figure 4 Frequency distribution of the item: I know a lot of people that I can fully rely on

In table 6 and figure 4 the frequencies of the item "I know a lot of people that I can fully rely on" is shown. This item is heavily rightly skewed. The majority of the respondents have enough people to fully rely on.

Table 7 Frequency table of the item: There are enough people to whom I feel closely connected

There are enough people to whom I feel closely connected

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	219	4,8	4,8	4,8
	More or less	1033	22,7	22,7	27,6
	Yes	3292	72,4	72,4	100,0
	Total	4544	100,0	100,0	
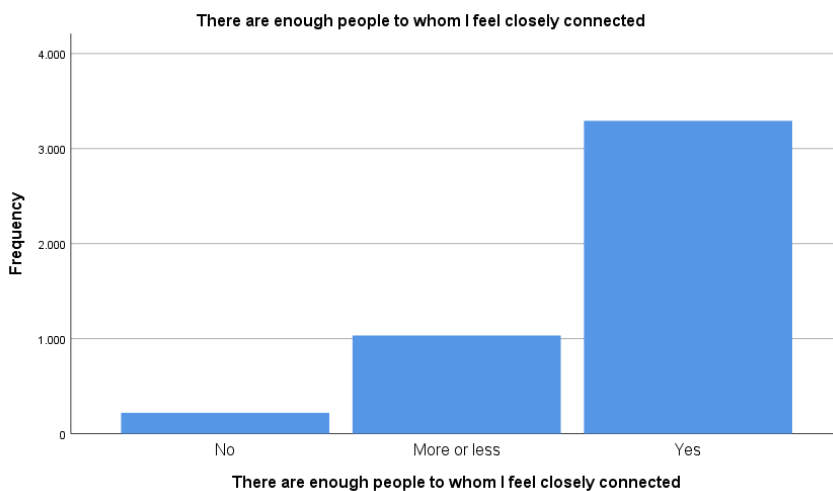
Figure 5 Frequency table of the item: There are enough people to whom I feel closely connected

In Table 7 and figure 5 the frequencies of the item "There are enough people to whom I feel closely connected" is shown. This item is also very rightly skewed. Most people feel like they are closely connected to enough people.

Table 8 *Frequency table of the item: I miss having people around me*

I miss having people around me

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	499	11,0	11,0	11,0
	More or less	1050	23,1	23,1	34,1
	No	2995	65,9	65,9	100,0
	Total	4544	100,0	100,0	
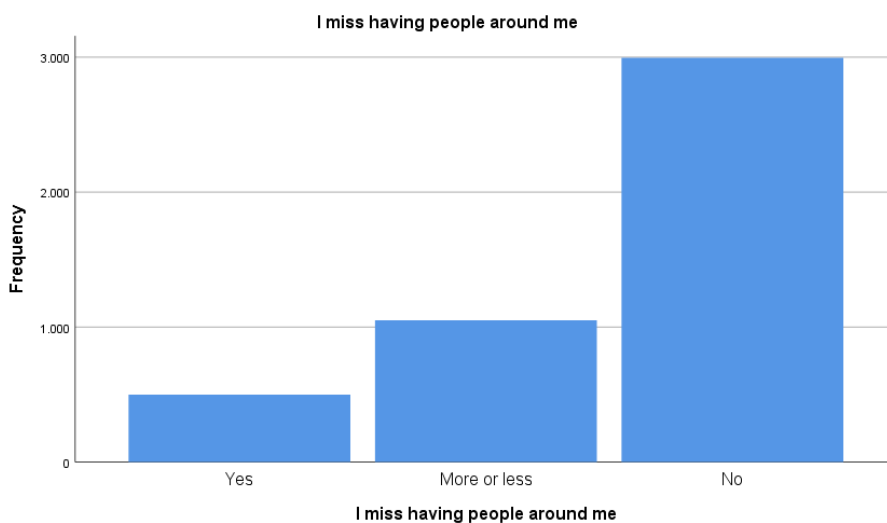
Figure 6 *Frequency distribution of the item: I miss having people around me*

In table 8 and figure 6 the frequencies of the item "I miss having people around me" is shown. This item is also very rightly skewed distributed. Most people do not miss having people around them.

Table 9 Frequency table of the item: I often feel deserted

I often feel deserted

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	162	3,6	3,6	3,6
	More or less	451	9,9	9,9	13,5
	No	3931	86,5	86,5	100,0
	Total	4544	100,0	100,0	
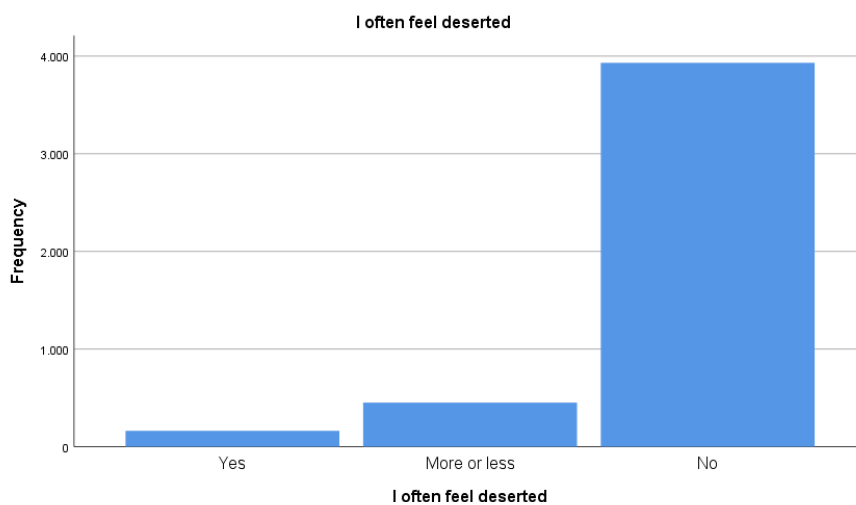
Figure 7 Frequency distribution of the item: I often feel deserted

Lastly, in table 9 and figure 7 the frequencies of the item "I often feel deserted" is shown. This variable is also very rightly skewed distributed. The vast majority of the respondents do not feel deserted often.

1.3 Operationalisations of the items

cs20m285, cs20m286 and cs20m287 were mirrored, so that for all items a higher score represents a positive feeling. For the item cs20m283, I only recoded the answer category -9 into missing. Then, for constructing the eventual variable, I made a sum of all the seven items, which I called Social integration. It has a Cronbach's Alpha of 0,743, the scale looks thus reliable. This variable runs from 6-28, where a higher score represents a better social integration feeling.

For the eventual variable Social integration, the descriptive statistics are shown in table 10, the distribution is shown in figure 8. The operationalisations have been done with the following syntax.

SYNTAX

COMPUTE Social_Integration=cs20m283 + cs20m284 + cs20m285 + cs20m286 + cs20m287 + cs20m288 +

cs20m289.

EXECUTE.

FREQUENCIES VARIABLES=Social_Integration

/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN SKEWNESS SESKEW KURTOSIS SEKURT

/BARCHART FREQ

/ORDER=ANALYSIS.

RELIABILITY

/VARIABLES=cs20m283 cs20m284 cs20m285 cs20m286 cs20m287 cs20m288 cs20m289

/SCALE('ALL VARIABLES') ALL

/MODEL=ALPHA

Table 10 Descriptive statistics o the scale variable Social integration

Statistics

Social_Integration

N	Valid	4544
	Missing	0
Mean		19,2980
Median		20,0000
Std. Deviation		2,37468
Skewness		-,955
Std. Error of Skewness		,036
Kurtosis		1,915
Std. Error of Kurtosis		,073
Minimum		8,00
Maximum		28,00

In table 10 can be seen that on average respondents scored 19,298 on the scale of social integration and the variable runs from the score 8 till 28. This implicates a quite good feeling of social integration. In figure 8 the distribution can be seen. There seems to be a little bit of a left-skewed distribution. Also, almost no scores above the 25 were observed. In table 10 the skewness is -0,995 and the kurtosis is 1,915. For the skewness and kurtosis, I use the range of acceptance between values of -3 and 3. There seems to be no problems here as both values fall within the acceptable range. The distribution looks quite normal.
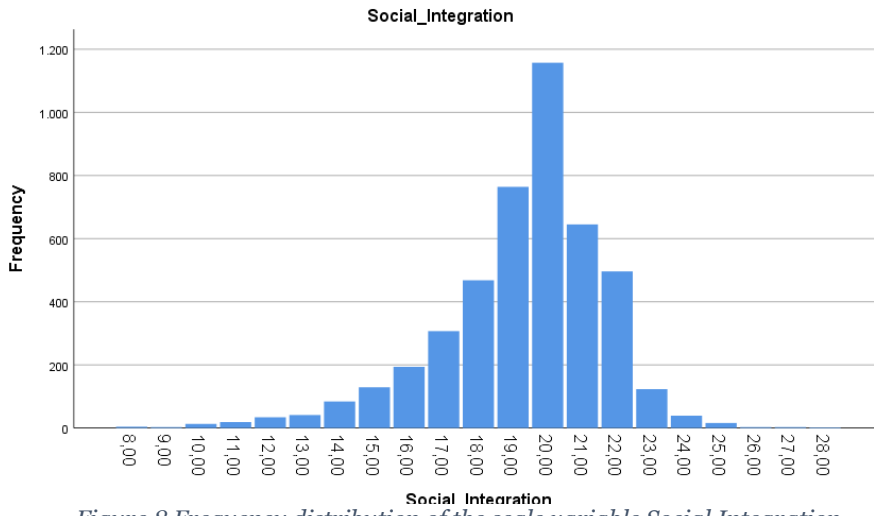
Figure 8 Frequency distribution of the scale variable Social Integration

2.   Network diversity in terms of employment

For network diversity in terms of employment the question: *"Does [person 1/2/3/4/5] work full-time, part-time, or not at all?"* was used. This question only took answers into account for respondents that assigned 5 persons as close contacts. 5 items belong to this question, one for every person. These are: cs20m327 (person 1), cs20m338 (person 2), cs20m349 (person 3), cs20m360 (person 4), cs20m371 (person 5). In table 10 the descriptive of these items are shown.

SYNTAX

FREQUENCIES VARIABLES=cs20m327 cs20m338 cs20m349 cs20m360 cs20m371

  /STATISTICS=STDDEV MEAN MEDIAN

  /BARCHART FREQ

  /ORDER=ANALYSIS.

In table 11 the descriptive statistics of the items can be seen. What immediately stands out is that the missing values of the items increases with the person mentioned. Respondents reported gradually reported less contacts. Therefore, the second till fifth contacts are less represented then the first. This can cause distorted results. Every person mentioned seems to averagely work part-time as the mean of every person is around the 1,9.

Table 11 Descriptive statistics of the items for network diversity in terms of employment

		Does person 1 work full-time, part-time, or not at all?	Does person 2 work full-time, part-time, or not at all?	Does person 3 work full-time, part-time, or not at all?	Does person 4 work full-time, part-time, or not at all?	Does person 5 work full-time, part-time, or not at all?
Statistics						
N	Valid	4534	3836	3046	2120	1499
	Missing	10	708	1498	2424	3045
Mean		1,95	1,92	1,93	1,95	1,96
Median		2,00	2,00	2,00	2,00	2,00
Std. Deviation		,855	,839	,828	,837	,837

In figure 9 till 13 the frequency distributions of the items are shown.
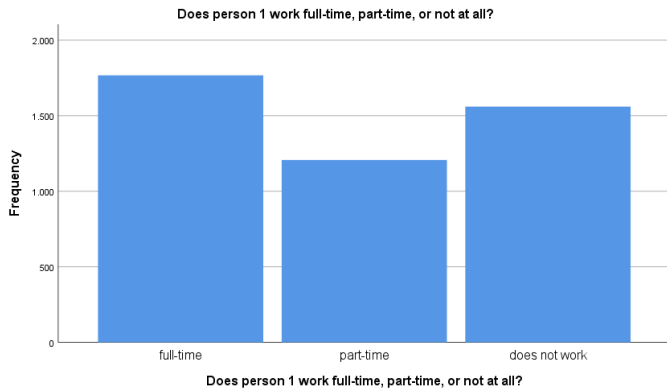

Figure 9 Frequency distribution of the item: Does person 1 work full-time, part-time or nor at all?
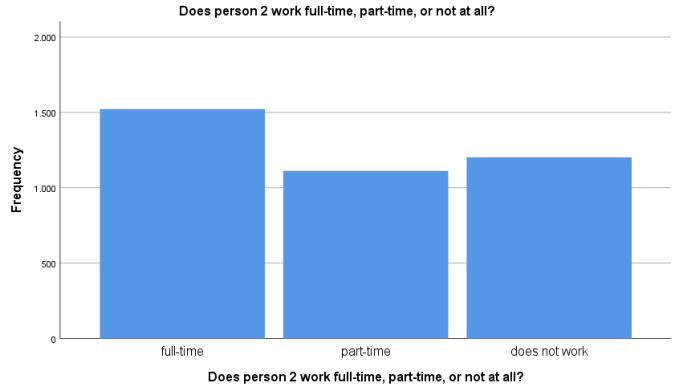

Figure 10 Frequency distribution of the item: Does person 2 work full-time, part-time or not at all?
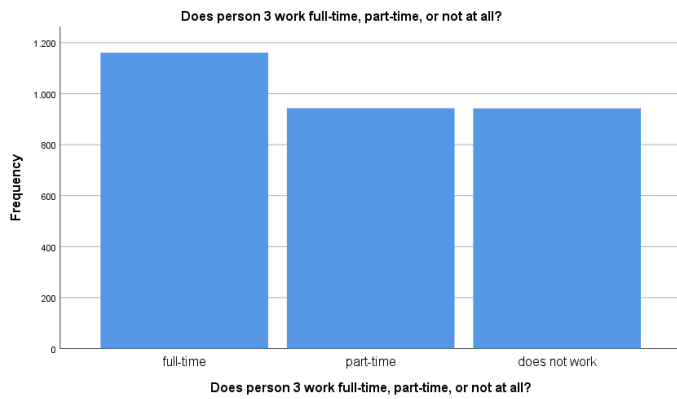

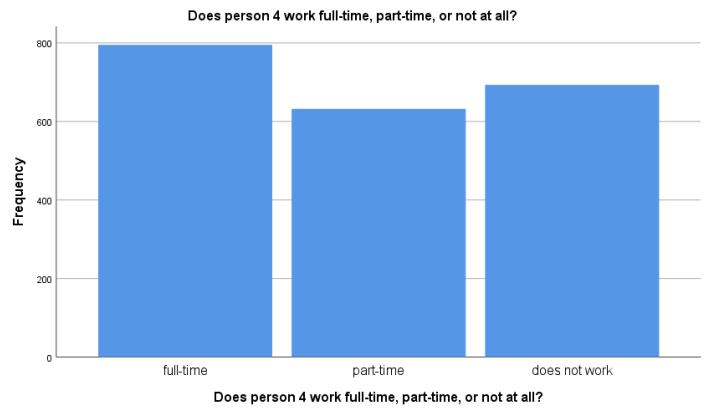*Figure 11 Frequency distribution of the item: Does person 3 work full-time, part-time or not at all?*

Figure 12 Frequency distribution of the item: Does person 4 work full-time, part-time or not at all?

The items are not normally distributed, but not clearly skewed to one side. In contrast with the mean, the distributions show that most people work full-time. For every person most of the people work full-time.
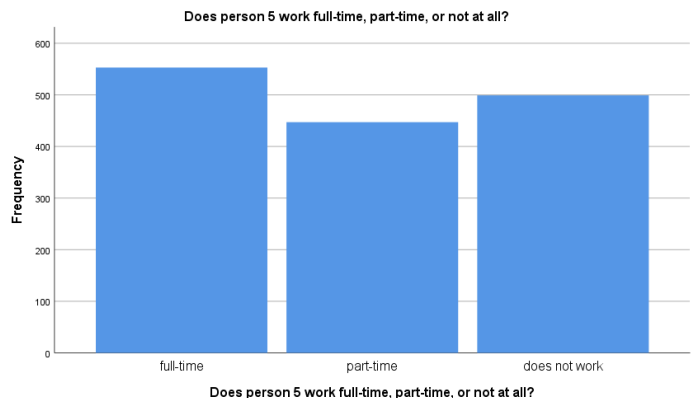
Figure 13 Frequency distribution of the item: Does person 5 work full-time, part-time or not at all?

2.2   *Operationalisations*

I recoded the answer category -9 into missing for all items, as this is answer is not useful.
Then I mirrored the answer categories 1-3 for all items, so that a higher score represents that
the respondent's contact works more. These operationalisations were already performed
before the descriptive and frequency analysis. Next, I made 2 dummies for every item. One
dummy for every item where 1=full-time and 0=part-time and not at all (fullt1, fullt2, fullt3,
fullt4 and fullt5). The second dummy for each item is coded into 1=part-time and 0=full-time
and not at all (partt1, partt2, partt3, partt4 and partt5). Eventually, I made a mean score
variables of the full-time dummies, called propfull, and the part-time dummies, called
proppart. This way, proportions of the employment types of the contacts are represented. The
rest proportion is the proportion of contacts that do not work. I also made centred versions of
propfull and proppart called Cpropfull and Cproppart, so that they can be used for the
moderation analysis.

SYNTAX

RECODE cs20m327 cs20m338 cs20m349 cs20m360 cs20m371 (-9=SYSMIS) (1=3) (3=1).
EXECUTE.

RECODE cs20m327 cs20m338 cs20m349 cs20m360 cs20m371 (1=0) (2=1) (3=0) INTO Partt1 partt2 partt3
    partt4 partt5.
VARIABLE LABELS Partt1 'parttime_1' /partt2 'parttime_2' /partt3 'partime3' /partt4 'parttime_4'
    /partt5 'parttime_5'.
EXECUTE.

RECODE cs20m327 cs20m338 cs20m349 cs20m360 cs20m371 (1=0) (2=0) (3=1) INTO fullt1 fullt2 fullt3
    fullt4 fullt5.
VARIABLE LABELS fullt1 'fulltime_1' /fullt2 'fulltime_2' /fullt3 'fulltime_3' /fullt4 'fulltime_4'
    /fullt5 'fulltime_5'.
EXECUTE.

COMPUTE proppart=MEAN(Partt1,partt2,partt3,partt4,partt5).
EXECUTE.

COMPUTE propfull=MEAN(fullt1,fullt2,fullt3,fullt4,fullt5).
EXECUTE.

```
DATASET ACTIVATE DataSet3.
COMPUTE Cproppart=proppart - 0.2768.
EXECUTE.


COMPUTE Cpropfull=propfull - 0.3851.
EXECUTE.


FREQUENCIES VARIABLES=proppart propfull
  /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN
  /BARCHART FREQ
  /ORDER=ANALYSIS.
```

The descriptive statistics of the eventual variables propfull and proppart are shown in table 12.

Table 12 Descriptive statistics of the variables propfull and proppart

Statistics

		proppart	propfull
N	Valid	4544	4544
	Missing	0	0
Mean		,2784	,3368
Median		,2000	,2500
Std. Deviation		,30347	,33752
Minimum		,00	,00
Maximum		1,00	1,00

The mean of proppart is 0,278. This means that averagely the proportion part-time working contacts of the respondents is 27,8%. The mean of propfull is 0,337. This means that averagely 33,7% of the contacts of the respondents work full-time. This leaves an average proportion of 38,5% of contacts that work not at all. On average the most contacts of the respondents do not work at all.
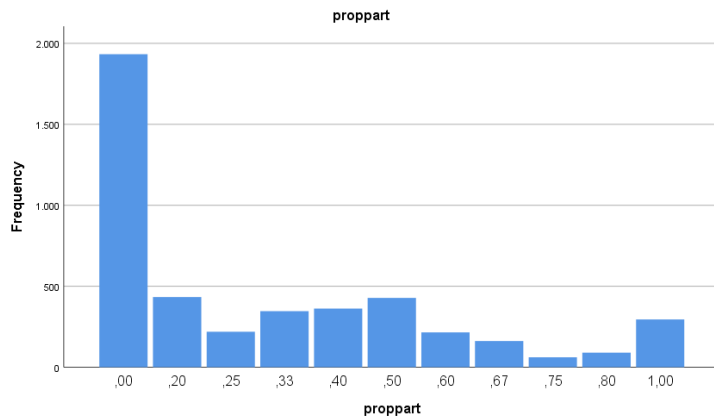
Figure 14 Frequency distribution of the variable proppart

In Figure 14 the frequency distribution of the proportion part-time working contacts (proppart) is visualised. The distribution looks rightly skewed. The proportions are mostly very close to 0.
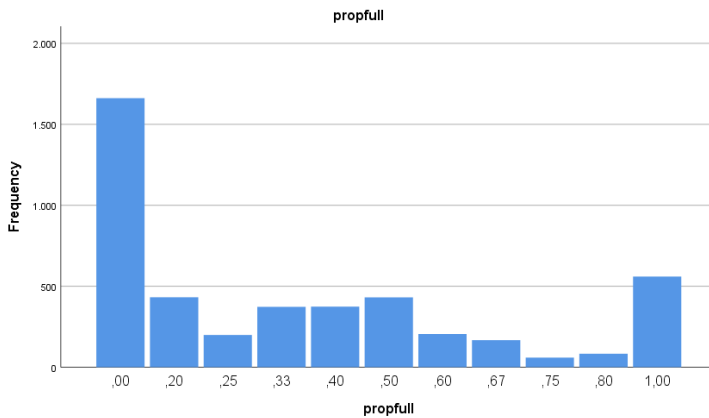

Figure 15 Frequency distribution of the variable propfull

In figure 15 the frequency distribution of the proportion full-time working contacts (propfull) is visualised. Also, this variable looks rightly skewed. Most proportions are close to zero. For both variables this is not very strange as within the answer category 0 of the dummies within the variables fall more categories that the score 1. Therefore, the chance on scoring a 1 is smaller than scoring a 0. The skewed distributions therefore are a consequence of the operationalisations and do not seem problematic.

3.   Gender

For the variable gender I used the variable 'geslacht' where 1= male and 2= female. I Made a new dummy called Gender where 1=female and 0= male. I Will analyse this dummy as the statistics and distributions of both geslacht and Gender are not different. In table 13 the descriptive statistics of Gender are shown.

SYNTAX

/STATISTICS=STDDEV MINIMUM MAXIMUM MODE

/BARCHART FREQ

/ORDER=ANALYSIS.

Table 13 Descriptive statistics of the dummy variable Gender

Statistics

Gender

N	Valid	4544
	Missing	0
Mode		1,00
Std. Deviation		,49538
Minimum		,00
Maximum		1,00

In table 13 can be seen that the mode of Gender is 1. This means that there are more female respondents than male.

In table 14 the frequency distribution of the variable is shown. In figure 16 the frequency distribution is visualised.

Table 14 Frequency table of the dummy variable Gender

Gender

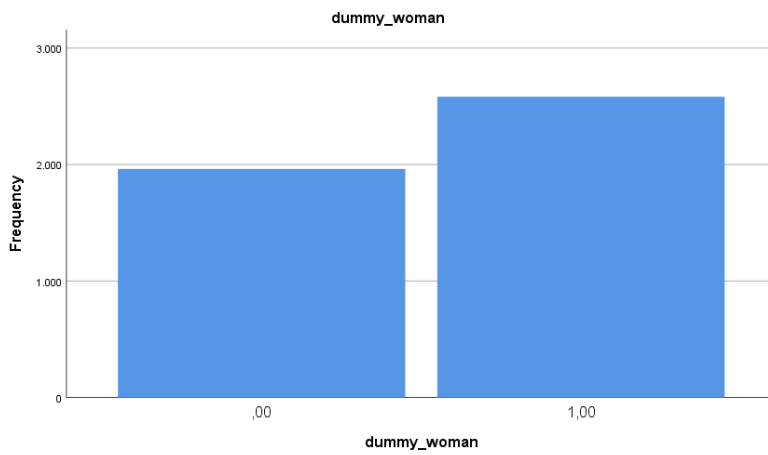|		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	1962	43,2	43,2	43,2
	1,00	2582	56,8	56,8	100,0
	Total	4544	100,0	100,0	

Figure 16 Frequency distribution of the dummy variable Gender

In table 14 can be seen that 56.8% of the respondents are female and 43.2% are male. In figure 16 it also can be seen that he majority of respondents is female. The distribution is therefore left skewed.

3. Age

For the variable age I used the variable "leeftijd" which represents the age of the household member, otherwise said, it represents the age of the respondent. No operationalisations were performed for this variable. The descriptive statistics of leeftijd are shown in table 15.

SYNTAX

FREQUENCIES VARIABLES=leeftijd

 /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN

 /BARCHART FREQ

 /ORDER=ANALYSIS.

Table 15 Descriptive statistics of the variable leeftijd

Statistics

Age of the household member

N	Valid	4544
	Missing	0
Mean		51,16
Median		53,00
Std. Deviation		18,444
Minimum		16
Maximum		103

The mean from table 15 is 51,16. The average age of the respondents is thus 51,16. The median is 53, which means that half of the respondents are younger than 53 and half of the respondents are older than 53. The oldest respondent is 103.

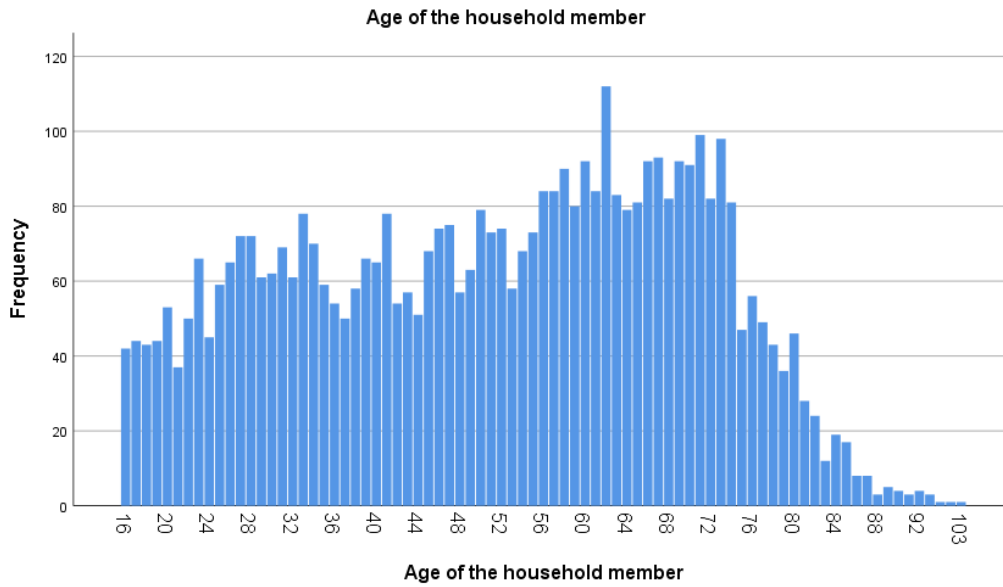In figure 17 the frequency distribution of leeftijd is visualized.

Figure 17 Frequency distribution of the variable leeftijd

In figure 17 can be seen that the variable leeftijd is not really normally distributed and looks right skewed. The observations drastically decrease after the age of 73.

4.   Domestic situation

For the domestic situation of the respondents, I used the variable woonvorm, which represents the domestic situation of the respondent. Where 1=single, 2=(un)married co-habitation, without child(ren), 3=(un)married co-habitation with child(ren), 4=single, with child(ren) and 5=other. I first recoded the answer category 5=other into missing as this information is not convenient.

SYNTAX

RECODE woonvorm (5=SYSMIS).

EXECUTE.

In table 16 the descriptive statistics of woonvorm are shown.

SYNTAX

FREQUENCIES VARIABLES=woonvorm

 /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN

 /BARCHART FREQ

 /ORDER=ANALYSIS.

Table 16 Descriptive statistics of the variable Domestic situation

Statistics

Domestic situation

N	Valid	4544
	Missing	0
Mean		2,19
Median		2,00
Std. Deviation		,880
Minimum		1
Maximum		4

In table 16 can be seen that on average people score a 2,19 which means that averagely the respondents live in cohabitation, married or not married without any children.

In table 17 the frequency table is shown and in figure 18 the frequency distribution is visualized.

Table 17 Frequency table of the variable Domestic situation

Domestic situation

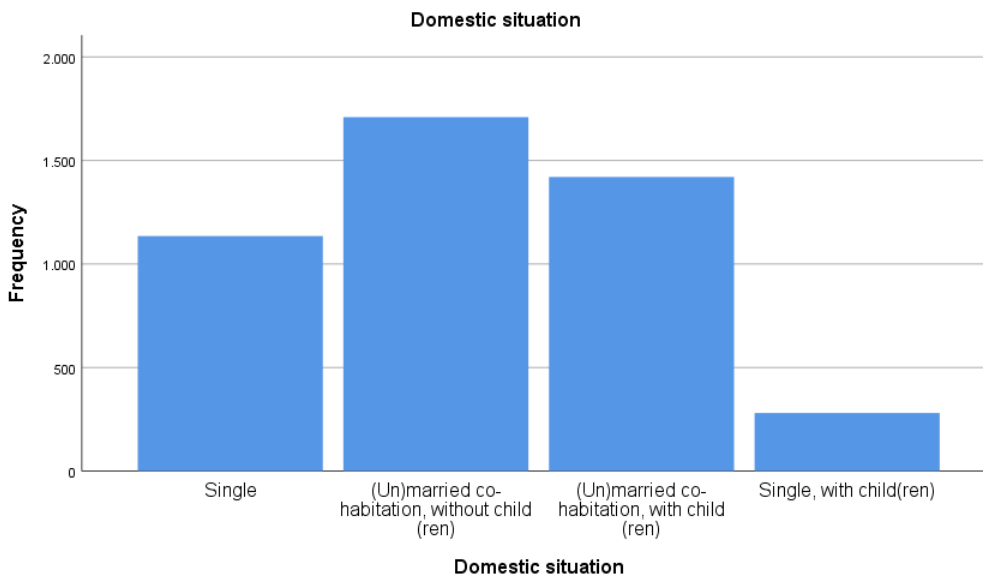|		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Single	1135	25,0	25,0	25,0
	(Un)married co-habitation, without child(ren)	1709	37,6	37,6	62,6
	(Un)married co-habitation, with child(ren)	1420	31,3	31,3	93,8
	Single, with child(ren)	280	6,2	6,2	100,0
	Total	4544	100,0	100,0	

Figure 18 Frequency distribution of the variable Domestic situation

In table 16 can be seen that 25% of the respondents live alone, 37,6% live in (un)married co-habitation without children, 31,3% live in (un) married co-habitation with children and 6,2% live single with children. Most respondents thus live in (un)married co-habitation without children. This also can be seen in figure 18. The distribution seems a little bit rightly skewed. Single living respondents with children are the least represented group.

4.2   Operationalisations

To incorporate this variable into the regression analysis a few operationalisations needed to be made. I made 3 dummies. The first dummy is called Singlewith, 1= single, with children and 0= the rest of the categories Singlewith represents the single living respondents with children. The second dummy is called Cowith, where 1= (un)married co-habitation with children and 0=the rest of the categories. Cowith represents the (un)married co-habitation living respondents with children. The third dummy is called Cowithout, where 1= (un)married co-habitation without children and 0= the rest of the categories. Cowithout represents the respondents who live in (un)married co-habitation without children. The reference group (where all dummies score 0) are the respondents who live alone without children.

SYNTAX

RECODE woonvorm (1=0) (2=1) (3=0) (4=0) INTO Cowithout.

EXECUTE.

RECODE woonvorm (1=0) (4=0) (2=0) (3=1) INTO Singlewith.

EXECUTE.

RECODE woonvorm (1=0) (2=0) (4=1) (3=0) INTO Cowith.

EXECUTE.

In table 18 the descriptive statistics of the dummies are shown.

SYNTAX

FREQUENCIES VARIABLES=Singlewith Cowith Cowithout

 /STATISTICS=STDDEV MINIMUM MAXIMUM MODE

 /BARCHART FREQ

 /ORDER=ANALYSIS.

Table 18 Descriptive statistics of the dummies of Domestic situation

Statistics

		Singlewith	Cowith	Cowithout
N	Valid	4544	4544	4544
	Missing	0	0	0
Mode		,00	,00	,00
Std. Deviation		,46356	,24049	,48446
Minimum		,00	,00	,00
Maximum		1,00	1,00	1,00

Zero is the most common score for all three dummies, which is logical as more answer categories score a 0 than a 1 for each dummy.

It is more informative to look at the frequency distributions. In table 19 till 21 and figure 19 till 21 the frequency distributions are shown and visualised.

Table 13 Frequency table of the dummy Singlewith

Singlewith

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	3124	68,8	68,8	68,8
	1,00	1420	31,3	31,3	100,0
	Total	4544	100,0	100,0	

Table 4 Frequency table of the dummy Cowith

Cowith

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	4264	93,8	93,8	93,8
	1,00	280	6,2	6,2	100,0
	Total	4544	100,0	100,0	

Table 21 Frequency table of the dummy Cowithout

Cowithout

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	2835	62,4	62,4	62,4
	1,00	1709	37,6	37,6	100,0
	Total	4544	100,0	100,0	

The tables 19 till 21 show the same statistics as the original variable. 25% of the respondents live alone, 37,6% live in (un)married co-habitation without children, 31,3% live in (un) married co-habitation with children and 6,2% live single with children.
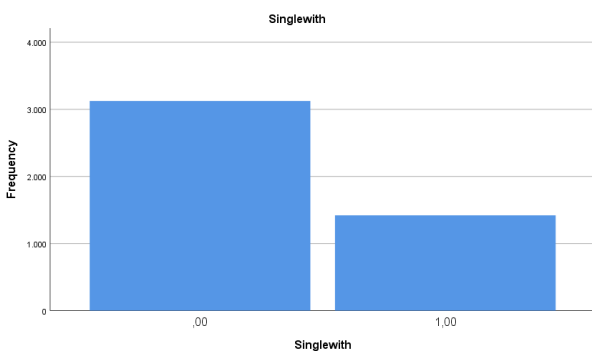

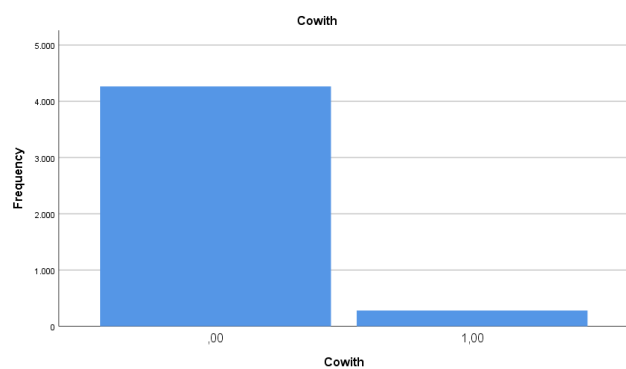*Figure 19 Frequency distribution of the dummy Singlewith*

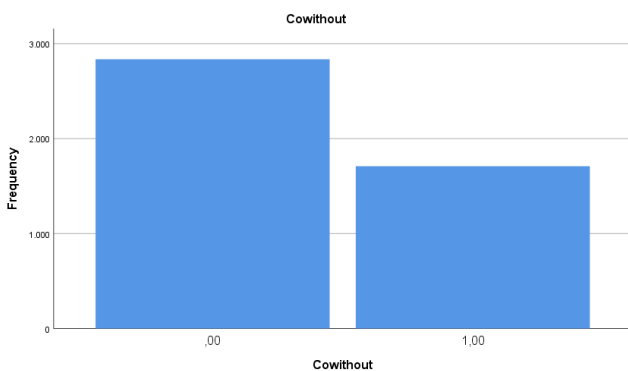*Figure 20 Frequency distribution of the dummy Cowith*

All three dummies are right skewed as the most common observation is 0.

Figure 21 Frequency distribution of the dummy Cowithout

## 5.	Income

For the control variable income, I used the variable nettoink_f. Which represents the personal net monthly income in Euros of the respondent. In table 21 the descriptive statistics of the original variable nettoink_f is shown and the distribution is visualized in figure 22. No further operationalisations have been performed.

SYNTAX

FREQUENCIES VARIABLES=nettoink_f

 /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN

 /BARCHART FREQ

 /ORDER=ANALYSIS.

Table 22 Descriptive statistics of the variable income

Statistics

Personal net monthly income in Euros, imputed

N	Valid	4544
	Missing	0
Mean		1869,95
Median		1750,00
Std. Deviation		3840,456
Minimum		0
Maximum		146652

The mean net income per month is 1869,95 euros. Half of the respondents earn less then 1750 euros per month and half earns more.

In figure 22 the frequency distribution of income is visualised. Here can be seen that a lot of people earn 0 euros per month. Therefore, the distribution is highly rightly skewed.

Figure 22 Frequency distribution of the variable income

6.   Dataset construction

A few modifications to the original LISS dataset have been made before the analyses. First, I exported the items and variables is need to another dataset which I called dataset 1. Then I did the operationalisations and constructed the scales and saved this in dataset2. Then I checked for respondents who gave no answer an all my items. No respondents had to be deleted. The checked dataset is saved as dataset3. Lastly, I made a dataset with no missing values. Below the syntax can be found. The sample size went from N=6795 to N=4544. In table 23 the differences in means of the complete (dataset 3, N=6795) and the not complete (dataset 4, N=4544).

SYNTAX

RECODE proppart propfull Social_Integration Singlewith Cowith Cowithout leeftijd nettoink_f
   geslacht (SYSMIS=1) (ELSE=0) INTO misproppart mispropfull missocint missinglwith misscowith
   miscowithout misage misincome misgender.
EXECUTE.

COMPUTE SumMissing=SUM(misproppart,mispropfull,missocint,missinglwith,misscowith,miscowithout,
   misage,misincome,misgender) .
EXECUTE.

USE ALL.
COMPUTE filter_$=(SumMissing = 0).
VARIABLE LABELS filter_$ 'SumMissing = 0 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.

DATASET COPY BAW4nomis.
DATASET ACTIVATE BAW4nomis.
FILTER OFF.
USE ALL.
SELECT IF (SumMissing = 0).
EXECUTE.
DATASET ACTIVATE DataSet1.

Table 23 Descriptive statistics of the complete and incomplete data

Variable	Mean (standard deviation)	
		Incomplete (N=4544)
	Complete (N=6795)	
Social integration	23,06 (3,89)	19,30 (2,37)
Network diversity employment	38,5% Propfull	33,68% Propfull
	27,7% Proppart	27,84% Proppart
	33,8% No work	38,48% No work
Gender	45,7% Male	43,2% Male
	54,3% Female	56,8% Female
Age	50,51 (18,83)	51,16 (18,44)
Income	1805,89 (3289,60)	1869,95 (3840,456)
Domestic situation	25,2% Single, without children	25,0% Single, without children
	34,4% co-habitation, without children	37,6% co-habitation, without children
	6,3% co-habitation, with children	6,2% co-habitation, with children
	33,1% Single, with children	31,3% Single, with children

There is one relatively big difference between the means of the complete and incomplete dataset. The means of social integration have changed approximately 4 points. This means that very socially integrated respondents

Appendix 2
Syntax, output and explanation of the analysis.

2.1 Descriptive and bivariate statistics

First a descriptive analysis of the variables has been done with the following syntax.

SYNTAX

FREQUENCIES VARIABLES=Social_Integration

 /STATISTICS=STDDEV RANGE MEAN

 /HISTOGRAM NORMAL

 /ORDER=ANALYSIS.

FREQUENCIES VARIABLES=proppart propfull

 /STATISTICS=STDDEV RANGE MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT

 /HISTOGRAM NORMAL

 /ORDER=ANALYSIS.

FREQUENCIES VARIABLES=Gender

 /STATISTICS=STDDEV RANGE MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT

 /BARCHART FREQ

 /ORDER=ANALYSIS.

FREQUENCIES VARIABLES=nettoink_f Age Cowithout Cowith

 /STATISTICS=STDDEV MEAN SKEWNESS SESKEW KURTOSIS SEKURT

 /HISTOGRAM

 /ORDER=ANALYSIS.

FREQUENCIES VARIABLES=woonvorm

 /ORDER=ANALYSIS.

The results can be seen in table 24.

Table 24 Descriptive statistics of all the variables included in the analysis with N=4544

Variable	Mean or proportion (standard deviation)[a]	Minimum	Maximum
Social integration	19,30 (2,37)	8	28
Network diversity of employment	33,68% Full-time	0	1
	27,84% Part-time		
	38,48% Does not work		
Gender	43,20% Man		
	56,80% Women	0	1
Age	51,16 (18,44)	16	103
Income	1869,95 (3840,46)	0	146.652
Domestic situation	25,0% Single, without children	0	1
	37,60% Co-habitation, without children		
	31,2% Co-habitation, with children		
	6,20% Single, with children		

N=4544

In table 1 can be seen that the mean for social integration is 19,3. This means that people feel quite socially integrated as the scale of the variable runs from 8-28. Then the employment types statistics of the close contacts are as follows. Most contacts do not work (38,48%) followed by the percentage contacts who work full time (33,68%). 27,8% of the contact works part-time. Most of the respondents were female (56,8%) and the average age of the respondents is 51,16. Averagely the respondents earned €1869,95 per month. However, the standard deviation is very big (s=3840,46). This indicates that the distribution of the incomes of the respondents is very skew. Some people earn much more than the average and some people earn much less. Lastly, the domestic situations of the respondents are as follows. Most of the respondents live in co-habitation, without children (37,6%). Then, 31,2% live in co-habitation with children, 25% lives alone and the least of the respondents live as a single parent (6,2%).

Then the correlations between the variables have been analysed. The following syntax has been used. For the correlations between the continuous variables.

SYNTAX

CORRELATIONS

 /VARIABLES=Social_Integration proppart propfull Age nettoink_f

 /PRINT=TWOTAIL NOSIG

 /MISSING=PAIRWISE.

Table 25 Correlations between the continuous variables incorporated in the analysis

Correlations

		Social_Integration	proppart	propfull	Age of the household member	Personal net monthly income in Euros, imputed
Social_Integration	Pearson Correlation	1	-,005	,038*	,125**	,027
	Sig. (2-tailed)		,741	,010	,000	,066
	N	4541	4541	4541	4541	4541
proppart	Pearson Correlation	-,005	1	-,452**	-,201**	-,036*
	Sig. (2-tailed)	,741		,000	,000	,014
	N	4541	4541	4541	4541	4541
propfull	Pearson Correlation	,038*	-,452**	1	,322**	-,009
	Sig. (2-tailed)	,010	,000		,000	,540
	N	4541	4541	4541	4541	4541
Age of the household member	Pearson Correlation	,125**	-,201**	,322**	1	,103**
	Sig. (2-tailed)	,000	,000	,000		,000
	N	4541	4541	4541	4541	4541
Personal net monthly income in Euros, imputed	Pearson Correlation	,027	-,036*	-,009	,103**	1
	Sig. (2-tailed)	,066	,014	,540	,000	
	N	4541	4541	4541	4541	4541

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

In table 25 can be seen that the proportion variables of employment are negatively correlated with each other (r=-0,452). The relationship is significant and quite big. The negative relationship is logical as the variables are constructed in such a way that if the value of the one increases, the other decreases. The correlation between the proportion part-time working contacts and the social integration of the respondents is r=0,005 and not significant. This is a very small correlation. No relation between the proportion part-time working contacts and social integration is present. As regards to the proportion full-time working contacts, the relation with social integration is bigger and significant as r=0,038.

Still not a very strong relation, but it can be said that an increase in the proportion full-time working contacts sometimes goes together with a higher score on social integration. The correlation between the part-time working contacts and age is $r=-0,201$ and significant. This is a moderately strong relation. It can be said that an increase in age goes together with a higher proportion of part-time working contacts. Also, the correlation between the proportion of full-time working contacts and age is quite strong and significant as is $r=-0,322$. This relation is stronger than the relation with the part-time proportion an age. An increase in age goes together with an increase of the proportion of full-time working contacts. The relationships between the employment type proportions and income is only significant for the part-time proportion. The $r=-0,036$, very small. The relationship is negative. An increase in part-time working contacts rarely goes together with a decrease in Income. The relationship between social integration and income is small and not significant ($r=0,027$). No relationship between social integration and income is found. A relation between age and social integration has been found. The correlation is significant and moderately strong ($r= 0,125$). An increase in age goes together with an increase in social integration. The relationship between income and age is moderately strong, positive and significant ($r=0,103$). An increase in income goes together with a higher age.

The relationships between gender and the continuous variables have been tested with a paired t-test with the following syntax.

SYNTAX

T-TEST PAIRS=Gender Gender Gender WITH Age Social_Integration Cproppart

   Cpropfull nettoink_f (PAIRED)

  /CRITERIA=CI(.9500)

Table 26 Paired sample correlations between gender and the continuous variables incorporated in the analysis

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Gender & Age of the household member	4541	-,085	,000
Pair 2	Gender & Social_Integration	4541	-,053	,000
Pair 3	Gender & proppart	4541	,106	,000
Pair 4	Gender & propfull	4541	,010	,487
Pair 5	Gender & Personal net monthly income in Euros, imputed	4541	-,167	,000

In table 26 the output of the paired sample correlations can be seen between gender and the continuous variables. The relationship between gender and the employment type proportions is only significant for the part-time proportion ($r=0,106$). This is a positive and moderately strong relationship. An increase in the proportion part-time working contacts is related to female egos. The

relationship between social integration and gender is negative, weak and significant (r=-0,053). An increase in gender (being a women) rarely goes together with less social integration. The relationship with income and gender is stronger and significant (r=-0,167). Being a woman is related to a lower income. The relationship between age and gender is weak and significant (r=-0,085). An increase in age is rarely related to being a man.

The relationships between the categorical variable domestic situation and the continuous variables have been tested with a Anova F-test with the following syntax.

SYNTAX

UNIANOVA Age BY woonvorm

  /METHOD=SSTYPE(3)

  /INTERCEPT=INCLUDE

  /CRITERIA=ALPHA(0.05)

  /DESIGN=woonvorm.

UNIANOVA Social_Integration BY woonvorm

  /METHOD=SSTYPE(3)

  /INTERCEPT=INCLUDE

  /CRITERIA=ALPHA(0.05)

  /DESIGN=woonvorm.

UNIANOVA propfull BY woonvorm

  /METHOD=SSTYPE(3)

  /INTERCEPT=INCLUDE

  /CRITERIA=ALPHA(0.05)

  /DESIGN=woonvorm.

UNIANOVA proppart BY woonvorm

  /METHOD=SSTYPE(3)

  /INTERCEPT=INCLUDE

  /CRITERIA=ALPHA(0.05)

/DESIGN=woonvorm.

UNIANOVA nettoink_f BY woonvorm

/METHOD=SSTYPE(3)

/INTERCEPT=INCLUDE

/CRITERIA=ALPHA(0.05)

/DESIGN=woonvorm.

In tables 27 till 30 the output is shown.

Table 27 Anova F-test between domestic situation and proppart

Tests of Between-Subjects Effects

Dependent Variable: proppart

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	13,611ᵃ	3	4,537	50,936	,000
Intercept	227,358	1	227,358	2552,600	,000
woonvorm	13,611	3	4,537	50,936	,000
Error	404,106	4537	,089		
Total	769,684	4541			
Corrected Total	417,717	4540			

a. R Squared = ,033 (Adjusted R Squared = ,032)

Table 28 Anova F-test between domestic situation and propfull

Tests of Between-Subjects Effects

Dependent Variable: propfull

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	19,797ᵃ	3	6,599	60,230	,000
Intercept	315,591	1	315,591	2880,529	,000
woonvorm	19,797	3	6,599	60,230	,000
Error	497,074	4537	,110		
Total	1032,082	4541			
Corrected Total	516,870	4540			

a. R Squared = ,038 (Adjusted R Squared = ,038)

Table 29 Anova F-test between domestic situation and social integration

Tests of Between-Subjects Effects

Dependent Variable: Social_Integration

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	325,277[a]	3	108,426	19,557	,000
Intercept	1025954,144	1	1025954,144	185057,511	,000
woonvorm	325,277	3	108,426	19,557	,000
Error	25153,013	4537	5,544		
Total	1717021,000	4541			
Corrected Total	25478,289	4540			

a. R Squared = ,013 (Adjusted R Squared = ,012)

Table 30 Anova F-test between domestic situation and age

Tests of Between-Subjects Effects

Dependent Variable: Age of the household member

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	276086,076[a]	3	92028,692	329,192	,000
Intercept	6722928,626	1	6722928,626	24048,307	,000
woonvorm	276086,076	3	92028,692	329,192	,000
Error	1268360,694	4537	279,559		
Total	13424845,000	4541			
Corrected Total	1544446,770	4540			

a. R Squared = ,179 (Adjusted R Squared = ,178)

The relationships between the employment type proportions and the domestic situation are both significant and fairly small (Cproppart r=0,033; Cpropfull r=0,038). This indicates that it does not really matter which employment type the contacts have with regard to the domestic situation of the ego, as both employment types have a similar relationship with domestic situation. The relationship between social integration and the domestic situation is also very small and significant (r=0,013). An increase in domestic situation (living with more children) is sometimes related to higher social integration. The relationship between income and domestic situation is very small but significant (r=0,005). Higher income rarely goes together with a higher domestic situation (living with children). The relationship between age and domestic situation is stronger and also significant (=0,179). Being older is related to living with more children

Lastly the relationship between gender and domestic situation is analysed with a crosstab with the following syntax.

SYNTAX

CROSSTABS

  /TABLES=woonvorm BY Gender

  /FORMAT=AVALUE TABLES

  /STATISTICS=CHISQ PHI

  /CELLS=COUNT

  /COUNT ROUND CELL.

Table 31 Symmetric measures of crosstab analysis between gender and domestic situation

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,084	,000
	Cramer's V	,084	,000
N of Valid Cases		4541	

Table 32 Chi square table of the crosstab analysis between gender and domestic situation

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	32,037[a]	3	,000
Likelihood Ratio	32,429	3	,000
Linear-by-Linear Association	1,507	1	,220
N of Valid Cases	4541		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 120,48.

In table 31 can be seen that the Cramer's V is 0,084 and significant. There is a weak relationship between gender and domestic situation.

Table 33 Cross tab of the variables domestic situation and gender

Domestic situation * Gender Crosstabulation

Count

		Gender		
		,00	1,00	Total
Domestic situation	Single	455	679	1134
	(Un)married co-habitation, without child(ren)	812	897	1709
	(Un)married co-habitation, with child(ren)	605	814	1419
	Single, with child(ren)	89	190	279
Total		1961	2580	4541

There is a significant but small relation between gender and the domestic situation (Cramer's V=0,084; p<0,001; Chi square=32,037; p<0,001)

2.2   Regression analysis

The regression analysis has been done by following the analysis plan and has been runned in SPSS with the following syntax.

SYNTAX

REGRESSION

 /MISSING LISTWISE

 /STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE ZPP

 /CRITERIA=PIN(.05) POUT(.10)

 /NOORIGIN

 /DEPENDENT Social_Integration

 /METHOD=ENTER Age Singlewith Cowith Cowithout nettoink_f

 /METHOD=ENTER Age Singlewith Cowith Cowithout nettoink_f Cproppart Cpropfull

 /METHOD=ENTER Age Singlewith Cowith Cowithout nettoink_f Cproppart Cpropfull Gender

 /METHOD=ENTER Age Singlewith Cowith Cowithout nettoink_f Cproppart Cpropfull Gender

  GenXCproppart GenXCpropfull

 /PARTIALPLOT ALL

 /SCATTERPLOT=(*ZRESID ,*ZPRED)

 /RESIDUALS HISTOGRAM(ZRESID)

 /RESIDUALS NORMPROB(ZRESID).

 /SAVE COOK LEVER RESID ZRESID DFBETA DFFIT.

The results of the regression analysis are shown in table 34 and 35.

Table 34 Model summary of the regression analysis

Model Summary[e]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,161[a]	,026	,025	2,33917	,026	24,272	5	4535	,000
2	,163[b]	,026	,025	2,33921	,000	,928	2	4533	,395
3	,167[c]	,028	,026	2,33785	,001	6,257	1	4532	,012
4	,167[d]	,028	,026	2,33815	,000	,424	2	4530	,654

a. Predictors: (Constant), Personal net monthly income in Euros, imputed, Cowithout, Cowith, Age of the household member, Singlewith

b. Predictors: (Constant), Personal net monthly income in Euros, imputed, Cowithout, Cowith, Age of the household member, Singlewith, Cproppart, Cpropfull

c. Predictors: (Constant), Personal net monthly income in Euros, imputed, Cowithout, Cowith, Age of the household member, Singlewith, Cproppart, Cpropfull, Gender

d. Predictors: (Constant), Personal net monthly income in Euros, imputed, Cowithout, Cowith, Age of the household member, Singlewith, Cproppart, Cpropfull, Gender, GenXCpropfull, GenXCproppart

e. Dependent Variable: Social_Integration

Table 35 Coefficients table of the regression analysis

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Correlations Zero-order	Correlations Partial	Correlations Part	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	18,065	,134		134,389	,000					
	Age of the household member	,016	,002	,124	7,607	,000	,125	,112	,111	,814	1,228
	Singlewith	,488	,097	,095	5,006	,000	-,019	,074	,073	,590	1,694
	Cowith	,256	,158	,026	1,614	,107	-,030	,024	,024	,832	1,202
	Cowithout	,591	,090	,121	6,564	,000	,105	,097	,096	,633	1,580
	Personal net monthly income in Euros, imputed	1,813E-5	,000	,019	1,310	,190	,027	,019	,019	,987	1,014
2	(Constant)	18,077	,138		130,927	,000					
	Age of the household member	,016	,002	,123	7,256	,000	,125	,107	,106	,751	1,332
	Singlewith	,480	,098	,094	4,895	,000	-,019	,073	,072	,583	1,716
	Cowith	,245	,159	,025	1,545	,122	-,030	,023	,023	,830	1,205
	Cowithout	,591	,090	,121	6,557	,000	,105	,097	,096	,632	1,582
	Personal net monthly income in Euros, imputed	1,898E-5	,000	,020	1,370	,171	,027	,020	,020	,983	1,017
	Cproppart	,172	,129	,022	1,333	,183	-,005	,020	,020	,785	1,275
	Cpropfull	,097	,120	,014	,811	,417	,038	,012	,012	,734	1,362
3	(Constant)	18,216	,149		122,392	,000					
	Age of the household member	,015	,002	,120	7,089	,000	,125	,105	,104	,748	1,337
	Singlewith	,467	,098	,091	4,758	,000	-,019	,071	,070	,581	1,720
	Cowith	,251	,159	,025	1,582	,114	-,030	,023	,023	,830	1,205
	Cowithout	,577	,090	,118	6,399	,000	,105	,095	,094	,630	1,588
	Personal net monthly income in Euros, imputed	1,344E-5	,000	,014	,958	,338	,027	,014	,014	,959	1,043
	Cproppart	,210	,130	,027	1,613	,107	-,005	,024	,024	,774	1,292
	Cpropfull	,120	,120	,017	,994	,320	,038	,015	,015	,730	1,370
	Gender	-,180	,072	-,038	-2,501	,012	-,053	-,037	-,037	,948	1,055
4	(Constant)	18,220	,149		122,262	,000					
	Age of the household member	,015	,002	,121	7,118	,000	,125	,105	,104	,747	1,339
	Singlewith	,465	,098	,091	4,738	,000	-,019	,070	,069	,581	1,721
	Cowith	,253	,159	,026	1,593	,111	-,030	,024	,023	,829	1,206
	Cowithout	,575	,090	,118	6,367	,000	,105	,094	,093	,629	1,589
	Personal net monthly income in Euros, imputed	1,294E-5	,000	,014	,921	,357	,027	,014	,013	,957	1,045
	Cproppart	,204	,190	,026	1,075	,282	-,005	,016	,016	,363	2,758
	Cpropfull	,218	,167	,031	1,305	,192	,038	,019	,019	,378	2,648
	Gender	-,189	,073	-,040	-2,596	,009	-,053	-,039	-,038	,923	1,083
	GenXCproppart	-,006	,260	-,001	-,021	,983	,000	,000	,000	,348	2,876
	GenXCpropfull	-,193	,232	-,020	-,833	,405	,023	-,012	-,012	,363	2,755

a. Dependent Variable: Social_Integration

In table 35 model 1, the effects of the control variables on social integration have been tested. What first stands out is the really small effect of income (b=1,813E-5; p=0,190). This effect is not significant but indicates when earning €1000 more per month, people will feel 0,0813 more socially integrated, which is still a small effect. The effect of age is, positive and significant (b=0,016; p<0,001), however still small. When being 10 years older, you will feel 0,16 more socially integrated. This means that the older people get, the more socially integrated they will feel. However, this is a small effect and slow process, as it takes 10 years to feel a little bit more socially integrated. On the scale of social integration these effects are neglectable. The effects of the domestic situations are however bigger. Being single with children will lead to feeling b=0,488 more socially integrated, compared to living alone. This effect is significant (p<0,001) and means almost half a point increase of social integration. The effect of living in co-habitation without children is comparable (b=0,591; p<0,001). Living in co-habitation without children will lead to 0,0591 on the 8-28 scale of social integration, compared to living alone. This far, not living alone seems to lead to feeling better socially integrated. This also applies to the last category of domestic situation, however this effect is much smaller and not significant (b=0,256*; p=0,190). The effect implies that living in co-habitation with children will cause a 0,256 increase in feeling socially integrated, compared to living alone. However, this effect is not significant.

In model 2 of table 35, the main effects, centred proportion part-time working contacts and the centred proportion full-time working contacts, are added. Here the first and second hypotheses will be tested: "More part-time working contacts, will lead to feeling better socially integrated, compared to more full-time working contacts." And: "As a person's network is more diverse in terms of employment types, the ego will feel less well socially integrated." Both effects of the proportions part-time and full-time working contacts are positive, but not significant. The effect of the proportion part-time working contacts is the biggest of the two (b=0,172; p=0,183). As the close contact network contains of 5 people and the effect based on 1% increase, an increase of 1 person will mean a 20% increase in proportion. Therefore, the effect implies that a 20% increase in the proportion part-time working contacts above the average will lead to a 3,44 (0,172*20) point increase of social integration, compared to the proportion not working contacts and controlled for age, income and domestic situation. On the 6-28 scale of social integration, this is quite a big effect. However, the effect is not significant. The effect of the proportion full-time working contacts on social integration is also positive (b=0,097; p=0,320). The effect implies that a 20% increase proportion full-time working contacts above the average, leads to an increase of 1,94 (0,097*20) points on the scale of social integration, compared to the proportion not working contacts and controlled for age, income and domestic situation. This too, is quite a big effect but not significant. The effects are in line with the first hypothesis, as more part-time

contacts leads to a bigger increase in feeling socially integrated than the proportion full-time contacts. However, the effects are not significant. The effects are not in line with the second, as an increase in the proportion of both employment types will lead to an increase in feeling socially integrated.

In model 35 of table 3, the effect of gender has been added. The effect of gender on social integration is negative and on the edge of being significant with an alpha of $p=0,01$ and is significant with an alpha of $p=0,05$ ($b=-0,180$; $p=0,012$). Scoring higher on gender, being a female leads to a decrease of 0,180 points on the social integration scale, compared to being male and controlled for age, income and domestic situation. Women feel less socially integrated than men. This is in line with the theory. However, this effect is not significant. The effects of the proportion part-time working contacts and the proportion full-time working contacts have increased (Proportion part-time: $b=0,210$; $p=0,107$; Proportion full-time: $b=0,120$; $p=0,320$). Adding gender to the model had an effect on the employment type proportion variables.

Lastly the complete model is visible in table 35 as model 4. Here the interaction variables between gender and the employment type proportion variables have been added. Here, there will be tested if there is a difference between men and women for the effect of the employment type proportions on social integration. Specifically, the hypothesis: "The diversity of employment types of the network has less effect on the feeling of being socially integrated for women then for men", will be tested. The interaction variables are both negative and not significant. The negative interaction effect between gender and the proportion part-time working contacts (GenXCproppart: $b=-0,006$; $p=0,192$) indicates that the effect of the proportion part-time working contacts on social integration will decrease with 0,006 given that the respondent is a woman and that the proportion part-time working contacts is average (Cproppart=0). On the 8-28 scale of social integration and looking at the effect size of the variable Cproppart this is a very small effect and is not significant. The interaction effect is stronger between gender and the proportion full-time working contacts (GenXCpropfull: $b=-0,193$; $p=0,405$). The effect of the proportion full-time working contacts decreases with 0,193 given that the respondent is a woman and that the proportion full-time working contacts is average (Cpropfull=0). Compared to men, the effect of the proportion full-time working contacts on social integration is much weaker for women. Given the effect size Cpropfull, this is a fairly big effect. The effect will would more than double its size. The effects are in line with the theory. However, this effect is not significant. The effect of gender has increased and is now significant ($b=-0,189$); $p=0,009$). Females score 0,189 lower on social integration than men. This is in line with the theory. The effect sizes of the proportion variables have become bigger and are now more or less equally strong (Cproppart: $b=0,204$;

p=0,282; Cpropfull; *b*=0,218; *p*=0,192). This is not in line with the first and second hypotheses.

Throughout the models, the effects of the control variables have not changed much and are therefore not mentionable.

4.2.2 Model evaluation

In table 34 can be seen that the first model explains 2,5% of the variance within social integration (*r²adjusted*=0,025). This is not much, therefore the first model is not a very good model for predicting social integration. Then, the second model has the same amount of explained variance as the first model. Model 2 also explains 2,5% of the variance within social integration (*r²adjusted*=0,025). Model 2 does not add any explained variance. This can also be seen in the F change which is very small and not significant (*Fchange*=0,928; *p*=0,395). Model 2 is thus not a very good model for predicting social integration.

In model 3 the proportion explained variance has gone up compared to the second model, but not significantly (*r²adjusted*=0,026; *Fchange*=6,257; *p*=0,012). Still, the explained variance is very small. Model 3 also does not a very good job in predicting social integration, but there seems to be a small improvement in comparison to model 2, even though it is not significant. Lastly, compared to model 3, model 4 did not add any explained variance as it stayed the same. There was also no significant F change (*r²adjusted*=0,026; *Fchange*=0,424; *p*=0,654). Therefore, model 3 can be seen as the best model for predicting social integration.

Appendix 3
Assumptions, outliers and multicollinearity

The assumptions of linear regression have been tested with the following syntax.

SYNTAX

REGRESSION

 /MISSING LISTWISE

 /STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE ZPP

 /CRITERIA=PIN(.05) POUT(.10)

 /NOORIGIN

 /DEPENDENT Social_Integration

 /METHOD=ENTER Age Singlewith Cowith Cowithout nettoink_f

 /METHOD=ENTER Age Singlewith Cowith Cowithout nettoink_f Cproppart Cpropfull

 /METHOD=ENTER Age Singlewith Cowith Cowithout nettoink_f Cproppart Cpropfull Gender

 /METHOD=ENTER Age Singlewith Cowith Cowithout nettoink_f Cproppart Cpropfull Gender

  GenXCproppart GenXCpropfull

 /PARTIALPLOT ALL

 /SCATTERPLOT=(*ZRESID ,*ZPRED)

 /RESIDUALS HISTOGRAM(ZRESID)

 /RESIDUALS NORMPROB(ZRESID).

 /SAVE COOK LEVER RESID ZRESID DFBETA DFFIT.

To evaluate the quality of this analysis, the assumptions of linear regression will be tested. The syntax for this that can be seen above. First off all, the assumption of independency will be discussed. This assumption assumes that the observations are done independently. This means that the answers of one respondent does not give information about another respondent. The independency assumption can not be tested with hundred precent certainty, as the data used for this study is secondary. However, it is known for this data set that the LISS data panel works with households and holds respondents within these households. Therefore, it can occur that more people of one household are selected for a sample. In this study this can mean that the close contacts networks can have some overlap. It is quite plausible that one or more close contacts are people within the household as people tend to have intensive contact with each other and close relationships. However, the samples are randomized so this compensates. Though, as said, it can not be known for certain if this is also the case for this sample. So, this assumption can not be violated completely. But since

this study investigates ego networks, this overlap will not cause any problems as the effects on individuals are studied and not the effects on the ties in the networks.

The next assumptions that will be discussed are linearity and homoscedasticity. The linearity assumption tests if the dependent and independent variables have a linear relationship with each other. The homoscedasticity tests if the dispersion of the dependent variable is the same for all the data. In figure 23 a scatterplot of the standardized residuals of the complete model (model 4) of the regression analysis is shown. The assumption of homoscedasticity does not seem violated. No pattern can be seen in the dispersion of the residuals. The dispersion is quite even. Looking at figure 23, it also looks like that the average of the residuals is zero. The linearity assumption is therefore not violated.

The last assumption that will be discussed is the normality assumption. This assumptions tests if the dependent variable is normally distributed. For this assumption a histogram of the standardized residuals of the dependent variable social integration and a PP-plot have been made. These can be seen in figure 24 and figure 25. In figure 24 can be seen that the variable social integration is left skewed. Figure 25 also show that the points do not fit the line well. A little reversed S pattern can be seen. This indicates that the dependent variable is not distributed normally, the assumptions is therefore violated. This can affect the analysis and tests negatively. The power of the tests is weaker, even though the big sample size, and generalisations about the populations need to be made more carefully.

Interventions for this violation have been investigated. However, there was no clear connection between a certain variable and social integration that could explain the skewness. The variable is probably skewed from itself. Which is not odd as people can be quite extreme in answering questions like this. The only intervention that will be taken is that the tests will be tested more strictly. Therefore, an alpha of 0,01 will be used to determine whether an effect is significant or not.
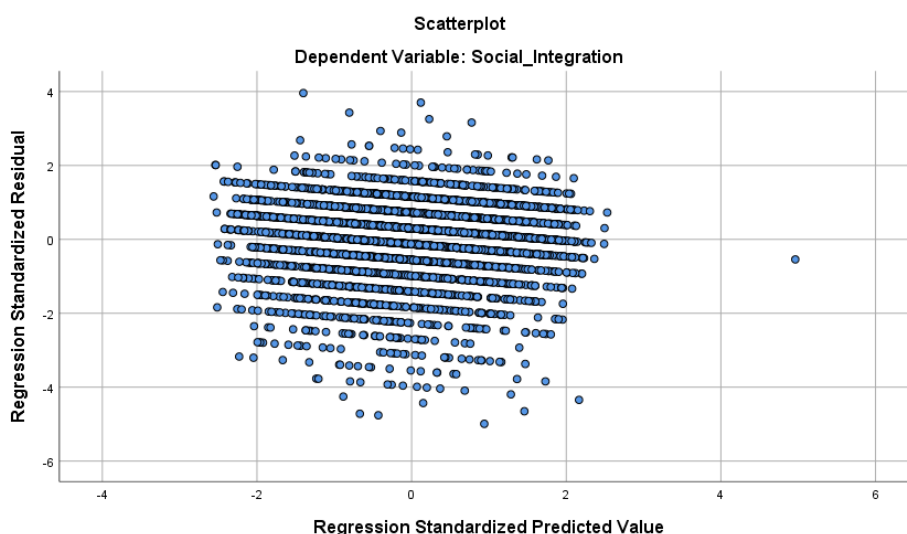

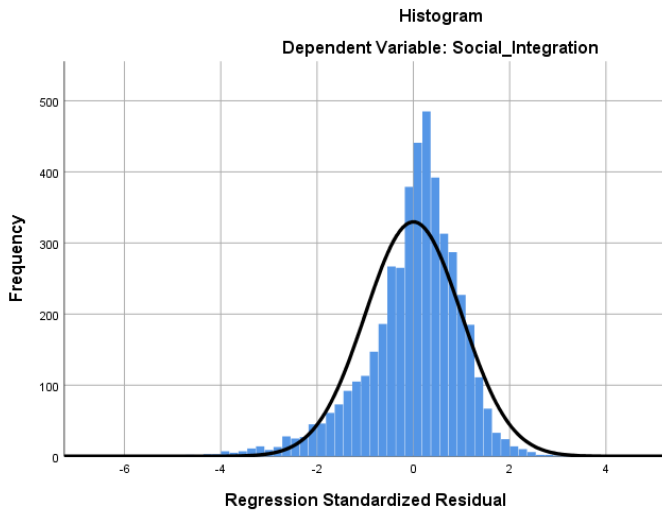*Figure 24 Residual plot with as dependent variable Social integration*

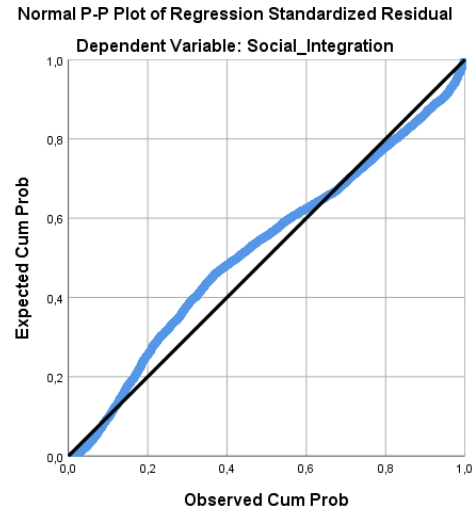*Figure 24 Histogram of the standardized residuals of the dependent variable social integration*

Figure 25 PP-plot of the standardized residuals

3.2 Multicollinearity

Multicollinearity occurs when a variable has a strong connection with another variable. This indicates that the variable explains the same part of variance of the independent variable as another predictor. To examine the interconnections between the independent variables, there will be looked VIF values of the independent variables. These values can be seen in table 10. in appendix 2 and in the results chapter table 3. The general limit value of 4 will be applied. VIF values above 4 will indicate multicollinearity. In table 12 in appendix 2 can be seen that no variable exceeds the limit value of the VIF. However, the variables of the proportion part-time, proportion full time and the interaction variables are higher than the others. This is however logical as these variables have some of the same components. No problems of multicollinearity seem to appear.

3.3 Outliers

To investigate whether there are some cases that influence the results very much and if there are outliers the syntax shown below has been runned. An analysis of the Cook's distances, leverage DFFI and the standardized residuals has been done. In table 12. the cases with the biggest cook's distance, leverage and DFFIT are shown.

SYNTAX

REGRESSION

 /MISSING LISTWISE

 /STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE ZPP

```
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT Social_Integration
/METHOD=ENTER Age Singlewith Cowith Cowithout nettoink_f
/METHOD=ENTER Age Singlewith Cowith Cowithout nettoink_f Cproppart Cpropfull
/METHOD=ENTER Age Singlewith Cowith Cowithout nettoink_f Cproppart Cpropfull Gender
/METHOD=ENTER Age Singlewith Cowith Cowithout nettoink_f Cproppart Cpropfull Gender
  GenXCproppart GenXCpropfull
/PARTIALPLOT ALL
/SCATTERPLOT=(*ZRESID ,*ZPRED)
/RESIDUALS HISTOGRAM(ZRESID)
/RESIDUALS NORMPROB(ZRESID).
/SAVE COOK LEVER RESID ZRESID DFBETA DFFIT.
```

First the tolerance values of the Cook's distance, leverage and DFFIT must be calculated. The maximum tolerance value of the Cook's distance is 0,00088 (4/4544). When a case exceeds this value, there might be a problem with this case. When the Cook's distance value is bigger than 1 there really is a problem. The maximum tolerance value of the leverage is 0,00726 ((3x (11/4544)). The maximum tolerance value for DFFIT is 0,098402 ((2x (11/4544). There are 235 values that exceed the tolerance value of the Cook's distance, 10 values that exceed the leverage tolerance and 1 that exceeds the DFFIT tolerance. The biggest values are shown in table 36. The range of the standardized residuals is 3,94475 till 4,96912. When using a tolerance of -3 and 3, 65 values exceed this. The extremes below -3 mostly scored 12 or lower on social integration. The 3 plus values scored 27 or higher on social integration. Thus, the residuals perhaps can be explained by extreme levels of social integration. Because not all the extreme residuals have around the same values for a certain variable, no real explanation and connection can be found for the residuals. It just shows that answers on questions regarding social integration are very widely answered by people in the population. Deleting this case will cause a not poorer representation of the population. Therefore, these cases will not be deleted.

In table 36 can be seen that the case numbers with the highest Cook's distance is 854346 with a Cook's distance of 0,11857, then case 873681 with a value of 0,02942 and case 899908 with 0,00637. Case 854346 has the second highest leverage value, 0,31814. The DFFIT of this case is -1,51063, this is quite big and shows that case 854346 is an influential point. There is probably a problem in the x and y direction. However, the residual is within the tolerance range of -3 and 3. The variable is an influential point but not an outlier

Case 821415 has the highest leverage value, 0,31856, however this case has a very small Cook's distance (0,00056). The DFFIT is big -1,0434, but the residual is within the tolerance range. Case 821415 is probably no outlier but is influential.

Case 873681 has both a high Cook's distance (0,02942) and leverage (0,26454). Just as case 854346. The DFFIT is 0,68624, so the case is quite influential. A predicted value would change with 0,68624 if this case would not be incorporated in the analysis. On the 8-28 scale of the social integration variable, this is not very big. This point will not be a very big problem.

The third highest Cook's distance is from case 899908 (0,00637). The leverage and DFFIT however, are not very big. The residual exceeds the tolerance range (-4,87306). This case is an outlier but is not very influential.

Table 36 Outlier statistics of mentionable cases

Case	Cook's Distance	Leverage	DFFIT	Standardized residual
854346	0,11857	0,31814	-1,51063	-1,37971
873681	0,02942	0,26454	0,68624	0,81292
899908	0,00637	0,00271	-0,03362	-4,87306
859329	0,00576	0,00374	-0,3714	-3,98597
821415	0,00056	0,31856	-1,0434	-0,09511

There has been looked at the incomes of these cases as income was a very skew distributed variable and the small effects of income in the regression. For this a boxplot has been made. This can be seen in figure 26. The syntax can be found below

SYNTAX

```
EXAMINE VARIABLES=nettoink_f
 /ID=nomem_encr
 /PLOT BOXPLOT STEMLEAF
 /COMPARE GROUPS
 /STATISTICS DESCRIPTIVES
 /CINTERVAL 95
 /MISSING LISTWISE
 /NOTOTAL.
```

In figure 26 can be seen that case number 821415 has the highest income of the data with €145.666 per month. But since this case is not really in influential point, based on the arguments and analysis earlier, this case will not be deleted. Case 873681 has the third highest income with €133.537 per month. Case number 899908 has an income of €0 per month.

Case numbers 854346 and 873681 are influential points and case 899908 an outlier. These cases are deleted with the following syntax.

SYNTAX

USE ALL.

COMPUTE filter_$=(nomem_encr ~= 854346 & nomem_encr ~= 873681 & nomem_encr ~= 899908).

VARIABLE LABELS filter_$ 'nomem_encr ~= 854346 & nomem_encr ~= 873681 & nomem_encr ~= 899908 '+

'(FILTER)'.

VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.

FORMATS filter_$ (f1.0).

FILTER BY filter_$.

EXECUTE.

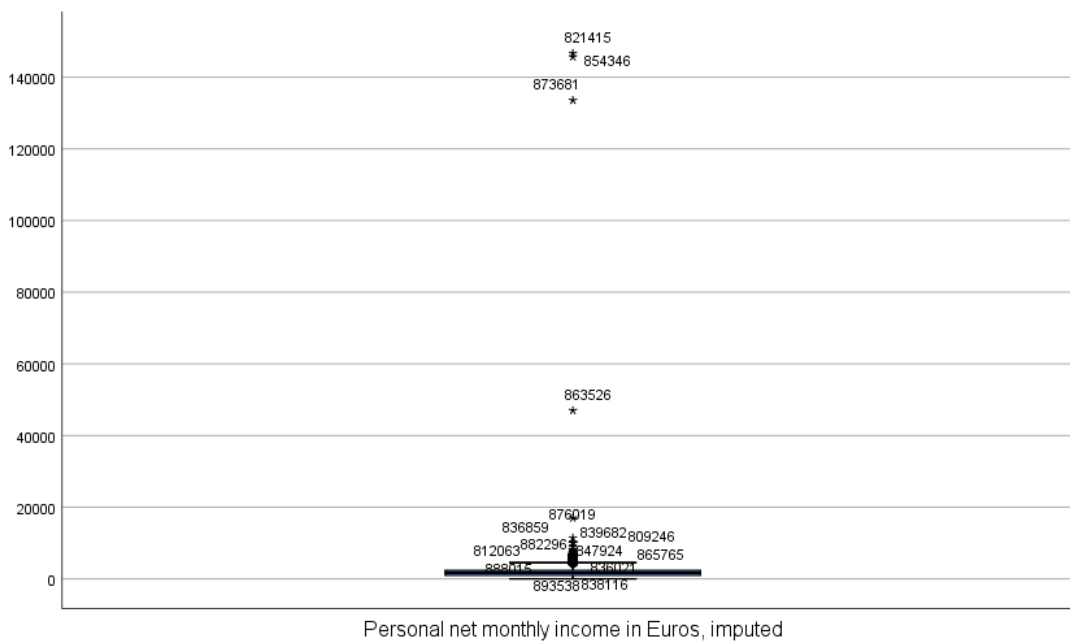After selecting the cases the case numbers were deleted with the option clear in SPSS.

Figure 26 Boxplot of netto_inkf labelled by case numbers