Lubbers, Miranda J.; Molina, José Luis; Valenzuela, Hugo. «When networks speak volumes: Variation in the size of broader acquaintanceship networks». Social networks, Vol. 59 (2019), p. 55-69. DOI 10.1016/j.socnet.2018.08.004

When Networks Speak Volumes:

Variation in the Size of Broader Acquaintanceship Networks

Authors: Miranda J. Lubbers, José Luis Molina, Hugo Valenzuela

Journal: Social Networks

Volume: 59

Page numbers: 55-69

DOI: https://doi.org/10.1016/j.socnet.2018.08.004

Date of publication: January 2019

Original publisher: Elsevier B.V.

Acknowledgements:

We are grateful to the Spanish Center for Sociological Research (Centro de Investigaciones Sociológicas -CIS-, Madrid) for conducting the National Barometer Survey Number 3036 “Social cohesion and trust”for which we have designed the Network Scale-Up module. The module was selected in a competitive call for proposals. Moreover, we thank Félix Requena-Santos, Berta Álvarez-Miranda, Carmen Lence, and Araceli Mateos-Díaz for their valuable help in designing the module. We are also grateful to Devon Brewer for his advice regarding the NSUM instrument for Spain. The first author was supported by a grant of the Spanish Ministry for Science and Innovation (RYC-2010-06081).

Corresponding author:

Miranda J. Lubbers, Department of Social and Cultural Anthropology, Edifici B - Facultad de Filosofía i Lletres, Campus de la Universitat Autònoma de Barcelona, 08193 Bellaterra (Cerdanyola del Vallès), Barcelona, Spain. E-mail: mirandajessica.lubbers@uab.cat

Abstract:

Personal network researchers have extensively studied the characteristics and effects of individuals' closest relationships, but they have paid much less attention to broader acquaintanceship networks, despite evidence that weak ties can also provide social support. In this paper we focus on one aspect of these networks: acquaintanceship volume. We estimate its distributional parameters for a large, representative sample of the general population of Spain, explore its variation across social groups as well as its implications for social support availability. We designed a survey instrument based on the Network Scale-Up Method and implemented it in a national survey in Spain. Our results suggest that Spaniards have approximately 530 acquaintances, with a large inter-individual variation, comparable to the estimates reported for the American population. Acquaintanceship volume vary with gender, age, education, and income. These differences are partially related to the unequal participation of social groups in voluntary associations, confirming the civic value of such associations, and in employment. Even with similar core network size, acquaintanceship volume increases the likelihood of having adequate social support available, suggesting that broader acquaintanceship networks also structure individual outcomes.

Keywords: Acquaintanceship; Personal Networks; Network Size; Weak Ties; Social Support

1. Introduction

Personal network research is based on the premise that individuals are not only influenced by their own attributes and macro-level characteristics, but also by their relationships with others. Traditionally, the lion share of this research has focused on the most intimate layers of the networks, primarily consisting of the romantic partner, family and close friends, which have been assumed to be most consequential for individual health and well-being (e.g., Berkman and Glass 2000; Kahn and Antonucci 1980). On a daily basis however, people spend far more time with a range of non-intimate others (e.g., colleagues, neighbors, parents of the other children at the school of yours, club mates, acquaintances; Kahneman et al. 2004; Jacobs and Gerson 2004), who are more numerous and less densely connected among each other. Some scholars have suggested that non-intimate ties are as important for health and well-being as intimate ties, serving both distinct and similar functions (cf. Fingerman 2009, for a review). On the one hand, non-intimate others (or weak ties) are better able than intimate others (strong ties) to provide access to novel information (Granovetter, 1973). On the other, weak ties can also provide support for which strong ties are thought to be better suited, for example during emergencies (Fingerman, 2009), when close ties are not around (Bojarczuk and Mühlau 2018; Desmond 2013; Small and Sukhu 2016), or when individuals do not expect their significant others to have cognitive empathy with the issue that bothers them (Small 2017).

Yet despite the functions that weak ties are supposed to have, empirical evidence about these broader acquaintanceship networks is scarce, presumably due to the difficulty of investigating extensive sets of social relationships. A first, straightforward question to ask about these larger networks is: how large are they? Dunbar (e.g., Dunbar 1993; Kudo and Dunbar 2001) argued that the upper bound of the number of people humans can know as individuals and with whom they can maintain meaningful contact is determined by their long-term memory capacity. This is now known as the “social brain hypothesis”. By regressing the average group size of hominoids on their neocortex ratio and by using the regression line to extrapolate the findings to humans, he predicted that the average group size of humans would be close to 150, now dubbed “Dunbar’s number”. Indeed, he observed that many tribal and traditional communities have approximately this size, and detected 500 as another typical group size (e.g., Dunbar and Sosis 2017). He suggested that these sizes also apply to the outer layers (i.e., weaker ties) of personal networks, respectively the “active network” and the “acquaintances layer” (Curry & Dunbar, 2013).

A wide variety of methods have been applied to estimate acquaintanceship volume empirically in modern societies, including contact diaries (Dávid, Huszti, Barna, & Fu, 2016; De Sola Pool & Kochen, 1978; Fu, 2005, 2007; Gurevitch, 1961; Lonkila, 1997; Pachur, Schooler, & Stevens, 2014), participant observation (Boissevain, 1974), experiments (Bernard & Shelley, 1987; Freeman & Thompson, 1989; Killworth & Bernard, 1978; Killworth, Johnsen, Bernard, Ann Shelley, & McCarty, 1990), enquiries about the number of sent Christmas cards (Hill and Dunbar 2003), free lists of all related and unrelated network members (Lu, Roberts, Lio, Dunbar, & Crowcroft, 2009; Roberts, Dunbar, Pollet, & Kuppens, 2009), surveys (DiPrete, Gelman, McCormick, Teitler, & Zheng, 2011; McCarty, Killworth, Bernard, & Johnsen, 2001; Shati, Haghdoost, Majdzadeh, Mohammad, & Mortazavi, 2014; Shokoohi, Baneshi, & Haghdoost, 2010; Van Tubergen, Ali Al-Modaf, Almosaed, & Said Al-Ghamdi, 2016), and social media data (Ellison, Steinfield, and Lampe 2007; Arnaboldi, Guazzini, and Passarella 2013; Dunbar et al. 2015), often combined with some form of extrapolation to estimate the total set of contacts. These studies led to vastly different estimates of average network size (see Table 1): from less than 100 (free recall; contact diaries for a limited period; online social networks) up to thousands (extrapolation from telephone book experiments or from prolonged contact diaries, participant observation), depending among others on the method of estimation, the characteristics of the sample, and the underlying definition of "knowing someone" (the network boundary). However, many studies found much higher averages than 150, concluding that Dunbar's number is on the low side for modern societies (e.g., Wellman 2012). A mechanism that may explain the higher numbers given limited cognitive capacity is the use of compression heuristics among humans (Brashears, 2013), allowing the storage of larger amounts of information about social relationships in the brain.

--PLEASE INSERT TABLE 1 ABOUT HERE--

Although estimates of average acquaintanceship volume vary widely across studies, researchers coincide in observing large inter-individual variation (see Table 1). This raises other questions, concerning its potential causes and consequences. Scholars have mostly explored biological (e.g., gray matter volume; Bickart et al. 2011; Brashears, Hoagland, and Quintane 2016; Lewis et al. 2011) and psychological explanations for this variation (e.g., perspective-taking, or extraversion; Pollet, Roberts, and Dunbar 2011; Roberts et al. 2008; Lu et al. 2009; Stiller and Dunbar 2007). Another factor that may constrain the size of networks at different levels of tie strength is time (Roberts et al. 2009).

Yet so far, little research has explored whether acquaintanceship volume differs among social groups. In contrast, for core networks, a large body of evidence shows inequalities in the size and other characteristics of core networks in terms of gender, race, age, education and income (e.g., Fischer 1982; Marsden 1987). Among others, core networks have been found to be larger among the higher educated (Fischer, 1982; Grossetti, 2007; Marsden, 1987; Moore, 1990) and the higher income groups (e.g., Van den Berg and Timmermans 2015). Furthermore, they seem to shrink with age (Marsden 1987; Smith et al. 2015; Cornwell, Laumann, and Schumm 2011; Harling et al. 2018). Consequently, even though core networks are assumed to function as safety nets on the micro-level in terms of the social support they offer, on a macro-level they can reproduce or exacerbate inequalities (cf. DiMaggio and Garip 2012), in the sense that people who may need more support from their relationships have smaller networks. A sociological explanation for the differences between social groups in the size and other characteristics of core networks lies in the unequal access of these groups (Chua, 2013) to the social contexts in which ties are formed (e.g., Blau 1977; Feld 1981; Grossetti 2005; Small 2009). Feld argued that friendship ties in modern life are organized around “social foci”, defined as “social, psychological, legal or physical [entities] around which joint activities are organized (e.g., workplaces, voluntary organizations, hangouts, families, etc.)” (Feld 1981, p. 1016). Consequently, these foci provide opportunities for interaction to people, such that people associated with the same focus are more likely to develop a relationship. This theory has been cited extensively in research into core networks (e.g., Mollenhorst, Völker, and Flap 2011; Marsden 1987; Small and Sukhu 2016; Smith, McPherson, and Smith-Lovin 2014), to explain why strong relationships are often homophilous, and also why persons in poverty or the elderly have smaller networks (e.g., Van Eijk 2010).

Yet it is equally likely that the unequal access of social groups to social foci (Chua, 2013) affects the broader acquaintanceship networks. Weak ties, too, are created in social contexts, be it neighborhoods, schools, work places, churches, or other contexts where individuals interact. Do the differences in network size observed for core ties also apply to the larger acquaintanceship networks? The scarce empirical evidence for inequalities in broader network size (Dávid et al., 2016; DiPrete et al., 2011; Roberts et al., 2009; Shati et al., 2014; Shokoohi et al., 2010; Van Tubergen et al., 2016) is mixed for gender, age, and education (Van Tubergen et al. 2016; DiPrete et al. 2011; Shati et al. 2014;

Shokoohi, Baneshi, and Haghdoost 2010; Dávid et al. 2016; Roberts et al. 2009), but consistent for income (for only two studies; DiPrete et al. 2011; Van Tubergen et al. 2016), in the sense that people with higher incomes tend to have larger acquaintanceship networks. With regard to social foci, the scarce empirical evidence suggests that religious service attendance (DiPrete et al. 2011; in the US) and being employed (Van Tubergen et al. 2016; among youths aged 18-25) increase network extensity, suggesting that participation in religious communities and work environments has beneficial effects on acquaintanceship volume. Living in a couple also increases network extensity for young people in Saudi Arabia in the context of family relationships (Van Tubergen et al. 2016), but decreases it among Kermanian men in Iran (Shokoohi et al., 2010), and was not related to network size in the US (DiPrete et al. 2011). Except for the study of DiPrete and colleagues, none of these studies is based on representative samples of the national population though (see Table 1). Also, none of these studies relates acquaintanceship volume with support availability. However, if weak ties provide individuals with information, and complement strong ties in the provision of material and emotional support, does having more of them, beyond core network ties, affect the adequacy of social support?

Considering the voids in the literature about acquaintanceship networks, with this paper we first aim to estimate the distribution of acquaintanceship volume for the general population of Spain, on the basis of a large, nationally representative sample and following the recommendations with regard to instrument design by McCormick, Salganik, and Zheng (2010). Second, we aim to explore variation in acquaintanceship volume across gender, age, education, and income. Third, we explore whether these differences are related to differential participation in social foci: civil status, as a proximal variable for participation in family networks; having children of minor age, as an indicator of participation in parental and school networks; employment status, for participation in work environments or studies; religious service attendance, for participation in religious communities, and active membership in associations. Last, we tested the association between acquaintanceship volume and the availability of adequate social support (Fischer, 1982) for four dimensions, namely help to find a job, lending money, practical help during illness, and emotional support. For these analyses, we control for individual attributes and core network size (the number of friends and relatives), to test whether the acquaintanceship volume *beyond* core ties is associated with social support availability. We hypothesized that due to its reliance on novel and specified information, help with finding a job would be particularly related to acquaintanceship volume (more than with the number of friends or relatives), whereas financial and emotional support and support during illness relies more on relatives and friends, with a complementary task of the broader acquaintanceship network.

The following section describes the method we use for the estimation of acquaintanceship volume in more detail. Subsequently, we present the design of our instrument for the population of Spain (in Section 3) and the characteristics of the national survey (N~2,500) in which it was administered by the Spanish Center for Sociological Research (CIS). Section 4 describes the results, and Section 5 our conclusions.

2. The known population method

As indicated in Section 1, the literature about acquaintanceship volume is predominantly based on relatively small convenience samples or on specific samples (see Table 1), as nearly all methods involve a heavy respondent burden and are therefore difficult to integrate in large-scale surveys. Consequently, we still know relatively little about the distributional parameters, predictors and effects of acquaintanceship volume for larger populations. Data about social media networks (Hofstra, Corten, Van Tubergen, & Ellison, 2017) are a notable exception in terms of respondent burden, but the extent to which they accurately represent acquaintanceship networks depends among others on the time informants and their network members spend on social media and the selectivity they employ in befriending and following others (REFERENCE OMITTED FOR PEER REVIEW). Furthermore, social media networks do not provide all the individual attributes that researchers may be interested in for explaining network size. Therefore, for the time being, such data complement, rather than substitute other types of data on acquaintanceship networks.

So far, two methods have been developed for estimating acquaintanceship volume in large-scale surveys, the summation method (cf. Bernard et al. 2010) and the known population method (Bernard et al. 1989, 1991, 2010; Killworth et al. 1998, 2003, 2006; McCarty et al. 2001). The summation method asks respondents to estimate the number of people they know in various categories (e.g., relatives, friends, neighbors etc.) and then sums the estimates over all categories (e.g., Van Tubergen 2016). The known population method, which is used in this paper, involves asking respondents how many people they know in various rare subpopulations of known size. Under the assumption that on average, acquaintanceship networks are representative, scaled-down versions of the society in which they are embedded (the assumption of *random mixing*), personal network size can be estimated. The assumption of random mixing can be expressed as follows:

$$\frac{y_{ik}}{d_i} = \frac{N_k}{N}$$

where y_{ik} represents the number of people that population member i knows in a subpopulation k, d_i the total number of people that individual i knows ("*d*" for degree), N_k the total size of the subpopulation k and N the total size of the population. In other words, the fraction that a subpopulation represents in the total population is assumed to be the same in the personal networks of the individuals who compose the population. Under this assumption, we can estimate acquaintanceship volume (or *degree*) in surveys by asking respondents "How many people do you know [in subpopulation *k*]?", where "knowing" should be defined a priori, and by complementing these data with official statistics of the size of the total population N and of known subpopulation N_k. The inclusion of questions about multiple subpopulations of known size $K=\{1, ..., k\}$ makes the estimate more accurate. Thus, the number of people an individual i knows in a subpopulation k follows a binomial distribution with parameters d_i and N_k/N, and degree can be estimated as:

$$\hat{d}_i = \frac{\sum_{k=1}^{K} y_{ik}}{\sum_{k=1}^{K} N_k} \cdot N$$

and its approximate standard error as:

$$SE(\hat{d}_i) \approx \sqrt{\hat{d}_i} \cdot \sqrt{\frac{1 - \sum_{k=1}^{K} N_k / N}{\sum_{k=1}^{K} N_k / N}}$$

(Killworth et al., 1998; McCormick, Salganik, & Zheng, 2010). The known population method is a byproduct of the Network Scale-Up Method (NSUM), originally designed by Bernard, Killworth, and McCarty (Bernard et al. 1989, 1991, 2010; Killworth et al. 1998, 2003, 2006; McCarty et al. 2001) to estimate the size of hard-to-count populations, i.e., social groups in the general population for which there are no reliable official estimates of their size.

Both Killworth et al. and McCormick et al. stressed that the validity of the method depends heavily on the validity of the assumptions on which the method is based. First, it is assumed that respondents are fully informed about who in their networks is part of the subpopulation, so respondents should be asked about easily observable features. If this is not the case, or if people try to hide the characteristic, "transmission error" can occur, causing an underestimation of network size. Second, the random mixing assumption implies that all people are equally likely to know someone from a subpopulation. If this is not likely for the given subpopulation, “barrier effects” occur, which is when some social groups know fewer people from the subpopulation than others. Third, it is assumed that respondents can provide the number of acquaintances in the subpopulation accurately, that is, they do not have “recall bias”. Recall bias occurs when an individual knows more people from a given subpopulation than he or she remembers.

Taking these possible errors into account, McCormick et al. (2010) recommended the use of first names (“How many people do you know [called X]?”) for several reasons. National statistical institutes often keep updated onomastics data that allow researchers to know their prevalence in the population. Names are characteristics people tend to know about their contacts, even about their weak ties (thus avoiding transmission errors), and for most names the assumption of random mixing is reasonable. However, it should be noted that the use of names may be problematic for ethnic minority groups, when mixing between ethnic groups is not random and naming patterns diverge. To further minimize barrier effects, McCormick et al. recommended to take into account the gender associated with the names and their distribution over birth decades, so that the socio-demographic profile of the combined set of names used for estimation represents at a small scale the society (e.g., if 7% of the people in a society are women between the ages of 21 and 40, it is recommended that 7% of the population represented by the set of names are also women of these age categories).

Furthermore, to minimize recall bias, the authors recommended the use of relatively rare names (representing 0.1-0.2% of the total population). It is easier for respondents to remember the number of people they know who have a rare name than it is to remember the number of people with a popular name.

Finally, novel extensions of the initial estimation method have been developed (DiPrete et al., 2011; Feehan & Salganik, 2016; Maltiel, Raftery, McCormick, & Baraff, 2015; McCormick et al., 2010; Zheng et al., 2006b). Maltiel et al. (2015) developed Bayesian methods using MCMC algorithms to estimate the size of unknown populations, implemented in the R library NSUM. These methods also capture more accurately the uncertainty in network size, by controlling for possible biases. First, they introduced a random effect for the size of the networks d_i, thus incorporating the uncertainty in the sizes. d_i thus follows a log normal distribution with an average μ and a variance σ^2. This is the “random degree model”. Second, while the model of Killworth et al. is based on the assumption that the probability of knowing someone of a certain subpopulation is constant (N_k/N), barrier effects are very common. Therefore, Maltiel and colleagues proposed non-random mixing models, where the overdispersion in individual tendencies to form links with subpopulations is explicitly modeled, above the variation in expected responses based on the size of their networks and the size of subpopulations. Therefore, they allowed that the probability of knowing someone varies interindividually following a Beta distribution with an average m_k and a measure for over-dispersion, ρ_k. Indeed, their simulations showed that the barrier model worked better than the random-degree model in most cases. Finally, they proposed taking into account transmission errors in the case of stigmatized or hidden characteristics (which we will not need in this paper). Last, they proposed to adjust for recall bias based on the relation between the total number of people remembered in the known populations and the size of the known populations. Previous empirical studies have found that people overestimate the number of people they know in very small populations and underestimate the number of people they know in large populations. The models are estimated using Markov Chain Monte Carlo (MCMC) (see Maltiel et al. 2015 for further details).

The known population method has been applied and validated in a large number of investigations, especially in the field of health (see for overviews Bernard et al. 2010, and Feehan & Salganik 2016). It has also been included as a special module in the General Social Survey of the United States in 2006 (DiPrete et al., 2011). However, many studies that apply the method do not focus on describing social structure per se, but on estimating the size of hard-to-count populations. Exceptions are the studies of DiPrete et al. (2011) and Zheng, Salganik, and Gelman (2006), for the US, although earlier studies may be more limited by the validity of the assumptions.

To our knowledge, the method has been applied only once in Spain, but the NSUM results of this study have not yet been published. Spain is an excellent context for using the method, since it has publicly available onomastics data that also provide the profiles associated with each name in terms of gender, birth decade, province of birth, and province of residence, as well as the frequency of compound names (Instituto Nacional de Estadística 2016). In the next section, we will explain the instrument design.

3. Methods

Sample

For the study, we have designed a special module for the National Barometer in Spain, a national survey that is regularly conducted by the Spanish Center for Sociological

Research (*Centro de Investigaciones Sociológicas, CIS*)[1]. This study was administered to a multi-stage stratified sample of the adult population with Spanish nationality in Spain. The decision to focus exclusively on the population with Spanish nationality was based on the problems we foresaw for estimating acquaintanceship volume among people of other nationalities, with foreign names. If, as prior research has suggested, a large share of the network members of immigrants is foreign-born (Bolíbar, Martí, & Verd, 2015; De Miguel Luken & Tranmer, 2010; Lubbers, Molina, & McCarty, 2007), with a variety of foreign names, respondents' network size would be underestimated.

Primary sampling units -municipalities-, and secondary units -sections-, have been randomly selected proportionally, and the last units, individuals, have been selected by random routes and gender and age quotas. The strata were formed by crossing the 17 autonomous communities with the size of the town, divided into 7 categories.

The questionnaire was administered through personal computer-assisted interviews (CAPI) in the respondents' homes between December 11, 2014, and January 20, 2015. The sample size is 2,468. For this paper, we excluded respondents who were little or not at all sincere in their answers according to the interviewers (*N*=53), who had missing values on the NSUM instrument (an additional *N*=26), or inconsistent response patterns[2] (an additional 113 cases). The effective sample size is therefore 2,276. 48.5% of the respondents were men and 51.5% women, and age ranged between 18 and 93 years (*M*=47.8, *SD*=17.3). For the analysis of social support provision, there are a few more missing cases, lowering the effect sample size (see Tables 5 and 6).

Measures

The NSUM instrument. For the NSUM instrument, we used a series of questions about names as explained above. The general question and definition of "knowing" have been formulated as follows:

"*(...) First we will ask you how many people you know with certain names to help us estimate the number of people you know. By "knowing someone" we mean that you know this person by name and you would stop and talk to this person if you´d see him or her on the street, in a shop, or wherever. This includes PROXIMAL relations such as your partner, your relatives, friends, neighbors and work or study mates but also people YOU DON´T KNOW SO WELL. These people can live close to you or in other cities or countries. How many people of 15 years or older do you know who(se name is) …?*"

Responses could be given as numbers (0, 1, 2...) but responses larger than 10 were coded as “11 or more”. We decided to combine responses greater than 10 as recall problems may be larger for respondents who know many people in a subpopulation (DiPrete et al. 2011). This decision has affected only 1 in every 408 responses (0.2%).

To select names, we used the onomastics data from the Continuous Register Statistics on the 1st of January, 2013 (National Statistical Institute, 2014), and an initial list provided by Devon Brewer, used for another study in Spain (see Brewer, 2016, for information on the study in general), which consisted of 12 names. Of these two bases, names were selected that had a prevalence of 0.1-0.2% of the total population (men and women), as recommended by McCormick, Salganik, and Zheng (2010). Consequently, we calculated the distribution by gender, birth decade, and autonomous community, both for the total population and for each of the names. To achieve gender balance, a similar number of female and male names was selected. In addition, the set of female and male names (separately) were chosen in such a way that their distribution across birth decades coincided with that of the total population. Half of the names we selected also appeared in Brewer’s set.

Several other considerations concerning the selection of names deserve mention. First, compound names (e.g. “Maria Antonia”) are common in Spain, although in everyday life people use only one (e.g., “Antonia”). When asked “How many people do you know whose name is Antonia?”, respondents are likely to report about both Antonia and María Antonia. We took this into account by asking respondents about the single *and* the compound name (“How many people do you know whose name is Antonia or María Antonia?”) in these cases, and we collected in the onomastics database the prevalence of both the singular name, “Tomás”/“Antonia”, and the compound name, “José Tomás”/“María Antonia”. Second, we have selected names that were not particular for a specific geographic region. Nevertheless, there are variations across autonomous communities in the use of names, especially in Catalonia and the Basque Country. In some cases, Catalan or Basque variations were explicitly included in the questionnaire (e.g. “Ricardo or Ricard”; “Gonzalo or Gonçal”). Again, in these cases we aggregated the data about the prevalence of both names. Last, we generally avoided names that are associated with shorter calling names for informal use (nicknames) that suffer phonetic alterations (e.g., Francisco - Paco-, or Dolores - Lola-).

After the survey was administered, the Statistics of the Continuous Register of January 1, 2015 (Instituto Nacional de Estadística, 2016) were released. We updated our tables for estimation so that the distribution of names reflected more accurately the distribution at the time of the survey (December 2014 - January 2015). Conveniently, as of January 1, 2015, the birth decades of 2000 and 2010 together represent the population under the age of 15. For obtaining greater consistency with the survey question ("How many people over the age of 15 do you know?"), we excluded these two decades in determining the total prevalence of names in the population.

Together, the selected names represented 852,929 people resident in Spain over the age of 15 years (that is, 2.2% of the total population over 15 years). Figure 1 shows the fractions that men and women of different birth decades occupy in the total population over 15 years old, as well as the fractions that men and women of different birth decades occupy in the subpopulation of people with the 14 selected names, older than 15 years. The two lines for the male population hardly deviate from one another, and neither do the lines for the female population, showing that the set of names represents on a small scale the Spanish population in terms of gender and birth decade.

-- PLEASE INSERT FIGURE 1 ABOUT HERE; IN COLOR--

We observed a considerable variation in the prevalence of the selected names across geographic regions, although our observation is approximate since we could not select the population older than 15 years in the case of regions[3]. Regional variations have not been reported in other studies, so we do not know if this is a Spanish particularity. However, the number of names given by women and men by province (for the 43 provinces that had at least 5 respondents in the sample) did not correlate significantly with the prevalence of these names in the same provinces (r_{men}=-0.144, ns; r_{women}=0.034; ns).

Overall, the average number of people that respondents knew with these names correlated strongly with the prevalence of these names in the population (r=0.82, $p <$ .0001), supporting their validity. Figure 2 shows that the female name Consuelo is the furthest away from the trend line. The correlation would be r=0.92 ($p <$ .0001) when removing the name. However, this name is associated with relatively older women (in the population, the mean age of women called Consuelo is 65.3 and Maria Consuelo 56.5 years). Excluding the name would unbalance Figure 1 and therefore we decided to keep the name. Figure 2 also provides information to correct for memory error (Maltiel et al., 2015). The trend line does not cross the origin, suggesting a certain over-estimation of the less common names.

-- PLEASE INSERT FIGURE 2 ABOUT HERE--

Core network size was measured using the simpler summation method (e.g., Bernard et al. 2010): We asked respondents to estimate (a) the *number of relatives* over the age of 15 (s)he had with whom they have a relationship. Response categories were 0, 1-5, 6-10, 11-20, 21-30, 31-40, 41-50, more than 50; (b) the *number of friends* over the age of 15 with whom they meet or go out or talk about personal issues (categories 0, 1, 2, 3, 4, 5, 6-10, 11-15, 16-20, 21-30, more than 30). For the regression analyses, we recoded the categories to their midpoint values, and the highest values to 60 and 35, respectively. Even so, we repeated all analyses that included these variables, replacing the numerical variables by the categorical ones, as a control. Of the total effective sample, 4 and 14 cases, respectively, had missing values on the questions about relatives and friends.

Social groups. We measured four variables: (1) *gender*, (2) *age* in years, standardized for regression; (3) highest *level of education* completed (ordinal variable, ranging from no studies to higher education; recoded to a variable with 6 categories; primary education; secondary education 1st stage; secondary education 2nd stage, professional education; higher education; other - see Table 2); and (4) net monthly *household income* (ordinal variable with 11 categories, ranging from no income of any type to 6,000 Euros or more; recoded to a variable with 4 categories: less than 600; 601 to 1,200; 1,201 to 2,400; and more than 2,400 Euros). Age and gender had no missing values, education had 8 missing values and income 686.

Participation in social foci. We used five variables to measure participation in social foci: (1) *employment situation*, indicating participation in work and study contexts (categorical variable with 4 categories: inactive -unpaid domestic work, pensioners and retirees-; employed; 3 unemployed -after work or looking for first job-; student. The category other has been combined with missing values); (2) *civil status*, as a proxy for participation in family contexts (categories married, single, widowed, divorced/separated); (3) whether the respondent has *children of minor age* or not (dummy variable as a proxy for participation in parental and school networks); (4) *frequency of religious service attendance* (with the categories hardly - combining “hardly ever” and “a couple of times a year”-, approximately monthly, and approximately weekly - combining “almost all Sundays” and “more often”-), and (5) *membership of associations*. For the latter, we used nine questions about the membership of different types of associations: political parties, trade unions or associations of entrepreneurs, professional colleges, parishes or other religious organizations, sports clubs or groups, cultural or leisure groups, social or human rights organizations, youth or student associations, and other organizations (parent-teacher associations, neighborhood associations, etc.). For each, respondents were asked whether they (1) belonged and participated actively, (2) belonged but did not participate actively, (3) belonged in the past, but no longer, or (4) never belonged to such association. First, we counted the number of times respondents said they "belonged and participated actively". We collapsed responses 2 and higher to a single category due to the low frequency for higher categories. Respondents that had missing values on one or multiple of the underlying questions were coded as missing. This variable was used in the regression analysis. Afterwards, we repeated the analysis nine times, substituting the count variable by each type of association, separately, to investigate whether some types of associations were more consequential for acquaintanceship volume than others.

Adequate social support availability was measured on four dimensions. Respondents were asked "Now please think of your complete social circle of family, friends, neighbors, and other acquaintances. Approximately, to how many of them could you go (...)", which was completed with "(...) in case you needed it to care for you if you fell ill?", "(...) who could lend you money if you needed it?", "(...) to talk if you had a problem, felt sad or depressed?", and "(...) who could help you find a job?". Responses varied from "no one", with 1 unit increases up to "11 or more". For the analyses, we dichotomized the responses to differentiate between *unavailable and marginal support* (0-1 persons) and *adequate support* (2 or more persons; after Fischer 1982). We thus focus on a minimum threshold as an indicator of social support availability, assuming that the difference between say 4 and 5, or 9 and 10 support providers, is not as important. Of the total effective sample, 82, 30 and 25 cases, respectively, had missing values on the questions about lending money, help with illness, and talking about problems, so for these analyses, the sample size was slightly lower. We restricted our analyses of help with finding a job to the persons who were currently employed, unemployed or students, N = 1,608. Of these people N = 1,502 answered the question about available support with finding a job.

Analysis

We used the estimation method developed by Maltiel et al. (2015; see Section 2) implemented in "NSUM" R library. Network size was estimated both with Killworth et al.´s (1998) original method, and with Maltiel et al.'s barrier model, in the latter case keeping the last name as unknown population. To determine the number of iterations for the MCMC, we used the Raftery-Lewis diagnostic and set the number at 40,000 with a burn-in length of 1,000, and we kept 4,000 iterations after thinning for calculating network size. We ran the iterations twice. After estimation of the unknown population, we applied the Gelman-Rubin diagnostic to evaluate the convergence of the MCMC chains for the network size parameters. For the vast majority of individuals, the MCMC algorithm converged well, with the Gelman-Rubin diagnostic < 1.015. For 79 cases (3.5%) however, the chains did not reach convergence for the estimation of degree. Closer inspection showed that these included all 47 respondents who did not know anyone with any of the given names, 17 of the 53 respondents who knew in total 1 person with the set of names, and 10 respondents who knew 2 to 6 persons in total with the set of names. We decided not to exclude them as this would restrict the range - the majority of the cases whose estimates did not converge knew the least names of all. Instead, we imputed for these cases the median estimated network size of the respondents who reported the same total number of names as they did, but whose estimates converged. For the 47 respondents who did not know anyone with these names, however, this strategy was not viable as the barrier model did not converge for any of them. We could have imputed the Killworth estimate of network size for these respondents, but decided against it as it is unrealistically low (namely, $\hat{d}$ =1). Instead, we imputed the value 113 for these 47 cases, which was at precisely the same distance (33) from the median degree of people who knew in total 1 person within the set of name (Median=146), as the latter was removed from the median of people who knew in total 2 persons with those names (Median=179). The latter decision is admittedly arbitrary. However, for the estimate of the median degree, the decision does not bear any consequences as these cases are at the bottom of the scale. For the negative binomial and logistic regressions, the decision had a negligible impact. To control for the robustness of the results, we also ran all regressions with the Killworth estimate of degree instead of the barrier model estimate, which gave highly similar results[4] - not surprising given the high correlation between the two estimates - (see Results section).

We used mostly non-parametric statistics to describe bivariate relations. To predict acquaintanceship volume, we used negative binomial regressions with a log link function, which is appropriate for overdispersed count data, and it allows estimation of the dispersion coefficient from the data. In Model 1, we included gender, age, education, monthly household income as explanatory variables, and in Model 2 we added employment situation, civil status, whether the respondent has children of minor age or not, frequency of attendance to mass and other religious events, and membership of associations. The only continuous variable, age, does not have missing cases. For the categorical variables, we excluded cases that had missing values on one or more explanatory variables, except for household income, which had a large number of missing cases (see Table 2) and the non-response appeared to be not random (e.g., when crossed with education, the persons who had missing values on income seemed to have a similar education profile as the €1,200-2,400 group). We therefore retained the category for missing values as a separate category in the analysis.

Finally, to predict social support availability (dichotomized variables), we used logistic regression analysis. In Model 1, we included gender, age in years, education, monthly household income as explanatory variables, but we also added employment situation, for its relevance for help with finding a job (we maintained it in the other analyses for comparison). In Model 2, we added three variables to the former model: the log-transformed estimated network size, our variable of interest, and the two variables of core network size.

4. Results

Central tendency and distribution of acquaintanceship volume and core network size

The 2,330 people who form the effective sample reported that they have between 0 and 138 acquaintances with the 14 names. 47 respondents (2.0%) indicated they did not know anyone with these names. We first estimated the size of acquaintanceship networks using the original method of Killworth, which resulted in a median network size of 531.7, with an interquartile range of 290.0 to 870.0 (M=654.1; SD=543.1). With Maltiel et al.'s (2015) more sophisticated barrier model, we observed a surprisingly similar median network size of 531.5, with a smaller interquartile range of 326.8-823.3 (M=657.0; SD=494.7). For the regressions, we used the degree estimated with the barrier model.

Comparing the distributions of acquaintanceship volume estimated with the two methods (see Figure 3), we also observed a more compressed range for the last method. Nevertheless, the correlation between the two measures is very high, with Spearman's ρ=0.99 (p <.01; scatter plot in Figure 4), since estimation with the method of Maltiel et al. (2015) showed that the overdispersion in the name populations was negligible. Our estimates are quite similar to those of US studies: DiPrete et al. (2011) observed that a median size of 550, with an interquartile range of 400 to 780. Our median is close to these values, although the interquartile range is larger in our case. Zheng et al. found a median of 610, and McCormick et al. (2010), on the basis of the same data as Zheng et al., of 472, but they did not indicate the IQR. However, as for McCormick et al., our maximum is also close to 6,000. In fact, 390 persons had values of 1,000 and more, while 54 persons had values over 2,000, 12 over 3,000, 4 over 4,000 and 1 over 5,000. Although these high-connectivity nodes differed significantly from the lower-connectivity nodes on many of the explanatory variables, there was not a single variable (e.g., income, or associations) that set them clearly apart from the rest. As we used non-parametric models and/or log-transformation of acquaintanceship volume, there was no need to exclude or recode the higher values from the data base.

-- PLEASE INSERT FIGURES 3 AND 4 ABOUT HERE--

In comparison, respondents' self-estimates of the number of relatives with whom they had a personal relationship has a median of 11-20, and an interquartile range varying from 6-10 to 21-30 relatives. When recoding each category to the center of the range and the highest value to 60, we obtained an average of 17.4. Nine respondents indicated they did not have a personal relationship with any relative, whereas 96 indicated they had a personal relationship with more than 50 relatives. The self-estimates of the number of friends, on the other hand, has a median of 4 and an interquartile range from 2 friends to 6-10 friends. When recoding each category to the center of the range and the highest value to 35, we obtained an average of 5.8. In total, 175 respondents indicated they did not have any friend and 24 indicated they had more than 30. Summing over relatives and friends, core networks have a median size of 19.1 and an average of 23.1, roughly corresponding to the size of what Dunbar called *sympathy groups,* on the inner circles of the networks (6-20 people).

Acquaintanceship volume was positively, but not strongly, related with the number of relatives and the number of friends individuals reported to know (With the number of relatives, Spearman's $\rho = 0.235$; $p < .001$; with friends, $\rho = 0.158$; $p < .001$; Table 2; with the sum of relatives and friends, $\rho = 0.251$; $p < .001$). The number of relatives was also positively related with the number of friends ($\rho = 0.251$; $p < .001$), in contrast with findings of Roberts et al. (2009), who found a negative relation between the two and interpreted it in terms of time constraint (see Introduction).

Differences in acquaintanceship volume between social groups

Table 2 presents the descriptive statistics of the explanatory variables and their bivariate relations with acquaintanceship volume. With regard to the social groups, it shows that younger, male, higher educated people, and people in higher income groups report larger acquaintanceship volumes.

We then regressed the size of acquaintanceship networks on these attributes using negative binomial regression analysis. The model (see Table 3, Model 1) has a significantly better fit than the empty model (χ^2=117.3; *df*=11, *p*<.001). We estimated the dispersion coefficient to be 0.391; the 95% confidence interval does not include 0, which implies overdispersion, confirming the appropriateness of the negative binomial regression model. The regression model largely confirms the conclusions drawn from bivariate relations: social groups (of gender, age, education and income) differ in acquaintanceship volume. However, for income, the median network size increases more or less monotonically when considered bivariately (see Table 2), but when controlling for other variables, only the highest income category has a significant regression coefficient at *p*<.01. A quadratic effect of age was tested and was found significant for Model 1 (higher ages having increasingly smaller networks) but not for Model 2 (additional analyses not reported in table).

--PLEASE INSERT TABLES 2 AND 3 ABOUT HERE--

Social contexts as potential mediators of network inequality

As Table 2 showed, the five variables we use as proximal variables for social foci are all bivariately significantly related with acquaintanceship volume, except attendance to religious services. Relations for having minor children, employment situation, and membership of associations were as expected, but for civil status we observed that not married, but single people had the largest networks.

In the regression model (Table 3, Model 2, model fit χ^2=187.4; *df*=22, *p*<.001), controlling for all other variables, membership in associations has a clearly significant, positive relation with acquaintanceship volume: being an active member of one association is associated with an estimated increase in network size of 11% (multiplication factor 1.113), and participating actively in multiple associations with an increase of 30% (multiplication factor 1.302). Thus, participation in associations is related with acquaintanceship volume, in line with what the literature on the benefits of associative fabric for social cohesion suggests (e.g. Putnam 2000). We further analyzed whether all types of associations were associated for network size (not in table), and observed that political parties, parishes or other religious organizations, sports clubs or groups, cultural or leisure groups, youth or student associations (*p*<001), and other organizations (PTAs, neighborhood associations, etc.; *p*<.01) are significantly associated with network size. For most of these cases, people who participate actively have larger estimated network sizes compared to those who have never been active. For political parties however, non-active members have larger networks than active members, and in religious associations, former members have larger networks than current members. Participation in social or human rights organizations has a marginally significant regression coefficient (*p*<.05; marginally given the sample size), while the coefficients of participation in trade unions, associations of entrepreneurs or professional colleges were not significant.

Employment situation also has a significant regression coefficient. Against our expectations, students, who we supposed would have many opportunities to meet new people at their universities/colleges, do not significantly differ in estimated network size from people who are inactive at the labor market when we control for age and other variables. On the other hand, the employed (21% increase), but to our surprise also the currently unemployed (19% increase) have larger networks than the inactive population.

Having children of minor age, civil status, and attendance to religious services -proximal variables for participation in parental and school networks, family networks, and religious communities- are not associated with acquaintanceship volume when controlled for the other variables. More precisely, the bivariate relation between having children of minor age and acquaintanceship volume was significant (Table 2), but its significance disappeared in a regression analysis when controlling for other variables (Table 3).

Model 2 further shows that the inclusion of social foci attenuates the coefficients of the social groups, such that the coefficients of age and high income are no longer significant, while the coefficients of education decreased in size. This (in addition to the significant relations between social groups and social foci; not in table) suggests that social foci explain, at least partially, the difference between social groups. The main effect of gender, however, maintains its original effect.

Acquaintanceship volume and the availability of adequate support

We now turn to the relation between acquaintanceship volume and social support. Table 4 shows descriptive statistics of the social support variables. For the support with finding a job, the table shows that three quarters of the respondents who were employed, unemployed, or students (73%) perceived to have adequate support (defined as having at least two persons who could help them), while 27% thought they had no (19%) or marginal (8%) help. When people who mentioned at least one person were asked who came to mind first, 41% thought of a friend, 26% a first-degree family member, 14% of their partner, 7% of other relatives, and 11% mentioned of other people (colleagues, neighbors and others).

--PLEASE INSERT TABLE 4 ABOUT HERE--

When we regressed the variable on individual attributes (see Table 5), Model 1 shows that the likelihood of having at least two people able to help individuals find a job differed slightly across social groups. The highest income groups and the currently employed were more likely, while older respondents were less likely to know at least 2 persons who could help them find a job. In Model 2, we added core network size (family and friends) and acquaintanceship volume as explanatory variables. This model (Model 2) showed that controlling for individual attributes, the number of friends is related with a larger probability of adequate support with finding help, while the coefficients of the number of relatives, and surprisingly, acquaintanceship volume, were smaller. The two models have a good fit (Model 1: χ^2=110.0; *df*=13, *p*<.001; Model 2: χ^2=144.8; *df*=16, *p*<.001). As a control, we repeated the analysis of Model 2 replacing the numerical variables of friends and relatives by the categorical variables. In this analysis, the number of relatives is no longer significant, but acquaintanceship volume remains significant.

--PLEASE INSERT TABLE 5 ABOUT HERE--

For lending money when needed, descriptive statistics (see Table 4) show that again three quarters of the population (72%) perceived to have at least two persons who could help them, while 28% had no one (13%) or only marginal help (15%). Model 1 of Table 5 shows that there are considerable differences across social groups in the availability of financial support. In particular, the higher income groups and the higher educated had 2 to 3 times higher probabilities to know people who could help them than people with no or low education and income. Furthermore, age was negatively associated with such support. To illustrate, descriptive bivariate statistics showed that 82% of the 18-35 years old, 78% of the 36-50 years old, 65% of the 51-65 years old, and 56% of the persons older than 65 years had this type of support. Those who mentioned they had at least one person were also asked who the first person was who came to mind. While parents were mentioned by the majority of 18-35 year olds and by a minority of people of 50+ as could be expected, children replaced parents progressively in the older cohorts, and siblings played a substantial role in the 51-65 years category, such that for all groups, a majority (between 66 and 79%) of the first mentioned was first-degree kin. This suggests that the result for age may be more complex than the decreasing presence of parental help. Another explanation may be that younger Spaniards are -at least in the current context of high precariousness and youth unemployment- more likely to ask for help, shaping their perception of support availability. Overall, the models have a good fit (Model 1: χ^2=221.8; *df*=14, *p*<.001; Model 2: χ^2=283.7; *df*=17, *p*<.001). Again, we repeated the analysis of Model 2 replacing the numerical variables of friends and relatives by the categorical variables. In this case, the number of friends was also significant at the *p*<.001 level, similar to the number of relatives and acquaintanceship volume.

In comparison to the former two support types, the likelihood that adequate social support is available during illness is generally higher (see Table 4): 89% of the respondents thought they had at least two persons who could help them out, while 11% reported they had nobody (2%) or 1 person (9%) who could help them. Model 1 in Table 6 shows that there are only small differences across social groups in this likelihood, indicating that there is little social inequality. Furthermore, Model 2 shows that the size of the two core networks have significant associations with social support availability, especially friends. In contrast, acquaintanceship volume was not significantly related to support availability, after controlling for core network size and individual variables. Overall, the models have a good fit (Model 1: χ^2=65.9; *df*=14, *p*<.001; Model 2: χ^2=138.6; *df*=17, *p*<.001). Replacing the numerical variables of friends and relatives by the categorical variables gave similar results for both core and total network size.

The likelihood of having at least two persons to talk to about problems was 85%, with 15% reporting they had nobody (3%) or only one person (12%) to talk to about such problems (see Table 4). As Model 1 in Table 6 shows, gender, and surprisingly, household income matter considerably. All else being equal, men were 34% less likely, and people with no or very low household incomes were considerably less likely to report they had at least two persons with whom they could talk about personal problems. Model 2 shows that the number of friends, but not relatives, was significantly related to the number of people to whom respondents could turn if they needed to talk. Furthermore, acquaintanceship volume has a significant association with support availability even after controlling for the numbers of relatives and friends and for individual attributes. Both models have a good fit (Model 1: χ^2=76.9; *df*=14, *p*<.001; Model 2: χ^2=141.6; *df*=17, *p*<.001). Replacing the numerical variables of friends and relatives in Model 2 by the categorical variables, we found that the coefficient of the number of relatives is also significant at the *p*<.001 level, although none of the categories is significant, while acquaintanceship volume remained significant at the *p*<.001 level.

--PLEASE INSERT TABLE 6 ABOUT HERE--

5. Conclusion

This paper aimed to estimate the distribution of broader acquaintanceship volume for the general population of Spain with Spanish nationality, and to investigate its variation among social groups and its relations with social support availability. For this aim, we designed a Network Scale-Up Method module that was incorporated in a National Barometer administered at the end of 2014 and beginning of 2015 by the Spanish Center for Sociological Research to a large, representative sample. To estimate acquaintanceship volume, we used the generalized estimation methods developed by Maltiel et al. (2015).

Our results show that the median estimated network size in Spain is 532, which comes quite close to the estimates of DiPrete et al. (550) and McCormick et al. (472) for the American population - a far stretch from Dunbar´s number (150), but nonetheless close to Dunbar's other theoretical sweet spot of 500. Individuals' self-estimates of the number of friends and relatives with whom they have a personal relationship summed up to roughly 23 on average, which suggests that most of the ties in the broader acquaintanceship networks can be considered weak.

Individual variation in acquaintanceship volume is large (interquartile range of 500) and follows a positively skewed distribution, similar to the one observed by McCormick, Salganik, and Zheng (2010) for the US, and also rather similar to that of core network size on a smaller range (Van Tilburg, 1995). The distribution has a long tail: Although relatively few persons exceed 1,000 acquaintances (*N* = 390), those who do reach estimates of up to 6,000.

A relevant question is therefore, what predicts inter-individual differences? While other research has focused mainly on biological and psychological characteristics, in this paper we explored differences across social groups in acquaintanceship volume. We expected that acquaintanceship volume would (as core network size) show inequalities by gender, age, income, and education. Indeed, we found bivariate associations with acquaintanceship volume for all variables, suggesting that women, older people, and people with lower levels of education and income have smaller network sizes. The associations with income and education confirm results of DiPrete et al., while age and gender coefficients were not significant in their analysis. When entered together in a regression model, however, the regression coefficient of age becomes insignificant and that of income marginally significant (only the highest category is related to larger network sizes compared to the reference category). Together, these results suggest that the size of broader acquaintanceship networks is characterized to some degree by inequality.

Furthermore, we expected that social groups might differ in acquaintanceship volume due to unequal participation to social contexts. Our results show that active membership in associations of almost all types is associated with higher acquaintanceship volume. This result adds evidence to the empirical literature emphasizing the civic value of associational life for creating social capital by connecting individuals to a wide range of others. Current participation is most associated with acquaintanceship volume, although in some occasions, having participated in the past is also associated with acquaintanceship volume. Furthermore, not only working people, but also unemployed people had larger networks than people who are inactive at the labor market, perhaps showing an effect of current *and* former participation in work contexts - but this latter interpretation should be further explored. On the other hand, having children of minor age (which we theorized would give access to parental and school networks), being married (extended family networks), frequency of attendance to religious events (religious communities), and being a student (universities/colleges) did not have significant coefficients. Therefore, our findings on participation in social foci are mixed.

Last, we investigated whether the broader acquaintanceship volume has any relation, beyond the size of core networks, with the social support individuals can mobilize when they are ill or when they need someone to talk to. Research has suggested that broader acquaintanceship networks have distinct and similar functions to core networks in the provision of social support. Our results show that indeed acquaintanceship volume is associated with social support availability. Even when we control for individual characteristics and self-estimates of the number of friends and relatives, acquaintanceship volume has a significant and positive relation with the availability of financial and emotional support - but surprisingly, not with support with finding a job, for which we assumed weak ties would be most important. Thus, even with equal core network size, people who have more acquaintances are more likely to have adequate social support in the financial and emotional area.

Our paper has several strengths. Our results are based on a survey administered face-to-face by trained interviewers to a large, representative sample of the national population of Spain. The method that we employed for estimating acquaintanceship volume has the strength that it can be easily integrated in surveys as it takes very little time of the respondents. However, it also has limitations. First, it is difficult to estimate network size for social groups that have different naming patterns than the majority group if there are social barriers between these groups (McCarty, Bernard, Killworth, Shelley, & Johnsen, 1997). Our study therefore focuses only on the population whose nationality is Spanish (cf. Paniotto et al. 2009), but a better solution must be found to estimate acquaintanceship volume for minority populations as well. Second, we found regional variations in naming, and to our knowledge such variations have not been investigated before. The number of names people mentioned in the different regions did not correlate with the prevalence of the names, but even so it may be prudent to study such differences in more detail. Apart from acquaintanceship in particular, surveys have their limitations in the depth with which concepts can be investigated and the potential for exploring causality.

To conclude, we believe that the results call for more attention to broader acquaintanceship networks. As we mentioned in the introduction, personal network research has mostly focused on individuals' closest ties. These ties may well be more supportive when compared to weak ties on a relationship basis, but acquaintanceship networks as a set are assumed to have similar as well as distinct functions to core networks that are not well understood. For example, there seems to be an implicit assumption that acquaintanceship networks are heterogeneous and large enough for individuals to be similarly exposed to weak ties of all sorts, as the assumption of random mixing underlying the Network Scale Up Method illustrates. Our paper suggests that acquaintanceship networks may be similarly shaped by social contexts as core networks, such that social groups differ in acquaintanceship volume. Other research has questioned the assumed heterogeneity of acquaintanceship networks (DiPrete et al., 2011). As empirical evidence is scarce, there is a need for further research into the size, structure, composition, and functionality of acquaintanceship networks, extending our analysis beyond the constellation of its core. It is particularly important to investigate, at both the relationship and the network level, the interaction *between* strong and weak ties - or ties of varying strengths, to avoid further artificial dichotomization. This will allow us to obtain a deeper understanding of the distinct and similar functions of different sets of ties, and whether strong ties can be substituted for weaker ones when strong ones are unavailable as some researchers have suggested (see Introduction). It should also take into account the dynamics of tie strength itself: do weak ties morph into stronger ones or strong into weak ties, or do different sets of ties mostly have a life cycle of their own (McPherson, 2009)? And do strong ties affect (e.g., as gatekeepers, Chua 2013) the weak ties one acquires? Such research will advance our understanding of personal networks of larger volumes.

References

Arnaboldi, V., Guazzini, A., & Passarella, A. (2013). Egocentric online social networks: Analysis of key features and prediction of tie strength in Facebook. *Computer Communications*, *36*(10–11), 1130–1144. http://doi.org/10.1016/j.comcom.2013.03.003

Berkman, L. F., & Glass, T. (2000). Social integration, social networks, social support, and health. In L. F. Berkman & I. Kawachi (Eds.), *Social Epidemiology* (pp. 137–173). New York: Oxford University Press.

Bernard, H. R., Hallett, T., Iovita, A., Johnsen, E. C., Lyerla, R., McCarty, C., et al. (2010). Counting hard-to-count populations: the network scale-up method for public health. *Sexually Transmitted Infections*, *86 Suppl 2*, ii11-i15. http://doi.org/10.1136/sti.2010.044446

Bernard, H. R., Johnsen, E. C., Killworth, P. D., & Robinson, S. (1989). Estimating the size of an average personal network and of an event subpopulation. In M. Kochen (Ed.), *The Small World* (pp. 159–175). Norwood, NJ: Ablex Press.

Bernard, H. R., Johnsen, E. C., Killworth, P. D., & Robinson, S. (1991). Estimating the size of an average personal network and of an event subpopulation: Some empirical results. *Social Science Research*, *20*, 109–121. http://doi.org/10.1016/0049-089X(91)90012-R

Bernard, H. R., & Shelley, G. A. (1987). How much of the network does the GSS and RSW dredge up? *Social Networks*, *9*, 49–61.

Bickart, K. C., Wright, C. I., Dautoff, R. J., Dickerson, B. C., & Feldman Barrett, L. (2011). Amygdala volume and social network size in humans. *Nature Neuroscience*, *14*, 163–164. http://doi.org/10.1038/nn.2724

Blau, P. M. (1977). A macrosociological theory of social structure. *American Journal of Sociology*, *83*(1), 26–54.

Boissevain, J. (1974). *Friends of friends: Networks, manipulators and coalitions*. Oxford: Blackwell Publishing.

Bojarczuk, S., & Mühlau, P. (2018). Mobilising social network support for childcare: The case of Polish migrant mothers in Dublin. *Social Networks*, *53*, 101–110. http://doi.org/10.1016/j.socnet.2017.04.004

Bolíbar, M., Martí, J., & Verd, J. M. (2015). Just a Question of Time? The Composition and Evolution of Immigrants' Personal Networks in Catalonia. *International Sociology*, *30*(6), 579–598. http://doi.org/10.1177/0268580915579925

Brashears, M. E. (2013). Humans use Compression Heuristics to Improve the Recall of Social Networks. *Scientific Reports*, *3*(1513), 1–7. http://doi.org/10.1038/srep01513

Brashears, M. E., Hoagland, E., & Quintane, E. (2016). Sex and network recall accuracy. *Social Networks*, *44*(74), 84. http://doi.org/10.1016/j.socnet.2015.06.002

Brewer, D. D. (2016). Gynocentric communication among nuclear kin in the USA and Spain. *PeerJ Preprints*, *4:e647v2*. http://doi.org/https://doi.org/10.7287/peerj.preprints.647v2

Chua, V. (2013). Categorical sources of varieties of network inequalities. *Social Science Research*, *42*(5), 1236–1253. http://doi.org/10.1016/J.SSRESEARCH.2013.04.002

Cornwell, B., Laumann, E. O., & Schumm, L. P. (2008). The Social Connectedness of Older Adults: A National Profile. *American Sociological Review*, *73*, 185–203. http://doi.org/http://www.jstor.org/stable/25472522.

Curry, O., & Dunbar, R. I. M. (2013). Do Birds of a Feather Flock Together? The Relationship between Similarity and Altruism in Social Networks. *Human Nature*, *24*(3), 336–347. http://doi.org/10.1007/s12110-013-9174-z

Dávid, B., Huszti, E., Barna, I., & Fu, Y. C. (2016). Egocentric contact networks in comparison: Taiwan and Hungary. *Social Networks*, *44*, 253–265. http://doi.org/10.1016/j.socnet.2015.10.001

De Miguel Luken, V., & Tranmer, M. (2010). Personal support networks of immigrants to Spain: A multilevel analysis. *Social Networks*, *32*(4), 253–262. http://doi.org/10.1016/j.socnet.2010.03.002

De Sola Pool, I., & Kochen, M. (1978). Contacts and influence. *Social Networks*, *1*(1), 5–51.

Desmond, M. (2013). Disposable Ties and the Urban Poor. *American Journal of Sociology*,

117(5), 1295–1335. http://doi.org/http://www.jstor.org/stable/10.1086/663574 .
DiMaggio, P., & Garip, F. (2012). Network Effects and Social Inequality. *Annual Review of Sociology*, *38*(1), 93–118. http://doi.org/10.1146/annurev.soc.012809.102545
DiPrete, T. A., Gelman, A., McCormick, T., Teitler, J., & Zheng, T. (2011). Segregation in social networks based on acquaintanceship and trust. *American Journal of Sociology*, *116*(4), 1234–1283. http://doi.org/https://doi.org/10.1086/659100
Dunbar, R. I. M. (1993). Co-Evolution of Neocortex Size, Group Size and Language in Humans. *Behavioral and Brain Sciences*, *16*(4), 681–735.
Dunbar, R. I. M., Arnaboldi, V., Conti, M., & Passarella, A. (2015). The structure of online social networks mirrors those in the offline world. *Social Networks*, *43*, 39–47. http://doi.org/10.1016/j.socnet.2015.04.005
Dunbar, R. I. M., & Sosis, R. (2018). Optimising human community sizes. *Evolution and Human Behavior*, *39*(1), 106–111. http://doi.org/10.1016/j.evolhumbehav.2017.11.001
Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The Benefits of Facebook "Friends:" Social Capital and College Students' Use of Online Social Network Sites. *Journal of Computer-Mediated Communication*, *12*(4), 1143–1168. http://doi.org/10.1111/j.1083-6101.2007.00367.x
Ellison, N. B., Steinfield, C., & Lampe, C. (2011). Connection strategies: Social capital implications of Facebook-enabled communication practices. *New Media & Society*, *13*(6), 873–892. http://doi.org/10.1177/1461444810385389
Feehan, D. M., & Salganik, M. J. (2016). Generalizing the Network Scale-Up Method: A New Estimator for the Size of Hidden Populations. *Sociological Methodology*, *46*(1), 153–186. http://doi.org/10.1177/0081175016665425
Feld, S. L. (1981). The Focused Organization of Social Ties. *American Journal of Sociology*, *86*(5), 1015–1035. http://doi.org/10.1086/227352
Fingerman, K. L. (2009). Consequential strangers and peripheral ties: The importance of unimportant relationships. *Journal of Family Theory and Review*, *1*, 69–86. http://doi.org/10.1111/j.1756-2589.2009.00010.x
Fischer, C. S. (1982). *To dwell among friends: Personal networks in town and city*. Chicago: University of Chicago Press.
Freeman, L. C., & Thompson, C. R. (1989). Estimating Acquaintanceship Volume. In M. Kochen (Ed.), *The Small World* (pp. 147–158). Norwood, NJ: Ablex.
Fu, Y. C. (2005). Measuring personal networks with daily contacts: A single-item survey question and the contact diary. *Social Networks*, *27*, 169–186. http://doi.org/10.1016/j.socnet.2005.01.008
Fu, Y. C. (2007). Contact Diaries: Building Archives of Actual and Comprehensive Personal Networks. *Field Methods*, *19*(2), 194–217. http://doi.org/10.1177/1525822X06298590
Granovetter, M. S. (1973). The Strength of Weak Ties. *The American Journal of Sociology*, *78*(6), 1360–1380.
Grossetti, M. (2005). Where do social relations come from?: A study of personal networks in the Toulouse area of France. *Social Networks*, *27*(4), 289–300. http://doi.org/10.1016/j.socnet.2004.11.004
Grossetti, M. (2007). Are French networks different? *Social Networks*, *29*(3), 391–404. http://doi.org/10.1016/j.socnet.2007.01.005
Gurevitch, M. (1961). *The social structure of acquaintanceship networks.* Cambridge, MA: MIT.
Harling, G., Morris, K. A., Manderson, L., Perkins, J. M., & Berkman, L. F. (2018). Age and Gender Differences in Social Network Composition and Social Support Among Older Rural South Africans : Findings From the HAALSI Study. *Journals of Gerontology: Social Sciences*, 1–12. http://doi.org/10.1093/geronb/gby013
Hill, R. A., & Dunbar, R. I. M. (2003). Social network size in humans. *Human Nature*, *14*(1), 53–72. http://doi.org/10.1007/s12110-003-1016-y
Hofstra, B., Corten, R., Van Tubergen, F., & Ellison, N. B. (2017). Sources of Segregation in Social Networks: A Novel Approach Using Facebook. *American Sociological Review*, *82*(3), 625–656. http://doi.org/10.1177/0003122417705656
Huszti, É., Dávid, B., & Vajda, K. (2013). Strong Tie, Weak Tie and In-betweens: A Continuous

Measure of Tie Strength Based on Contact Diary Datasets. *Procedia - Social and Behavioral Sciences*, *79*, 38–61. http://doi.org/10.1016/j.sbspro.2013.05.056

Jacobs, J. A., & Gerson, K. (2004). *The Time Divide: Work, Family, and Gender Inequality*. Cambridge, MA: Harvard University Press.

Kahn, R. L., & Antonucci, T. C. (1980). Convoys over the life course: Attachment, roles, and social support. In P. B. Baltes & O. Brim (Eds.), *Life-span Development and Behavior* (pp. 253–286). New York: Academic Press. http://doi.org/10.1038/nature12518

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science*, *306*(5702), 1776–1780. http://doi.org/10.1126/science.1103572

Killworth, P. D., & Bernard, H. R. (1978). The reversal small-world experiment. *Social Networks*, *1*(2), 159–192. http://doi.org/http://dx.doi.org/10.1016/0378-8733(78)90018-7

Killworth, P. D., Johnsen, E. C., Bernard, H. R., Ann Shelley, G., & McCarty, C. (1990). Estimating the size of personal networks. *Social Networks*, *12*(4), 289–312. http://doi.org/10.1016/0378-8733(90)90012-X

Killworth, P. D., Johnsen, E. C., McCarty, C., Shelley, G. A., & Bernard, H. R. (1998). A social network approach to estimating seroprevalence in the United States. *Social Networks*, *20*(1), 23–50. http://doi.org/10.1016/S0378-8733(96)00305-X

Killworth, P. D., McCarty, C., Bernard, H. R., Johnsen, E. C., Domini, J., & Shelley, G. A. (2003). Two interpretations of reports of knowledge of subpopulation sizes. *Social Networks*, *25*(2), 141–160.

Killworth, P. D., McCarty, C., Johnsen, E., Bernard, H. R., & Shelley, G. (2006). Investigating the Variation of Personal Network Size Under Unknown Error Conditions. *Sociological Methods & Research*, *35*(1), 84–112. http://doi.org/10.1177/0049124106289160

Kudo, H., & Dunbar, R. I. M. (2001). Neocortex size and social network size in primates. *Animal Behaviour*, *62*(4), 711–722. http://doi.org/https://doi.org/10.1006/anbe.2001.1808

Lewis, P. A., Rezaie, R., Brown, R., Roberts, N., & Dunbar, R. I. (2011). Ventromedial prefrontal volume predicts understanding of others and social network size. *Neuroimage*, *57*, 1624–1629. http://doi.org/10.1016/j.neuroimage.2011.05.030

Lonkila, M. (1997). Informal Exchange Relations in Post-Soviet Russia: A Comparative Perspective. *Sociological Research Online*, *2*(2).

Lu, Y.-E., Roberts, S., Lio, P., Dunbar, R., & Crowcroft, J. (2009). Size Matters: Variation in Personal Network Size, Personality and Effect on Information Transmission. In *2009 International Conference on Computational Science and Engineering* (pp. 188–193). IEEE. http://doi.org/10.1109/CSE.2009.179

Lubbers, M. J., Molina, J. L., & McCarty, C. (2007). Personal Networks and Ethnic Identifications: The Case of Migrants in Spain. *International Sociology*. http://doi.org/10.1177/0268580907082255

Maltiel, R., Raftery, A. E., McCormick, T. H., & Baraff, A. J. (2015). Estimating Population Size Using the Network Scale Up Method. *The Annals of Applied Statistics Institute of Mathematical Statistics in The Annals of Applied Statistics*, *9*(3), 1247–1277. http://doi.org/10.1214/15-AOAS827

Marsden, P. V. (1987). Core Discussion Networks of Americans. *American Sociological Review*, *52*(1), 122–131.

McCarty, C., Bernard, H. R., Killworth, P. D., Shelley, G. A., & Johnsen, E. C. (1997). Eliciting representative samples of personal networks. *Social Networks*, *19*, 303–323.

McCarty, C., Killworth, P. D., Bernard, H. R., & Johnsen, E. C. (2001). Comparing two methods for estimating network size. *Human Organization*, *60*(1), 28–39.

McCormick, T. H., Salganik, M. J., & Zheng, T. (2010). How many people do you know?: Efficiently estimating personal network size. *Journal of the American Statistical Association*, *105*(489), 59–70. http://doi.org/10.1198/jasa.2009.ap08518

McPherson, M. (2009). A Baseline Dynamic Model for Ego Networks. *American Behavioral Scientist*, *52*(12), NP1-NP20. http://doi.org/10.1177/0002764209331530

Mollenhorst, G., Völker, B., & Flap, H. (2011). Shared contexts and triadic closure in core discussion networks. *Social Networks*, *33*(4), 292–302.

Moore, G. (1990). Structural Determinants of Men's and Women's Personal Networks. *American Sociological Review*. http://doi.org/10.2307/2095868
Pachur, T., Schooler, L. J., & Stevens, J. R. (2014). We'll meet again: Revealing distributional and temporal patterns of social contact. *PLoS ONE*, *9*(1), 1–10. http://doi.org/10.1371/journal.pone.0086081
Paniotto, V., Petrenko, T., Kupriyanov, V., & Pakhok, O. (2009). *Estimating the Size of Populations with High Risk for HIV Using the Network Scale-Up Method. Analytical Report, Kiev Institute of Sociology, Ukraine.* Kiev.
Pollet, T. V., Roberts, S. G. B., & Dunbar, R. I. M. (2011). Extraverts have larger social network layers: But do not feel emotionally closer to individuals at any layer. *Journal of Individual Differences*, *32*(3), 161–169. http://doi.org/10.1027/1614-0001/a000048
Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster.
Roberts, S. G. B., Dunbar, R. I. M., Pollet, T. V., & Kuppens, T. (2009). Exploring variation in active network size: Constraints and ego characteristics. *Social Networks*, *31*(2), 138–146. http://doi.org/10.1016/j.socnet.2008.12.002
Roberts, S. G. B., Wilson, R., Fedurek, P., & Dunbar, R. I. M. (2008). Individual differences and personal social network size and structure. *Personality and Individual Differences*, *44*(4), 954–964. http://doi.org/10.1016/J.PAID.2007.10.033
Rosenthal, H. (1960). *Acquaintances and contacts of Franklin Roosevelt*. Cambridge, MA: MIT.
Shati, M., Haghdoost, A., Majdzadeh, R., Mohammad, K., & Mortazavi, S. (2014). Social network size estimation and determinants in Tehran province residents. *Iranian Journal of Public Health*, *43*(8), 1079–90.
Shokoohi, M., Baneshi, M. R., & Haghdoost, A. A. (2010). Estimation of the active network size of kermanian males. *Addiction & Health*, *2*(3–4), 81–8.
Small, M. L. (2009). *Unanticipated Gains: Origins of Network Inequality in Everyday Life*. Oxford: Oxford University Press.
Small, M. L. (2017). *Someone to talk to*. New York: Oxford University Press.
Small, M. L., & Sukhu, C. (2016). Because they were there: Access, deliberation, and the mobilization of networks for support. *Social Networks*, *47*, 73–84. http://doi.org/10.1016/j.socnet.2016.05.002
Smith, E. J., Marcum, C. S., Boessen, A., Almquist, Z. W., Hipp, J. R., Nagle, N. N., & Butts, C. T. (2015). The relationship of age to personal network size, relational multiplexity, and proximity to alters in the Western United States. *The Journals of Gerontology Series B, Psychological Sciences and Social Sciences*, *70*(1), 91–99. http://doi.org/10.1093/geronb/gbu142
Smith, J. A., McPherson, M., & Smith-Lovin, L. (2014). Social Distance in the United States. *American Sociological Review*, *79*(3), 432–456. http://doi.org/10.1177/0003122414531776
Stiller, J., & Dunbar, R. I. M. (2007). Perspective-taking and memory capacity predict social network size. *Social Networks*, *29*(1), 93–104. http://doi.org/10.1016/j.socnet.2006.04.001
Van den Berg, P., & Timmermans, H. (2015). A multi-level path analysis of social networks and social interaction in the neighborhood. *Region*, *2*(1), 55–66.
Van Eijk, G. (2010). Does living in a poor neighbourhood result in network poverty? A study on local networks, locality-based relationships and neighbourhood settings. *Journal of Housing and the Built Environment*, *25*(4), 467–480. http://doi.org/10.1007/s10901-010-9198-1
Van Tilburg, T. G. (1995). Delineation of the social network and differences in network size. In C. P. M. Knipscheer, J. De Jong Gierveld, T. G. Van Tilbrug, & P. A. Dijkstra (Eds.), *Living arrangements and social networks of older adults* (pp. 83–96). Amsterdam: VU University Press.
Van Tubergen, F., Ali Al-Modaf, O., Almosaed, N. F., & Said Al-Ghamdi, M. Ben. (2016). Personal networks in Saudi Arabia: The role of ascribed and achieved characteristics. *Social Networks*, *45*, 45–54. http://doi.org/10.1016/j.socnet.2015.10.007
Wellman, B. (2012). Is Dunbar's number up? *British Journal of Psychology*, *103*(2), 174–176. http://doi.org/10.1111/j.2044-8295.2011.02075.x
Yen, T. J., Fu, Y. C., & Hwang, J. S. (2016). Alters as species: Predicting personal network size

from contact diaries. *Social Networks*, *45*, 78–88. http://doi.org/10.1016/j.socnet.2015.12.001

Zheng, T., Salganik, M. J., & Gelman, A. (2006a). How Many People Do You Know in Prison? *Journal of the American Statistical Association*, *101*(474), 409–423. http://doi.org/10.1198/016214505000001168

Zheng, T., Salganik, M. J., & Gelman, A. (2006b). How Many People Do You Know in Prison? Using overdispersion in count data to estimate social structure in networks. *Journal of the American Statistical Association*, *101*(474), 409–423. http://doi.org/10.1198/016214505000001168

Notes

[1] The data are openly accessible via the website of the Center for Sociological Research, http://www.cis.es/cis/opencm/ES/1_encuestas/estudios/ver.jsp?estudio=14102

[2] Response patterns were deemed inconsistent if the number of relatives and/or friends with a certain name was higher than the total number of people respondents reported to know with the same name, on one or more occasion, with a total difference of three or higher. Smaller inconsistencies (a difference of 1 or 2 names overall) were corrected by adapting the total number of names to that of the friends or family.

[3] The selection of names represents the areas in the center and north of the country slightly better, except Guipúzcoa. Catalonia is well represented among female names, but slightly underrepresented among male names. For the Basque Country, the opposite occurs. Overall, lower representations were found in the smaller provinces Las Palmas and Santa Cruz de Tenerife (the Canary Islands), Ceuta and Melilla (the two Spanish sections on the African continent) and also in the province of Almería.

4 When we replaced the barrier model estimate with the Killworth estimate, age in Model 1 of Table 3 had only a marginal effect (Exp(B) = 0.952; p<.05), and unemployed was significant at p<.001 (Exp(B)=1.217). Effects in Tables 5and 6 were comparable to the presented results.

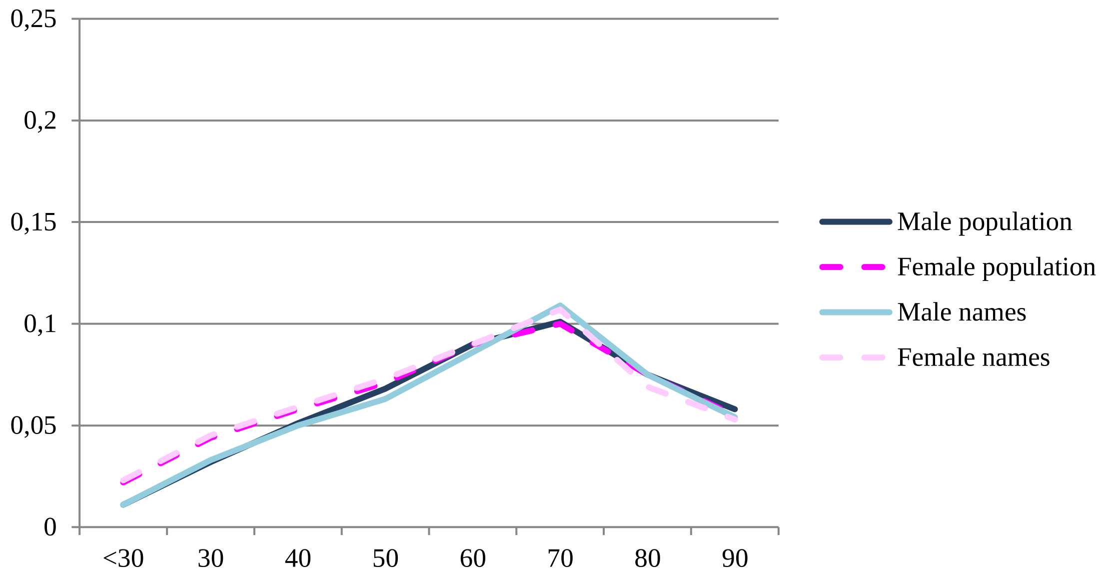

Figure 1. Proportion of the male and female population by birth decade in the total population, as well their representation in the set of 14 names selected for this study.

Note: Proportions per birth decade and gender sum up to 1 separately for the total population and for the subpopulation represented by the 14 names. Adjustment between the proportion of names in the population and the subpopulation represented by the 14 names is complete when the trend lines of "population" and "names" overlap. "<30" = birth year before 1930; "30" = birth year between 1930-'39; "40" = birth year between 1940-'49, etc.

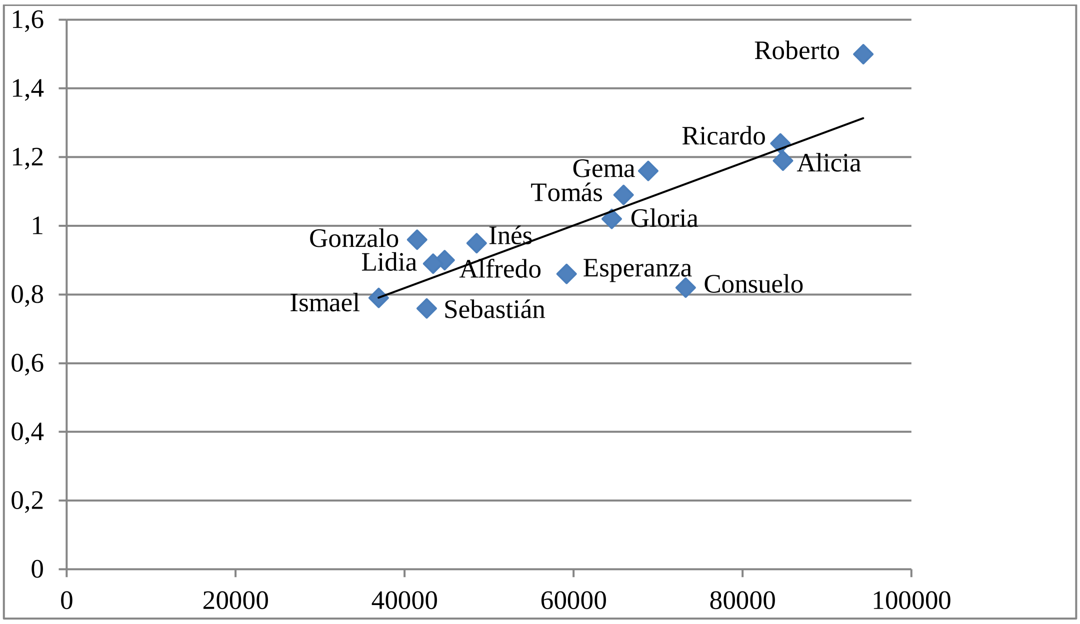

Figure 2. Association between the prevalence of the names in the population (X-axis), according to the data of INE, and the average number of persons with the same name known by the 2,330 respondents (Y-axis).

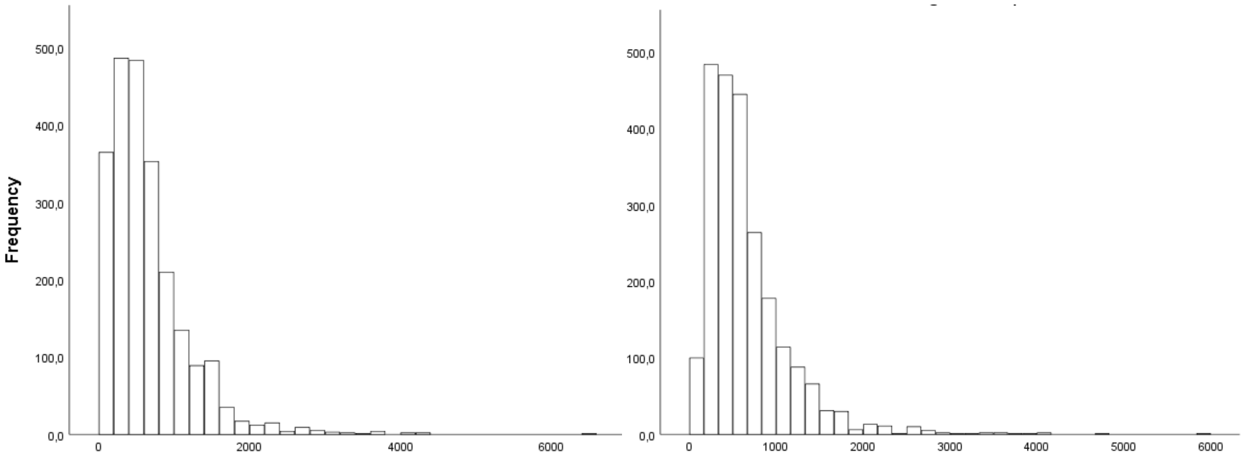

Figure 3. Histogram of the size of acquaintanceship networks as estimated with the method of Killworth et al. (1998; left) and that of the social barrier method of Maltiel et al. (2015, right); *N* = 2,330.

Note: The X-axis represents the size of the networks, the Y-axis the frequency of observation.

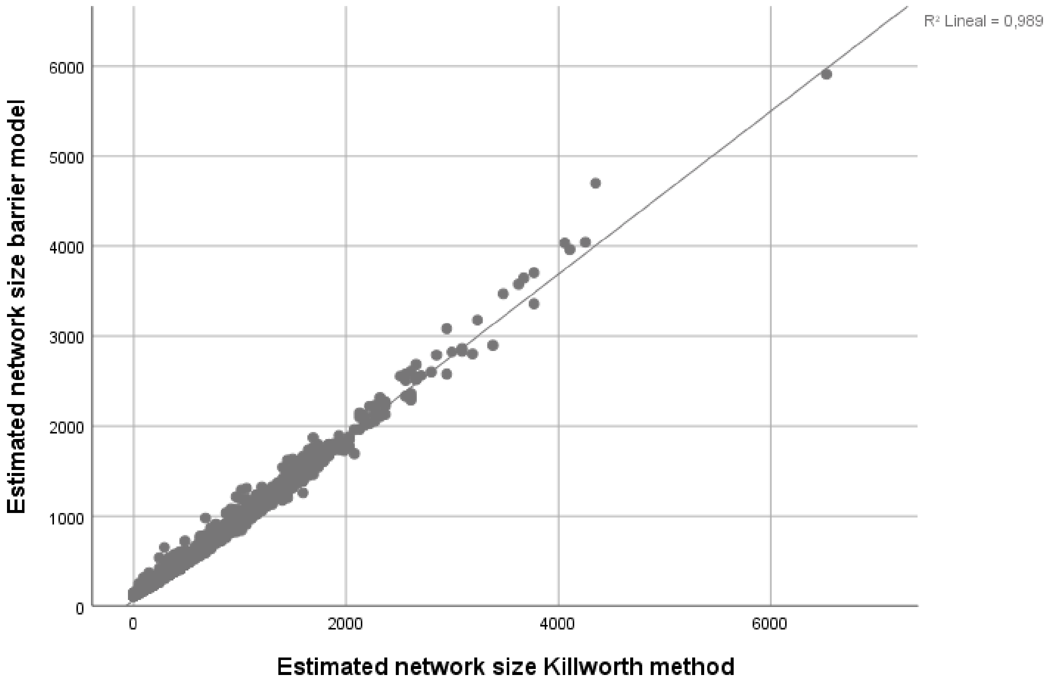

Figure 4. Relation between the estimated acquaintanceship volume of individuals with two estimation methods, that of Killworth et al., 1998 (X-axis), and that of the barrier model of Maltiel et al., 2015 (Y-axis), *N* = 2,330.

Author(s)	**Year pub.**	**Method**	**Sample *N*** *(Random sample of general population indicated in bold)*	**Estimated Acquaintanceship volume** Location	Dispersion	**Network boundary** Definition of relationship measured	**Country**
Rosenthal	1960	Records of President Roosevelt (85 days)	1	1,404		Name noted in the presidential appointment book	US
		Extrap. to 20 yrs		22,500			
Gurevitch	1961	Contact diaries (100 days)	18	*M*=370	Range 83-658	Person contacted one-on-one, either face-to-face, by phone or letter, except first-time contacts	US
De Sola Pool & Kochen	1978	Contact diaries (incl. R1960 + G1961) extrap. to 20 yrs	28	*M*=2,683 Me[*]=2,191	Range 377-6,371	See above (Rosenthal 1960 and Gurevitch 1961)	US
Fu	2007	Contact diaries (3-4 months)	54	*M*=227	Range 55-790	Person contacted one-on-one by any means of communication, whether personally known or not	Taiwan
Yen et al.	*2016*	Contact diaries (F2007) extrap. to 1,000 days		*M*=576	Range ±200 to >1,500		
Huszti et al.	2013	Contact diaries (1 week)	142	*M*=18	*SD*=13; Range 2-93	Person contacted by all means of communication, with contact for at least 5 min, or briefer but important	Hungary
			18	*M*=26	*SD*=10; Range 6-43		
Pachur et al.	2014	Contact diaries (100 days)	40	*M*=77	*SD*=33; Range 26-155	Person contacted face-to-face/by phone for at least 5 min., or electronically/on paper of at least 100 words	Germany
Boissevain	1973	Observation (1 year) and recall	2	1,751 & 638		Person over 14 years old whom R[*] has contact with	Malta
Killworth & Bernard	1978	Experiment (RSW[*];1,267 targets)	58	*M*=210	*SD*=168; Range 43-1,131	Person who could serve as intermediary for the presented list of targets	US
Killworth et al.	1984	Experiment (RSW; 500 targets)	40	*M*=134	*SD*=65	Acquaintance or relative somehow associated with the presented list of targets	US
Bernard & Shelley	1987	Experiment (RSW; 500 targets)	6	*M*=160		Person R feels comfortable asking to deliver message to selected targets/link in chain to list of targets	US
Freeman & Thompson	1989	Experiment (TB[*]; 305 targets), extrap.	247	*M*=5,520	*SE*=271	Person ever known by last name living in Orange County	US
Killworth et al	*1990*	Reestimation FT1989		*M*=2,025			
Bernard et al.	1990	Experiment 1 (RSW/TB)	98	*M*=148	*SD*=69	*RSW: see Killworth et al. 1984; TB: person known in the local area with the last name*	US
		Experiment 2 (RSW/TB)	99	*M*=82	*SD*=83		Mexico
Killworth et al.	*1990*	Exp. 1 B1990 extrap.	98	*M* ~1700	*SE*~400		US
		Exp. 2 B1990 extrap.	99	*M*~600	*SE*~460		Mexico

McCarty et al.	2001	Survey 1 (kpm* and sm*)	**796**	kpm M=291	SD=264	*Not given*	US
				sm M=291	SD=259		
		Survey 2 (kpm and sm)	**574**	kpm M=291	SD=259		US
				sm M=281	SD=255		
Zheng et al.	*2006*	Reestimation McC2001 (Survey 1 and 2; kpm)	**1,370**	M=750; Me=610;	90% range: 250-1,710		US
McCormick et al.	*2010*	Reanalysis McC2001 (Survey 1 and 2; kpm)	**1,370**	M=611; Me=472	Max >6,000		US
Shokoohi et al.	2010	Survey (kpm and sm)	500	kpm M=303	SD=189	Person mutually recognized by sign and name, contacted at least 1x in past year in person, face to face, by phone/email, and could be contacted again	Iran
				sm M=125	SD=284		
DiPrete et al.	2011	Survey (kpm)	**1,371**	Me=550	IQR*: 400-780	Person who R knows by name and for whom (s)he would stop and talk at least for a moment if (s)he ran into the person on the street/shopping mall	US
Shati et al.	2014	Survey (kpm)	829	M=259		Resident of Tehran province who R knows by name and face, who (s)he can visit, call, email when (s)he wants, and vice versa, contacted by phone, email or in person at least 1x in past 2 years	Iran
Hill & Dunbar	2003	Questionnaire (Christmas cards, household size receiving address counted)	43	M=154	SD=85	Person in household to whom R's household sends Christmas card or whom R sees during Christmas (and therefore does not receive card).	UK
Roberts et al.	2009	Questionnaire (Free recall)	160	M=72; Me=70	SD=33; Range 10-168	Known and living relative (genetic and affinal) or other unrelated person for whom R (i) has contact details; (ii) had some sort of contact in past year; (iii) wishes the relationship to continue.	Belgium
Lu et al.	2009	Questionnaire (Free recall)	30	M=52	Range 19-132	Living relative or unrelated individual with whom R maintains genuine relationship	*Not given*
Lewis et al.	2009	Questionnaire (Free recall)	45	*M=37*		Person with whom R has had contact in past 30 days	UK
Ellison et al.	2011	Questionnaire	436	*Me*=300	Max. 1,500	Facebook friend	US
Arnaboldi et al.	2013	Online Social Networks - egonetworks	28	*M=254*	SD=204; range 86-1,099	Facebook friend	Italy
Dunbar et al.	2015	Online Social Networks: Facebook 1	130,338	*M=41*		Facebook friend with at least one interaction per year; Relationship on twitter	-
		Facebook 2	5,761	*M=27*			US
		Twitter	60,790	*M=88*			-

Table 1. Overview of previous work on acquaintanceship volume. *Note* *Me: R: respondent; Median; kpm: known population method; sm: summation method; RSW: Reverse Small World; TB: Telephone Book; IQR: interquartile range; Extrap.: extrapolated. Cursive: (Partial) reanalysis of existing data.

Variable	Category	Univariate descriptive statistics		Bivariate relation with acquaintance-ship volume	
		M	***SD***	**Spearman´s *ρ***	**Sig[a]**
Number of relatives	(Categorical, recoded to midpoints; *N* = 2,326)	17,40	13,41	0.235	**
Number of friends	(Categorical, recoded to midpoints; *N* = 2,316)	5,78	5,56	0.158	**
Age	(Continuous variable, before standardization)	47.9	17.5	-0.151	**
		N	**%**	**Median**	**Sig[b]**
Gender	Women	1,200	51.5	506.0	**
	Men	1,130	48.5	563.0	
Highest education completed	No education	168	7.2	354.5	**
	Primary	326	14.0	463.5	
	Secondary 1st stage	589	25.3	515.0	
	Secondary 2nd stage	258	11.1	576.5	
	Professional educ.	453	19.4	551.0	
	Higher education	527	22.6	609.0	
	NA	*9*	*0.4*	*835.0*	
Monthly net household income	< 600€	230	9.9	453.0	**
	600 - 1,200€	552	23.7	507.5	
	1,200 - 2,400€	572	24.5	559.0	
	> 2,400€	266	11.4	648.5	
	NA	*710*	*30.5*	*509.0*	
Civil status	Married	1,234	53.0	535.5	**
	Single	767	32.9	558.0	
	Widow	158	6.8	399.5	
	Divorced/Separated	170	7.3	521.5	
	NA	*1*	*<0.1*	*289.0*	
Having minor children	No	1,663	71.4	510.0	**
	Yes	661	28.4	603.0	
	NA	*6*	*0.3*	*706.0*	
Employment situation	Inactive	722	31.0	441.5	**
	Unemployed	507	21.8	524.0	
	Employed	975	41.8	593.0	
	Studying	122	5.2	556.5	
	Other / NA	*4*	*0.2*	*894.5*	
Attendance to mass or other religious services	Never or hardly	1835	78.8	530.0	
	Approx. monthly	141	6.1	523.0	
	Approx. weekly	273	11.7	522.0	
	NA	*81*	*3.5*	*559.0*	
Number of types of associations of which respondent is an active member	None	1,635	70.2	502.0	**
	1	464	19.9	600.0	
	2 or more	211	9.1	715.0	
	NA	*20*	*0.9*	*591.5*	

Table 2. Descriptive statistics of the individual attributes and their bivariate relations with acquaintanceship volume (*N*=2,330).

[a]Significance level of Spearman´s rho / [b]Significance of the difference between medians of *k* independent samples (same results including and excluding the category NA or other/NA); ***p<.001*

Variable	Parameter	Acquaintanceship volume (*N*=2,216)							
		Model 1				**Model 2**			
		Exp(B)	**95% Wald CI**		**χ2 Wald**	**Exp(B)**	**95% Wald CI**		**χ2 Wald**
			Lower	Upper			Lower	Upper	
	Intercept	454.862	401.818	514.908		416.441	358.318	483.992	
Age	*(Continuous), std.*	**0.958***	0.928	0.988	7.4*	0.987	0.936	1.040	0.2
Gender (Ref.: Women)	Men	**1.143****	1.085	1.205	24.8**	**1.137****	1.077	1.201	21.4**
Highest education completed (Ref.: None)	Primary	**1.212***	1.075	1.366	28.6**	**1.185***	1.051	1.336	15.6*
	Secondary 1st stage	**1.183***	1.051	1.332		1.112	0.986	1.255	
	Secondary 2nd stage	**1.285****	1.122	1.471		**1.216***	1.061	1.395	
	Professional education	**1.319****	1.165	1.494		**1.216***	1.071	1.380	
	Higher education	**1.349****	1.190	1.528		**1.214***	1.068	1.379	
Household income (Ref: < 600€ /month)	600 - 1,200€ /month	1.058	0.959	1.167	11.3#	1.068	0.966	1.179	3.9
	1,200 - 2,400€ /month	1.090	0.987	1.205		1.069	0.963	1.188	
	> 2,400€ /month	**1.208***	1.071	1.363		1.128	0.991	1.284	
	NA	1.052	0.955	1.159		1.047	0.945	1.159	
Civil status (Ref.: Married)	Single					0.942	0.869	1.022	2.4
	Widow					0.966	0.859	1.087	
	Divorced/separated					0.969	0.871	1.078	
Having minor children (Ref.: No)	Yes					1.020	0.951	1.095	0.3
Employment situation (Ref: Inactive)	Unemployed					**1.190****	1.079	1.312	20.6**
	Employed					**1.212****	1.109	1.326	
	Studying					1.125	0.951	1.330	
Attendance to religious services (Ref: Never/ hardly)	Approx. monthly					1.027	0.922	1.145	2.1
	Approx. weekly					1.065	0.977	1.161	
Membership of associations (Ref: No)	Member of 1					**1.113***	1.041	1.190	33.8**
	Member of 2 or more					**1.301****	1.184	1.431	
		B	**95% Wald CI**			**B**	**95% Wald CI**		
			Lower	Upper			Lower	Upper	
Negative Binomial		0.387	0.366	0.410		0.376	0.356	0.398	

Table 3. Negative binomial regression of acquaintanceship volume on individual attributes

***p<.001; *p<.01; #p<.05*

Support type	Number of people who could provide type of support								*For those with 1 or more persons:* Relation to 1st person in %					
	Median	**IQR**	**Mean**	**SD**	**Frequency in %**			***N***	**Partner**	**First-degree family**	**Other relatives**	**Friend**	**Others**	***N***
					0	**1**	**2 +**							
Help with finding a job*	3	1-6	4.3	3.8	19	8	73	1,502	14	26	7	41	11	1,211
Lending money	2	1-4	3.2	2.8	13	15	72	2,248	12	73	7	7	1	1,975
Help during illness	4	3-7	5.1	3.3	2	9	89	2,300	43	49	3	2	3	2,267
Talk about problems	4	2-6	4.5	3.2	3	12	85	2,305	39	30	3	26	3	2,225

Table 4. Descriptive statistics of the number of people available per type of support

* For help with finding a job, we limited our analysis to people who were employed, unemployed or students.

Variable	Parameter	**To find work** (active population, N = 1,486)				**To lend money** (N = 2,225)			
		Exp(B)	**95% Wald CI**		**χ2 Wald**	**Exp(B)**	**95% Wald CI**		**χ2 Wald**
			Lower	Upper			Lower	Upper	
Model 1	Intercept	0.731	0.300	1.781		0.740	0.467	1.173	
Age	*(Continuous), std.*	**0.584****	0.480	0.710	29.1**	**0.752****	0.641	0.882	12.2**
Gender (Ref.: Women)	Men	1.363#	1.073	1.732	6.4#	0.837	0.686	1.022	3.1
Highest education completed (Ref.: None)	Primary	0.922	0.350	2.434	8.8	1.064	0.714	1.587	25.7**
	Secondary 1st stage	1.155	0.474	2.812		1.577#	1.052	2.363	
	Secondary 2nd stage	0.975	0.386	2.462		1.509	0.938	2.426	
	Professional education	1.190	0.482	2.935		1.805*	1.160	2.810	
	Higher education	1.654	0.668	4.094		**2.591****	1.645	4.080	
Household income (Ref: < 600€ /month)	600 - 1,200€ /month	1.261	0.825	1.928	18.1*	**1.982****	1.414	2.778	34.4**
	1,200 - 2,400€ /month	**1.944***	1.247	3.032		**2.352****	1.642	3.369	
	> 2,400€ /month	**2.777****	1.565	4.930		**3.907****	2.363	6.459	
	NA	1.353	0.880	2.080		**2.323****	1.646	3.278	
Employment situation (Ref: Unemployed for work; Inactive for money)	Unemployed	(Reference)			10.1*	1.124	0.789	1.600	3.6
	Employed	**1.546***	1.174	2.034		1.243	0.892	1.732	
	Studying	1.096	0.625	1.921		1.907	0.928	3.920	
Model 2	Intercept	0.120	0.028	0.513		0.072	0.025	0.202	
Age	*(Continuous), std.*	**0.593****	0.486	0.723	26.5**	**0.756****	0.642	0.889	11.4**
Gender (Ref.: Women)	Men	1.299#	1.018	1.658	4.4#	0.817	0.666	1.001	3.8
Highest education completed (Ref.: None)	Primary	0.870	0.322	2.349	7.0	0.954	0.634	1.437	19.2*
	Secondary 1st stage	1.098	0.441	2.735		1.426	0.943	2.157	
	Secondary 2nd stage	0.850	0.328	2.198		1.285	0.790	2.089	
	Professional education	1.094	0.433	2.762		1.589#	1.011	2.498	
	Higher education	1.417	0.559	3.596		**2.132***	1.338	3.398	
Household income (Ref: < 600€ /month)	600 - 1,200€ /month	1.258	0.816	1.937	14.9*	**1.927****	1.366	2.717	27.9**
	1,200 - 2,400€ /month	**1.871***	1.191	2.939		**2.146****	1.489	3.095	
	> 2,400€ /month	**2.540***	1.418	4.551		**3.364****	2.020	5.602	
	NA	1.322	0.854	2.049		**2.263****	1.593	3.214	
Employment situation (Ref: Unemployed for work; Inactive for money)	Unemployed	(Reference)			9.6*	1.050	0.733	1.503	3.4
	Employed	**1.545***	1.170	2.042		1.138	0.811	1.596	
	Studying	1.135	0.645	1.998		1.891	0.914	3.912	
N relatives	Number of relatives	1.010#	1.000	1.021	4.0#	**1.016****	1.008	1.025	13.3**
N friends	Number of friends	**1.050****	1.022	1.078	12.9**	**1.036***	1.014	1.058	10.4*
Log acq. volume	Log acq. volume	1.269#	1.044	1.543	5.7#	**1.405****	1.196	1.651	17.1**

Table 5. Logistic regression of the probability of having multiple persons who could provide the respondent with social support. ***p<.001; *p<.01; #p<.05*

Variable	Parameter	To receive care during illness (*N*=2,276)				To talk about problems (*N*=2,278)			
		Exp(B)	**95% Wald CI**		**χ2 Wald**	**Exp(B)**	**95% Wald CI**		**χ2 Wald**
			Lower	Upper			Lower	Upper	
Model 1	Intercept	4.183	2.327	7.521		2.451	1.461	4.112	
Age	*(Continuous), std.*	0.846	0.682	1.050	2.3	0.855	0.708	1.033	2.6
Gender (Ref.: Women)	Men	0.856	0.653	1.122	1.3	**0.665****	0.524	0.842	14.4**
Highest education completed (Ref.: None)	Primary	0.959	0.583	1.575	12.5#	1.279	0.807	2.026	9.6
	Secondary 1st stage	1.428	0.841	2.425		1.294	0.810	2.067	
	Secondary 2nd stage	1.674	0.861	3.255		1.188	0.681	2.071	
	Professional education	1.276	0.720	2.259		1.182	0.713	1.959	
	Higher education	**2.302***	1.238	4.280		2.001#	1.173	3.415	
Household income (Ref: < 600€ /month)	600 - 1,200€ /month	1.418	0.913	2.201	13.1#	**1.883***	1.279	2.771	23.5**
	1,200 - 2,400€ /month	**2.285***	1.377	3.789		**2.840****	1.841	4.379	
	> 2,400€ /month	2.074#	1.059	4.060		**2.273***	1.311	3.939	
	NA	1.285	0.823	2.005		**1.711***	1.159	2.525	
Employment situation (Ref: Inactive)	Unemployed	0.976	0.600	1.586	3.3	1.202	0.785	1.841	3.9
	Employed	0.955	0.605	1.508		1.085	0.732	1.608	
	Studying	2.507	0.774	8.113		2.314	0.942	5.681	
Model 2	Intercept	1.021	0.260	4.003		0.218	0.065	0.728	
Age	*(Continuous), std.*	0.860	0.691	1.070	1.8	0.884	0.731	1.071	1.6
Gender (Ref.: Women)	Men	0.859	0.652	1.132	1.2	**0.629****	0.494	0.801	14.1**
Highest education completed (Ref.: None)	Primary	0.847	0.509	1.408	8.2	1.135	0.710	1.816	5.6
	Secondary 1st stage	1.258	0.731	2.164		1.134	0.703	1.831	
	Secondary 2nd stage	1.399	0.708	2.761		0.945	0.535	1.670	
	Professional education	1.068	0.593	1.925		0.963	0.573	1.618	
	Higher education	1.728	0.912	3.273		1.483	0.856	2.570	
Household income (Ref: < 600€ /month)	600 - 1,200€ /month	1.287	0.821	2.018	8.2	**1.812***	1.221	2.687	18.7**
	1,200 - 2,400€ /month	1.943#	1.159	3.258		**2.601****	1.672	4.047	
	> 2,400€ /month	1.622	0.817	3.220		1.966#	1.121	3.448	
	NA	1.178	0.748	1.854		1.658#	1.115	2.465	
Employment situation (Ref: Inactive)	Unemployed	0.902	0.551	1.477	4.1	1.137	0.739	1.749	4.4
	Employed	0.870	0.545	1.389		1.005	0.672	1.501	
	Studying	2.426	0.748	7.873		2.295	0.930	5.659	
N relatives	Number of relatives	**1.031****	1.017	1.046	18.5**	1.007	0.997	1.018	2.0
N friends	Number of friends	**1.109****	1.063	1.156	23.3**	**1.090****	1.056	1.126	28.2**
Log acq. volume	Log acq. volume	1.138	0.916	1.415	1.4	**1.420****	1.172	1.721	12.8**

Table 6. Logistic regression of the probability of having multiple persons who could provide the respondent with social support. ***p<.001; *p<.01; #p<.05*