Journal of Transport Geography 31 (2013) 236–248

Contents lists available at SciVerse ScienceDirect

Journal of Transport Geography

journal homepage: www.elsevier.com/locate/jtrangeo

Distance patterns of personal networks in four countries: a comparative study

Matthias Kowald [a], Pauline van den Berg [b], Andreas Frei [a], Juan-Antonio Carrasco [c,*], Theo Arentze [b], Kay Axhausen [a], Diana Mok [d], Harry Timmermans [b], Barry Wellman [e]

[a] Institute for Transport Planning and Systems (IVT), ETH Zurich, 8093 Zurich, Switzerland
[b] Faculty of the Built Environment, Urban Planning Group, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
[c] Department of Civil Engineering, Universidad de Concepción, P.O. Box 160-C, Concepción, Chile
[d] Department of Geography, University of Western Ontario, Social Science Centre, ON London, ON N6A 5C2, Canada
[e] Department of Sociology, University of Toronto, 725 Spadina Avenue, Toronto, ON M5S 2J4, Canada

ARTICLE INFO

Keywords:
Social networks
Distance between social contacts
Comparative studies
Social activity
Travel behaviour
Data collection

ABSTRACT

Acknowledging the relevance of social networks on (social) travel behaviour, the objective of this paper is to comparatively study the distance patterns between the home locations of social contacts. Analyses are based on five recent collections of personal network data from four countries: Canada, Switzerland, the Netherlands, and Chile. Multilevel models, which explicitly account for the hierarchical structure of the data sets, are used to study the role of explanatory variables to understand the distance patterns of social contacts. Modelling results suggest that alters' characteristics (such as type of relationship, emotional closeness, and duration of the relationship) as well as personal network composition (alters with a certain relationship to the ego) constitute stronger predictors than an ego's socio-demographic information across these countries. In addition, comparative analyses suggest differences between countries on relevant key variables such as an ego's income and the ego–alter tie strength.

1. Introduction

In transportation research, traveling for social activities has received far less attention than traveling for other purposes, such as working or shopping. However, there has recently been recognition of the importance of social activity-travel, as such excursions not only account for a large number of trips, but constitute the fastest growing segment of travel (Axhausen, 2005).

A key determinant of social travel, influencing its spatial and temporal patterns, is the individual's social network. In fact, social contexts are often the main drivers to perform a social activity at a certain time and location. Location choice for social activities depends to a large extent on the home locations of the social network members performing the activity together (Carrasco et al., 2008a). Therefore, detailed information regarding the distance patterns and home locations of people's personal network members is a crucial element in the understanding of social travel.

Acknowledging the previously ignored importance of social networks on travel behaviour, data collection efforts have been undertaken in recent years in Canada (Hogan et al., 2007; Carrasco et al., 2008b), Switzerland (Ohnmacht and Axhausen, 2005; Frei and Axhausen, 2007; Kowald and Axhausen, 2012), the Netherlands (van den Berg et al., 2009), and Chile (Carrasco and Cid-Aguayo, 2012). These studies have highlighted the relevance that different aspects of an individuals' personal network have on their activity and travel behaviour.

Previous analyses from each of these data sets have demonstrated the role of socio-demographics and personal network characteristics on the distance patterns between social contacts. However, there is a need to disentangle the differences between these data sets, since a plausible hypothesis is that these spatial patterns are influenced by the city or national context where they are embedded. In fact, the bulk of reported contacts are local/regional, as the cost structure for local and regional travel is comparable, but not equal across each of the countries in question. Similarly, given the relative affordability of transport, as well as the availability of mobility tools (e.g., car ownership and Internet), one would expect the mean distances to be higher in Switzerland, Canada and the Netherlands than in Chile. In addition, the share of foreign nationals will influence the proportion of long distance and international contacts. Finally, factors related to socio-cultural contexts can also play a role in creating the different spatial patterns in networks; although these factors have not been incorporated in the data collection efforts that support this research.

* Corresponding author. Tel.: +56 41 2203603.
E-mail address: j.carrasco@udec.cl (J.-A. Carrasco).

0966-6923/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jtrangeo.2013.06.006

With this motivation, expectation, and scope, the objective of this paper is to comparatively study the observed distance patterns between the home locations of social contacts, or ego–alter pairs, in the different countries included.

The study is facilitated by the common personal network approach of these five data sets, which focuses on specific individuals (egos) and their social contacts (alters). In this approach, two levels of explanatory variables can be distinguished: ego-network and ego–alter. Ego-network attributes include the characteristics of the respondents' overall personal network and their socio-demographic characteristics. Ego–alter attributes include tie relationships, interaction patterns, and an alter's demographic characteristics. A third level can be added to account for the variability introduced by the different national contexts under study.

Multilevel models, which explicitly account for the personal network structure of the data sets, are used to study the role of explanatory variables to understand the distance between ego–alter pairs.

The paper is structured as follows. In the next section, we review the existing literature on social activity-travel and personal networks, with a main focus on geographical distance patterns. In Section 3 the five data collection efforts are described, followed by a description of the data in Section 4. Next, the methods of analysis as well as the main empirical results are presented, showing the relationship between individual characteristics, their personal networks, and the spatial structure of their contacts. Finally, the key conclusions are summarized.

2. Theoretical framework

2.1. Social networks in transportation research

The study of social networks has only recently been recognized as a research frontier in transportation (e.g. Dugundji et al., 2011). Two lines of research into the effect of social networks on travel behaviour have emerged recently. The first line of research focuses on the way people's social networks influence their travel decisions by the exchange of information and opinions (Dugundji and Walker, 2005; Páez and Scott, 2007; Páez et al., 2008; Schwanen, 2008; Dugundji and Gulyás, 2008), underlining that social networks, especially decisions of alters in the network, have an influence on people's decisions regarding trip destination, route, frequency, and mode. Agent-based simulations of social influence in transportation have been reported by Marchal and Nagel (2005), Hackney and Axhausen (2006), Ettema et al. (2011), Han et al. (2011), as well as Frei and Axhausen (2011).

The second line of research into the impact of social networks on travel behaviour focuses on their effect in generating immediate activity. In recent years, several data collections have been carried out in which respondents were asked to report on selected social network members (Larsen et al., 2006; Hogan et al., 2007; Frei and Axhausen, 2007; van den Berg et al., 2009; Carrasco and Cid-Aguayo, 2012; Kowald and Axhausen, 2012). Empirical analyses of these data sets have mainly focused on two aspects of social travel, namely travel frequency, as indicated by the frequency of face-to-face contact (in relation to contact frequency by ICT-mediated communication modes), and travel distance (or distance between the homes of the pair involved). Separate analyses have demonstrated that socio-demographics, personal network characteristics, and especially characteristics of the ego–alter link, play an important role in explaining the distance patterns between social contacts. However, as these analyses were based on data sets from different cultural contexts, there is a need to disentangle these differences in a comparative analysis of social network distance patterns.

2.2. Personal network approach

Although the study of social networks is relatively new to transportation research, it is rooted in a long tradition in the social sciences (e.g., Wasserman and Faust, 1994; Marsden, 1987, 2005; Degenne and Forsé, 1999). Drawing on this work, the datasets used in this paper have adopted the personal network approach, focusing on specific individuals (egos) and their social contacts (alters).

To define the boundaries of a network, a set of name generating questions is usually employed. Different name generating approaches can be distinguished according to the kinds of social contacts they capture (Van der Poel, 1993). For instance, the interaction approach elicits a record of all alters with whom the respondent (ego) interacts in a certain period (including casual and unknown contacts). Another approach is the role relation approach where a record of people with whom the individual has a certain role relationship, such as immediate family, relatives, neighbours or friends, is produced. Thirdly, the affective approach asks respondents to record the people with whom they have a close personal relationship, or who are especially important to them. Finally, another approach concentrates on people with whom the individual exchanges social support, both emotional and material. The choice between the different approaches depends on the aim of the study.

Once all alters are elicited, additional questions, called name interpreters, are used to gather information about the characteristics of the alter and the ego–alter relationship (tie).

However, using these methods to explore personal networks only captures realized patterns, that is, they only account for the social contacts and interactions in time and space that actually occur, without explicit consideration of the constraints or dynamics of these social processes. Similarly, the static nature of the personal network approach employed in this paper is unable to incorporate potentially important aspects when addressing the spatiality of interactions, such as the dynamics of relationships, power struggles, social influence, and other dynamic social processes.

2.3. Social activities and travel

The importance and need to study social activity and travel behaviour to improve models of travel demand has recently been emphasized because social activities account for a large portion of trips and constitute the fastest growing segment of travel (Axhausen, 2005). In addition, social activities and travel are important aspects of an individual's quality of life as interaction with other people provides access to a variety of resources, such as instrumental and emotional support (social capital), which is becoming a key topic of discussion from transportation policy point of view (Miller, 2006; Carrasco and Cid-Aguayo, 2012).

The relevance of social activity-travel is indicated by the fact that people's leisure time has increased over the last decades, as well as the rate of car ownership and usage (Schlich et al., 2004; Larsen et al., 2006). This tendency is expected to increase with an aging population (e.g. Banister and Bowling, 2004; Newbold et al., 2005). Moreover, travel distances for social activities are growing as people's social networks are spread over larger geographical areas than before (Schlich et al., 2004; Larsen et al., 2006; McPherson et al., 2006). The resulting travel demand for social activities is increasing in parallel.

Ettema and Schwanen (2012) argue that the analysis of leisure mobility, and thus social activities, has tended to ignore the joint character and dependence of other people on its decision making process. In their account, they recognize three recent approaches that – although with limitations – have given insights on these kinds of activities. The first approach corresponds to lifestyle approaches, including attitudes, values, and orientations (e.g., Ohnmacht et al., 2009). The second incorporates mental models, including learning and subjective perceptions (e.g., Arentze et al., 2008). The third approach entails social network structures, which enables an understanding of these activities by incorporating aspects of the individual's social context. Although the current state of the art in social network analysis is still limited by incorporating the role of the socio-cultural context factors in the decision making process, it constitutes a productive approach in overcoming the traditional, individualistic explanations of travel behaviour surrounding social activities.

2.4. The spatiality of social activities and social networks

Although social activities account for a large portion of travel, little is known about the factors that influence the spatial patterns of social activity-travel. The main premise in this work is that social activity-travel patterns emerge from individuals' social networks. Social activity space, or the set of potential locations to perform social activities (Horton and Reynolds, 1971), is defined, to a large extent, by an individual's social network, either directly (social network members' homes), or indirectly (pubs or restaurants close to the network members' homes or workplaces). Therefore, detailed information on the spatial distribution of people's personal networks is crucial to an understanding of social activity-travel.

However, studying the spatial dimension of social networks is not trivial. In fact, space is socially produced, and cultural differences play a relevant role. Therefore, studying the spatial patterns and distances between people needs to take into account the potentially high differentiation of socio-cultural areas, territories, and place-based communities, in conjunction with the material consequences for access to opportunities (Hanson, 1998). These non-linearities and discontinuities of space cannot be studied directly using the data and methodology employed in this study, but need to be borne in mind when interpreting the results of the analysis.

Another key dimension is location. In fact, in contrast to other activities, the locations of social purposes will not only depend on the accessibility and other traditionally measured attributes in the transport literature, but also on their functionality and affective relevance (Ettema and Schwanen, 2012). However, given that the personal network members considered here have a relatively strong emotional closeness, a plausible simplification consists of concentrating on homes as the key social location of activities.

In terms of developing empirical indicators of the distance patterns of social networks, this paper follows the current state of the art, using great circle distance indicators as dependent variables, and more specifically, making use of the logarithm of distances between homes. Although, as discussed before, there are limitations to this metric, logarithmic distance is a reasonable approach when working with a very wide range of distances from neighbourhoods to international distances (Latané et al., 2005; Mok et al., 2010). In fact, the logarithmic distance mimics the effect that travel speed generally grows with distance, as local modes are replaced by cars and trains and, at even greater distances by high-speed modes such as planes or high-speed trains, and for this reason, it is a usual metric when analysing distances in urban contexts (Schwanen and Mokhtarian, 2005; Daly, 2010; Naess, 2011).

There are alternatives to metric distance, such as the shortest free flow travel time, shortest congested travel time, shortest congested intermodal travel time, generalized costs of travel (depending on mobility tool ownership), and log-sums of a relevant mode choice/mobility tool ownership model. Yet, it is not clear to what extent this extra complexity generates better approximations of the subjectively perceived distance between social contacts, considering all other potential sources of bias, both in the data collection and modelling processes. Nevertheless, different studies have shown, that crow-fly distance is an appropriate measure for travel distance as it is highly correlated with network distance metrics (Rietveld et al., 1999) and deviates only substantially in short ranges (Chalasani et al., 2005).

In sum, and despite its theoretical limitations, the great circle indicator remains on balance the preferred variable, given the costs of the alternatives and their yet unknown quality in this research context. The detailed and unique nature of personal network data sets allow us to control for the socio-demographics and key aspects of an individual's social circumstances, such as the composition and size of their closer social contacts, thus overcoming the traditional, socially isolated approach of the study of social activities.

3. Data collection

Five datasets from four different countries are described below to compare the distance patterns of personal social networks.

3.1. Toronto

The data were collected In the East York area of Toronto, Canada between May 2004 and April 2005 as part of the ''Connected Lives Study'', a study about people's communication patterns. The study consisted of two stages: surveys of a random sample of 350 people from the East York area in Toronto, and interviews and observations of a sub-sample of them.

In these interviews, respondents were asked to name the persons who live outside their household, with whom they felt very close and somewhat close. Very close people consist of those persons with whom the respondent discusses important matters or regularly keep in touch with, or are there for them if they need help. Somewhat close people were described as persons who are more than just casual acquaintances, but not considered to be very close.

In total, 87 respondents completed the detailed questionnaire and named 1019 alters, which is an average degree of 11.7. Several attributes were collected for each alter, including home location, and frequency of contact by mode; more details can be found in Hogan et al. (2007) and Carrasco et al. (2008b).

3.2. Zurich

Between December 2005 and December 2006 social network data were collected in Zurich, Switzerland. Based on a random sample of the Zurich population, the participants were recruited on the telephone. The survey itself contained two parts. First, a written questionnaire was filled in by the respondents independently containing socio-demographic and travel related questions. Second, a face-to-face interview was conducted to collect the social network data.

Respondents were asked to name alters with whom they discuss important problems, with whom they stay in regular contact or whom they can ask for help. These questions cover the ''very close'' or ''most important'' contacts. A second name generator asked for persons with whom the respondents plan and spend leisure time. Both of these name generators targeted specifically social contacts influencing social activity behaviour.

In total, 307 respondents completed the questionnaire and named 3807 alters, which is an average degree of 12.4. Similar to Toronto, several alter attributes were collected among the exact home address location, and the contact frequency by mode. For more information, see Frei and Axhausen (2007).

3.3. Eindhoven

Between January and June 2008, social network data were collected in a number of neighbourhoods in the Dutch Eindhoven region. Eindhoven is a mid-sized city in the south of the Netherlands, with a population of 216,000.

The data were collected as part of a larger study which consisted of a 2-day social interaction diary (including a questionnaire on personal socio-demographic characteristics) and a follow-up questionnaire to capture the respondents' social network.

The participants of this social network study are a subset of the respondents who participated in the larger social interaction study. The social interaction study involved 747 respondents of which a subsample of 116 respondents completed the social network questionnaire. In this study a paper and pencil questionnaire was used in which respondents could self-report member of their social network members. The name generating questions used were similar to those in the Connected Lives Study.

Respondents could record up to 25 very close and 40 somewhat close social network members. Among 116 respondents a total of 2695 social network members were reported, which is an average network size of 23.28. For more information, see van den Berg et al. (2009).

3.4. Concepción

The study "Communities in Concepción" focused on the characteristics of social activity travel through the analysis of personal networks in different neighbourhoods of Concepción, Chile. The city is located 500 km south from Chile's capital, Santiago, and the Greater Concepcion Area has a population of around one million people, being the second largest city in the country. Data about personal networks were collected in four distinctive neighbourhoods as a way of capturing diverse income and accessibility to the CBD levels.

The data collection effort took place between August 2008 and April 2009. The data were collected in semi guided interviews with 240 people (60 from each neighbourhood), which elicited a total of 5053 personal networks members, an average network size of 22.24. Respondents were chosen by a random and socio-demographic quota based procedure. The study used the same name generators as in the Connected Lives Study, and included the networks spatial location, frequency of interaction, social support exchange, and a 2-day retrospective activity-travel survey. More details about this dataset can be found in Carrasco et al. (2013).

3.5. Switzerland

Snowball sampling was used to collect data on personal networks between January 2009 and March 2011. A stratified random sample of the Canton Zurich population was used to recruit 40 initial respondents. Two name generators for leisure and emotionally important contacts were used, asking respondents to mention people with whom they make plans to spend free time and those with whom they discuss important problems. The questionnaire provided space for 40 names and encouraged respondents to write down additional names on an extra sheet of paper if needed.

The snowball sampling method permitted the use all contacts mentioned in response to the name generator as the basis for further recruitments. In other words, these persons were asked to fill out the paper-based questionnaire as well. Repeating this process on five iteration levels resulted in a sample of connected personal networks containing information on 743 egos and 15,593 alters. Although recruitment efforts were taken to an international level, respondents and their social contacts are highly clustered in Switzerland. The data collected includes important information on a population wide leisure network structure, and an activity travel diary of 8 consecutive days. Besides questions on transport modes and types of activities, the instrument focused on accompanying people, trying to re-identify persons mentioned in the name generator of the questionnaire (Kowald and Axhausen, 2012).

4. Data description

A challenge during in the analysis of the datasets consisted of identifying the common variables available in all five datasets and adjusting common categories accordingly. In this process, there were some survey specific characteristics that needed special treatment. First, the Switzerland data have a large number of missing values in terms of distances between respondents' and their social contacts' home locations as the result of the snowball sampling strategy (17%). Respondents refused mentioning their postal addresses when they did not want their social contacts to become part of the sample. Second, distances were surveyed on an ordinal scale in Eindhoven. Only distances above 200 km were recorded on a metric scale. However, in order to model distances on a metrical scale, all values within the distance classes were replaced by the average within the class values. Third, distances were geo-coded in Toronto, Zurich, Switzerland, and Concepción using egos' and alters' addresses or in cases where the address was not available (7% in Zurich) the next best guess from the respondent was used (street-corner, closest landmark, etc.). By contrast, the Eindhoven dataset includes distances as reported by respondents. This practical issue limits the comparison of the spatial structure of the ego's social network to crow fly distances for each ego–alter pair. Alternatives that take the clustering effect of alters into account (see Schönfelder, 2006 for a comparison of different metrics; see Carrasco et al., 2006 and Frei and Axhausen, 2007 for the analysis of ego's network's spatial representation with confident ellipses) are not available in this comparison, but have been shown to correlate strongly with crow-fly distance (see Frei and Axhausen, 2007). The available data also does not account for cultural clustering or population density. In addition, clustering is difficult to integrate in the statistical model used in the paper since it follows the data structure with independent observations within the multilevel structure.

For the purpose of the analysis, observations with missing values were excluded from the datasets. A caveat of this approach is that the remaining observations may not be entirely representative within each dataset. However, as the comparative analysis includes regression models, employing an imputation method such as imputation based on an unconditional or a conditional distribution would have changed the dependency structure within a dataset. Furthermore, a multiple regression model would have resulted in confidence intervals indicating the range of the estimated parameters, resulting in more complex comparisons because of the uncertainty from the imputation. As a consequence, simply excluding missing values seems to be a more appropriate approach for the purposes of this study.

The descriptive analysis provides a comparative overview of the datasets, highlighting the dissimilar target populations in the four countries—as well as the differences the name generators and information collected in several units of analyses—all of which will have some impact on the results of the model.

Table 1 gives an overview of general statistics of each survey location. All datasets were surveyed in urban areas, except for the Switzerland study, which collected information at the national level. National statistics from the three developed countries show small differences in wage level and transportation costs, especially if their ratios are compared. The Chilean statistics for the Concepción case are different, as the wage level is much lower, and the

Table 1
Country specific statistics (2006). *Source*: UBS, Wealth Management Research (2006).

	Toronto	Zurich	Eindhoven	Concepción	Switzerland
Population (in 1000)	6054	372	213	292	7866
Population density	3972	4049	2407	1318	188
Wage level (Base: New York) Gross	74.2	115.1	77.0	21.2	115.1
Wage level (Base: New York) Net	80.4	124.2	72.7	24.3	124.2
Bus/Tram/Metro (Network ticket for a trip 10 km)	2.4	2.7	2.6	0.7	2.7
Taxi (per 5 km)	8.2	21.2	17.2	0.7	21.2
Train (200 km single ticket)	45.4	44.8	31.2	11.7	44.8
Ave					
Average cost of fuel per litre (US $)	0.89	1.22	1.72	1.06	1.22
Average Mid-Price Car (US $)	19,933	22,240	21,140	11,416	22,240
Tax on Car (USD/Year)	64	255	289	210	255

lower costs for individual transport do not cancel out the differences in wage. Finally, it should be noted that the Swiss national numbers shown in Table 1 may understate the incomes in Zürich.

Table 2 provides a comparison between egos' characteristics. The different datasets are quite similar and show only few major differences. The Concepción survey sample is younger than those of the other surveys as a result of the high share of participants in the youngest age class (<30 years). This is related to the phenomenon of the demographic transition of decreasing birth rates and an older overall population in most developed western countries, which has not yet been observed in Chile. In terms of household structure, the Zurich data show a nearly perfect split between respondents living with or without partner whilst there is a couple dominance in all other data sets. Another difference is the presence of children in the respondents' households. Households with young children represent a third in Eindhoven and Switzerland, half of the sample in Toronto, and two thirds in Concepción. In terms of the educational level, the four data sets from the developed countries show lowest shares for primary education, and – with the exception of Zurich – highest shares for educational levels with academic degrees. The income distribution – defined in terms of low, medium and high income categories within each country – follows different patterns in all data sets. The survey population from Concepción has more people in the lowest third income class, whilst the Switzerland study is dominated by respondents in the highest third income rank. Most households in the western world own a car, which is also true for Concepción, even though the dominance of car ownership in this latter case is less strong. No matter where respondents' households are located, most of them have Internet access. In addition, most respondents have a mobile phone, with the exception of the Toronto data set, which is also the oldest of the five. In terms of years living at the current location, the Switzerland data show the highest average.

Table 3 shows a synthesis of the characteristics of the social contacts in each data set. In terms of personal network sizes, the mean value is very similar among all data sets, with the exception of Zurich, where the value is much smaller. This similarity may be in part due to the same kind of name generators used by all data sets in terms of emotional closeness, although the Zürich and Switzerland datasets also added leisure contacts as another criterion. Another hypothesis about this similarity could be that there may be a "cognitive" threshold that causes the similar network sizes; however, the results from Zürich complicate this idea. In addition, the difference between the Zurich and Switzerland datasets could also be due to the fact that the name generator asking for leisure contacts is more detailed in the latter survey, supporting the idea that name generators are very sensitive to (even small) changes in wording. The previous mixed evidence about network sizes suggests the need of more research to better understand these issues.

The alters' age distribution shows similar patterns to the egos' age. Whilst a comparison between Toronto, Eindhoven, and Switzerland fits well, the share of social contacts in the youngest age

Table 2
Socio-demographic characteristics of the respondents.

	Toronto	Zurich	Eindhoven	Concepción	Switzerland
Number of respondents	84	265	106	241	426
	Sample (%)	Sample (%)	Sample (%)	Sample (%)	Sample (%)
Male	39.8	42.3	31.1	39.8	41.1
Young (<30)	9.8	19.6	7.5	24.1	5.9
Middle (30–60)	69.5	44.2	59.4	58.5	73.5
Old (>60)	20.7	36.2	33.0	17.4	20.6
Living with partner	61.9	48.3	72.6	58.1	80.3
Child(ren) under 18	46.4	–	34.0	61.8	36.6
Primary education	18.3	7.9	17.0	44.8	1.6
Secondary education	28.0	70.2	34.0	24.5	48.6
Higher education	53.7	21.9	49.0	30.7	49.8
Low HH-income	29.4	24.5	36.8	43.8	10.1
Medium HH-income	41.2	47.9	25.5	25.1	39.2
High HH-income	29.4	27.5	37.7	31.1	50.7
1 or more cars	–	63.4	83.0	56.4	89.0
Season ticket	–	38.9	42.5	–	82.9
Internet access	79.8	67.9	90.6	63.9	97.9
Mobile phone access	42.3	65.7	94.3	86.3	95.8
	Mean	Mean	Mean	Mean	Mean
Age	50.3	50.7	51.6	42.8	50.1
Work hours	22.5	–	14.6	21.8	–
Years in current location	12.8	–	13.3	16.4	25.4

Table 3
Characteristics of personal networks.

	Toronto	Zurich	Eindhoven	Concepción	Switzerland
Number of alters	1019	3156	2452	5038	7293
Average network size	23.8	11.9	23.9	20.9	21.6
	Sample (%)	Sample (%)	Sample (%)	Sample (%)	Sample (%)
Male	42.1	–	41.4	45.9	43.1
Young (<30)	11.8	–	13.0	32.2	9.2
Middle (30–60)	65.3	–	53.7	52.9	67.2
Old (>60)	22.9	–	33.3	14.9	23.6
Immediate family	25.4	18.3	18.8	20.2	13.4
Extended family	11.0	12.7	25.1	23.6	11.1
Friend or other	63.6	69.0	56.1	56.2	75.5
Very close	54.0	52.4	43.0	51.3	28.9
Somewhat close	46.0	47.6	57.0	48.7	71.1
Known <1 year	–	0.6	1.4	14.6	2.0
1–10 years	–	36.7	37.4	25.8	28.5
>10 years	–	62.7	61.2	59.6	69.5

class for the Concepción study is higher than the egos. Since the Zürich surveys did not include questions on alters' sex and age, this dataset cannot be used to calculate corresponding homophily values and employ them as dependent variables in the models.

In terms of the ego–alter relationships, Switzerland and Zurich include a high share of friends and only a low share of family members. In the case of Eindhoven and Concepción, the friendship category is dominant as well, but the distribution is more balanced with respect to family members. Toronto is placed somewhere between these two patterns. These differences can be, at least partly, due to the different name generators employed.

Regarding emotional closeness, the share of very close contacts is very similar across datasets, with the exception of Switzerland, which has a much lower share, possibly due to its stronger orientation towards leisure contacts.

Finally, in terms of the duration of the relationship, all data sets follow different patterns. A particular noteworthy case is Concepción, where the relations are much younger, which could be explained in part by the age distribution in this dataset.

Fig. 1 and Tables 4 and 5 show the general picture of the spatial distribution of the different personal networks. The geo-referenced values of the respondent's home and their social contacts were used to calculate great circle distances, employing an equidistant cylindrical projection to account for the shape of the Earth. Eindhoven is excluded in this analysis since their reported distance categories were used. The distribution of the great circle distances (Fig. 1c) between the respondents' residence and their alters has three elements. Around two-thirds of the alters live locally within 30 km (roughly 30 min to 1 h travel time by car) (Fig. 1a). The remaining distances are roughly equally divided into regional or national relationships (within 30–100 km) (Fig. 1a) and longer distance and international relationships (>100 km) (Fig. 1b). The 100 km threshold was chosen to mark the difference of long-distance travel and everyday travel, as this is a common cut-off point in the transportation planning literature. Overall, the respondents mix their daily life local/regional contacts with a multitude of non-local and often long distance contacts. In fact, there is also a noticeably high share of intercontinental links, especially in the case of Toronto, due to immigrants maintaining relationships at their birthplace.

In the spatial range of everyday life (less than 100 km), the overall differences in the distance distributions between the datasets are rather small. For longer distances, there are no visible trends except for the peaks in the longer distances for Toronto. Fig. 1d present in more detail the tie distance distribution within 100 km, showing the distance cuts and their shares on a log–log scale. The tie length distribution follows a power law distribution for all four datasets, defined by $p_{tie} \sim distance^{\beta}$, where β varies between -1.08 (Toronto) and -1.58 (Concepción). The order of the datasets is expected, given their settings. In fact, transportation costs compared to the wage levels are similar in Zurich, Switzerland and Toronto; also this latter dataset includes immigrants with a higher share of longer ties. The decay in the tie probability with distance is larger in Concepción, which could be explained in part by the higher ratio of transportation costs to wages compared to the other study areas. Table 4 gives the power law estimates and the tie distance comparison of the different datasets. As expected, the tie distance relationships between the datasets from Zurich and Switzerland are not substantially different from each other.

It is also interesting to note that in all data sets there are social contacts living at distances larger than 100 km, 1000 km, or even 10,000 km. A local peak can clearly be identified at around 8000–12,000 km, due to the distance between America and Europe.

The data from Eindhoven were excluded in Fig. 1 because the distances were recorded in discrete ranges. However, Table 5 presents the numerical comparison of the Eindhoven data with the other four datasets.

5. Methods and results

5.1. Hierarchical linear modelling

The datasets have an unbalanced hierarchical structure, resulting from respondents having different numbers of reported social contacts. The characteristics of alters and ego–alter relationships cannot be treated as independent observations as they depend on individual characteristics of the respondent, amongst other influences. Therefore, multilevel linear regression modelling techniques are employed, since they can estimate unbiased coefficients in hierarchically clustered dataset structures (for detailed information on multilevel modelling, see Snijders and Bosker, 1999; Goldstein, 1995). In particular, the datasets studied in this paper can be jointly structured in three levels. Level 1 includes information that depends on the characteristics of each alter and each ego–alter relationship. Level 2 includes the ego's socio-demographic as well as such personal network characteristics the number of contacts and the proportion of alters with common characteristics. Finally, level 3 includes information about the study area where each data collection took place, as previously shown in Table 1.

Two modelling approaches are used to study the ego–alter geographical distance as the dependent variable. The first approach models the pooled data jointly from all the study areas, using the common available variables from each of the datasets. Comparing

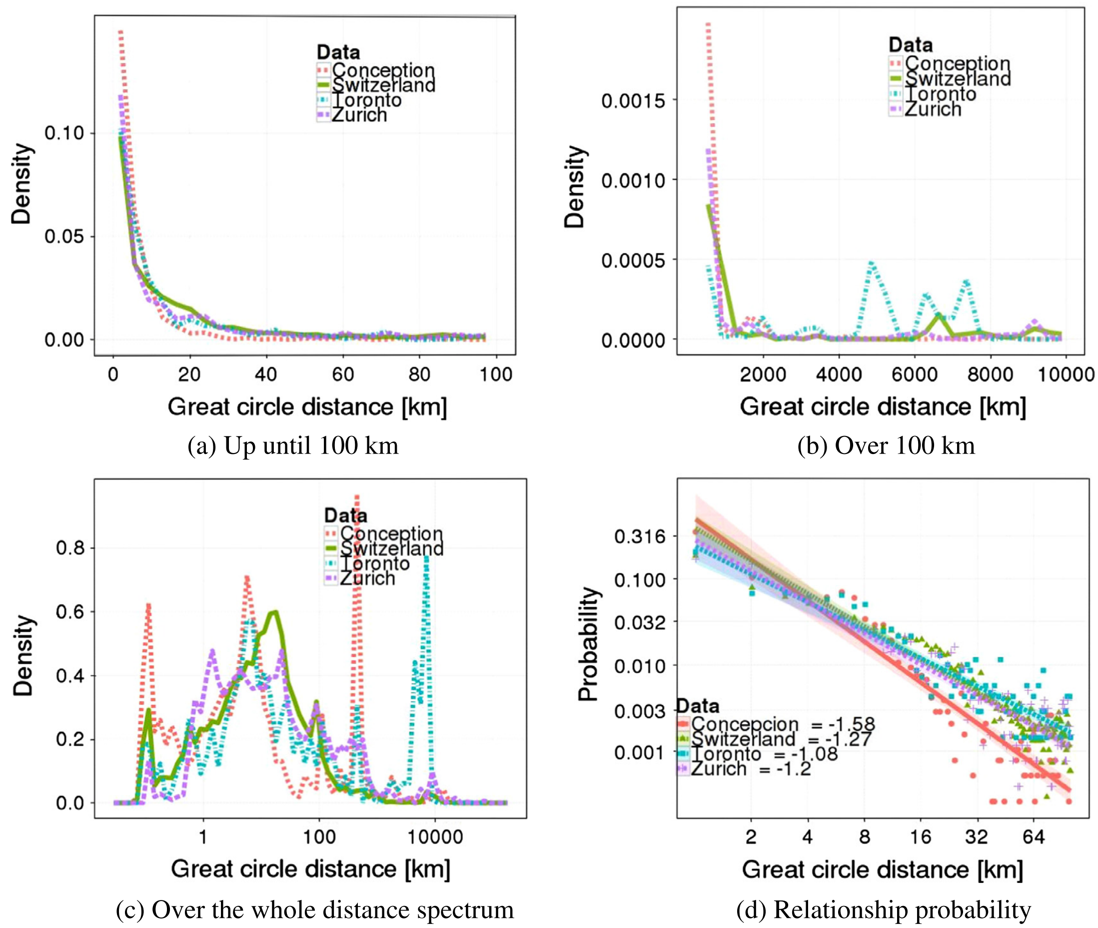

Fig. 1. Distance distribution between egos and alters.

Table 4
Tie distance distribution analysis within 100 km.

Distance power law β	Est.	Std.E.	$P(>\|t\|)$	Est. diff. from Concepción	$P(>\|t\|)$	Est. diff. from Switzerland	$P(>\|t\|)$
Concepción	−1.58	0.07	0.000	–	–	–	–
Switzerland	−1.27	0.06	0.000	0.31	0.001	–	–
Zurich	−1.20	0.06	0.000	0.38	0.000	0.07	0.387
Toronto	−1.08	0.07	0.000	0.50	0.000	0.19	0.032

Table 5
Distance distribution of social contacts.

	Toronto	Zurich	Eindhoven	Concepción	Switzerland
Distance <2 km	19.3	26.0	17.9	36.1	24.6
Distance 2–30 km	42.9	43.2	47.6	41.3	54.1
Distance 30–100 km	9.4	12.4	17.3	2.4	14.5
Distance >100 km	28.3	18.4	17.2	20.2	6.8
Distance mean	1036.0	286.6	152.9	222.9	106.3
Distance median	11.2	9.0	10.0	4.9	8.9

the estimates aims to identify influential parameters between the surveys and quantify their impact. The second approach consists of the five individual models which best describe each dataset using variables that were not necessarily collected in all the five collection efforts. Comparing the models with variables that are only available in some of the datasets allows us to recommend which set of variables should be included in future survey work and modelling of social networks and transportation.

Distances between egos and alters are often rather small, but can be very large as well, resulting in outliers and extreme values in all datasets. Despite this issue, no cut off points were defined; that is, all data were considered. The reason for this methodological decision is that an outlier may be at much higher distances in Toronto, where many respondents are immigrants, than in Switzerland, Eindhoven, or Zurich, where most personal networks have a local or regional character. As a consequence defining a unique cut off point involves having different numbers of omitted observations for each dataset, with a potential bias on the results.

A residual maximum likelihood estimator (REML) was preferred over the least squares method (OLS) as OLS estimation is known for having disadvantages when dealing with skewed distributions including multiple outliers. Although both REML and Maximum Likelihood (ML) estimators fit parameters to the overall dataset, REML is preferred since it is more appropriate when estimating mixed models including fixed as well as random effects (Snijders and Bosker, 1999).

A logarithmic transformation was employed to address the skewed distance distribution in each dataset. Distances with zero km were recorded due to rounding and aggregation, considering that not all data bases are geocoded at the household level, but to higher level units such as blocks and street corners. In order to avoid infinite values and the loss of observations, an empirically calculated constant was added to the smallest 2% of all distances in each dataset. The procedure helped to avoid adding a constant to each distance, which would have had a higher influence on small distances than on very large ones (Stahel, 2009).

5.2. Joint model on the pooled data

As explained before, the joint model has a three level structure, involving ego–alter relations nested in personal networks of egos that are nested in turn within specific study areas. Note that these regression techniques require models to have complete case observations without missing values. This limits the model to variables that can be considered in all five datasets.

In general, these hierarchical modelling techniques allow the employment of fixed and random effects on the independent variables. In case of such a mixed model approach, random group specific intercepts and slopes account for within group variations of lower level covariates and factors. To get non-biased estimates and test statistics, the intercept is specified as random. In addition, interaction terms and non-linear effects were studied in order to describe the independent variables as accurately as possible.

The results of estimation show both similarities and differences among the datasets. The estimated parameters and standard errors are presented in Table 6, together with the overall statistics of the model.

In level 3, separate intercepts for the different survey areas were calculated as a way of capturing the differences for the country and city specifics of each datasets. Even though the distance distributions from each dataset are similar, the level 3 variables account for the remaining variations. As previously mentioned, the datasets from Toronto and Concepción are especially different from the rest in their distance distribution, as the Toronto dataset includes much higher shares of very short distances and longer distances while Concepción shows a stronger decline in shares for shorter distances. The estimated negative coefficients in this level for those two areas can be partly attributed to those tendencies. Since there are only 5 observations available in level 3, the intercepts measure composition effects from the study area and the survey design discussed in chapter 3. Also because of the very small number of observations available for the third level, the intercept is estimated as a fixed effect to make sure, that the multilevel model is more efficient than a simple random sample designs (see Snijders, 2005). The variance's estimates should be looked at with caution due to the small sample size of the third level.

The coefficients in levels 2 and 1 were tested as constrained across all survey regions and also without the constraints. An ANOVA over all possible combinations was used to compare the different model structures. Although the models were estimated with REML, the log-likelihood from the profiled deviance is used to compare the models with a likelihood ratio test. Pinheiro and Bates (1995) showed that the deviance at the REML estimates for the complete parameter vector is close to the optimal profiled deviance. Only the effects of gender and internet access do not vary between the study areas, whereas all other effects do vary. Although the effects are very similar between the Zurich and Switzerland data, the varying effects among all datasets show that social network distance patterns and their manifestation are different across space. Transferability of results from these datasets in space is therefore limited. On the other hand, transferability on time seems rather applicable, as distance patterns and different influencing effects on these patterns are fairly stable in the case of Zurich and Switzerland from 2005 to 2010.

In level 2, female egos tend to maintain shorter distant contacts. The influence of the ego's age is dispersed, as it shows an increase in distance with age for Toronto, a non-significant effect for Eindhoven and Concepción, and a significant decrease in distance for Zurich and Switzerland. The quadratic non-linear influence of age shows that its marginal influence decreases for all the datasets with increasing values. Egos living with a partner show a decrease in distance for all of the datasets, except for Concepción; however, the parameter is not significant in this latter case, but it is significant for Eindhoven and Zurich. This result suggests that there is a trade-off between the travel time needed to maintain relationships and the actual available time, as it seems that people prefers maintaining a short distant network if they can invest the gained time in emotionally closer relationships, such as a partnership. Alternatively, this result could be explained with the presence of children, which often leads to more local social contacts. In the case of Eindhoven, this is the variable with the strongest influence on level 2. Internet access has a positive influence on distance. However, it should be noted that internet access might have a different meanings according to the year in which the survey was conducted, as the Zurich and Toronto datasets are older than Eindhoven and Switzerland. Although the difference in years seems small, the Internet access shares increased from 67% for the Zurich dataset to 96% in the Swiss data. On the other hand, Internet use has changed drastically in recent years, especially in the way people perceive and use social media. Internet access alone does not reflect on that, and internet use should be incorporated into future studies.

No significant income effects on distance are observed for Eindhoven, Zurich, and Switzerland. Even though this result seems to be counterintuitive, as higher income allows spending more on travelling, it can be explained by the low travel costs compared to the wage levels in those countries. This effect is different for the case of Concepción, where the income level becomes highly significant and important. In other words, the trade-off between monetary travel costs and maintaining contact over longer distances becomes measurable here. Finally, for the case of Toronto, the high income egos produce lower distances probably since income acts as a variable mediating immigration. Finally, higher educated people in general tend to maintain longer distance relationships, although this effect is only significant for Zurich and Toronto. The very high share of high education levels in Switzerland might influence this latter effect for the Swiss data. The effects of education on distance are expected, as higher education enables and demands more travelling in professional life and may involve a more diverse geographical biography.

The level 1 parameter estimates are similar among the five datasets for relatives, but different for emotional close relationships. Although the alter characteristics are relevant to understand ego–alter distances, the effect is not the same among the different study areas. Relatives live further away from egos compared to alters labelled as 'friends' in all of the datasets, except for extended family in the case of Concepción where family does not have influence at all. The overall influence of family is similar for immediate and extended family members. In general, people tend to maintain contact with their family members independent of geographical distance.

Table 6
Three level model on pooled data.

Number of observations (Level 1)	18,407				
Number of respondents (Level 2)	1099				
Number of study areas (Level 3)	5				
	BIC	Chisq.	Pr(>Chisq.)		
All coefficient constraint across regions	50,216				
Gender and Internet access fixed across regions	49,921	687.62	0.000		
No coefficients constraint across regions	49,994	5.14	0.743		
Characteristics of study area (Level 3)					
Intercepts	Estimate	Std.-Error			
Eindhoven	1.429**	(0.547)			
Zurich	1.324**	(0.293)			
Switzerland	1.104**	(0.362)			
Concepción	−0.526**	(0.289)			
Toronto	−1.235*	(0.772)			
	Eindhoven	Toronto	Concepción	Zurich	Switzerland
Egos' characteristics (Level 2)					
Female [y/n]	−0.078**				
	(0.034)				
Age [years]	−0.025	0.078**	0.031**	−0.022**	−0.022**
	(0.021)	(0.034)	(0.014)	(0.011)	(0.011)
Age^2/100 [years]	0.028	−0.077**	−0.027*	0.017*	0.021**
	(0.020)	(0.034)	(0.014)	(0.011)	(0.011)
Living with partner [y/n]	−0.265*	0.021	0.051	−0.194**	−0.060
	(0.142)	(0.154)	(0.08)	(0.07)	(0.071)
Internet access [y/n]	0.213**				
	(0.059)				
Medium income [y/n]	0.076	−0.793	0.437**	0.074	−0.003
	(0.146)	(0.199)	(0.102)	(0.085)	(0.096)
High income [y/n]	0.222	−0.708**	0.575**	0.132	0.068
	(0.145)	(0.207)	(0.113)	(0.132)	(0.100)
Medium education [y/n]	0.047	0.448**	−0.127	0.189	−0.003
	(0.156)	(0.223)	(0.108)	(0.133)	(0.222)
High education [y/n]	0.156	0.765**	0.015	0.337**	0.130
	(0.155)	(0.201)	(0.100)	(0.145)	(0.220)
Alters' characteristics (Level 1)					
Immediate family [y/n]	0.438**	0.839**	0.002	0.024	0.278**
	(0.053)	(0.083)	(0.022)	(0.044)	(0.035)
Extended family [y/n]	0.485**	0.948**	−0.002	0.513**	0.379**
	(0.046)	(0.108)	(0.022)	(0.054)	(0.035)
Very close tie (emotionally) [y/n]	0.068*	0.159**	0.200**	−0.143**	−0.032
	(0.041)	(0.070)	(0.027)	(0.034)	(0.027)
Random effects					
	Variance	Std. deviation			
Level 3 Intercept	0.0005	(0.023)			
Level 2 Intercept	0.2112	(0.456)			

Without flag = not significant ($p > 0.10$).
* Nearly significant ($p < 0.10$).
** Significant ($p < 0.05$).

Distances between strong ties are smaller than average in Switzerland and Zurich, and larger than the average in Eindhoven, Toronto and Concepción. This result could be an effect of the survey instruments: emotional closeness was implied in the name generator of Eindhoven, Toronto, and Concepción, but measured in the name interpreter with the help of proxy questions in the case of Zurich and Switzerland. Similarly, the explicit consideration of leisure in the name generator of the Switzerland and Zurich studies can also influence this result. Despite this consideration, the evidence from the models suggests that people prefer living close to those important to them, but maintain contact to strong ties independently of their geographical distance.

5.3. Parameter comparison between single dataset estimations

Independent two-level hierarchical linear regression models were calculated for each of the five datasets, without considering country specific information explicitly. Each model used a broader set of variables available on hierarchy levels 1 and 2 for each dataset, and thus can incorporate variables which may not be present in all datasets.

The models for each dataset are restricted to achieve maximum explanatory power by employing as few variables as possible to provide a better comparison of relevant effects. No random variations (in intercept and/or slope) were specified, as mixed model estimations for each of the different datasets could result in dissimilar random intercepts and/or slopes. Although the statistical tests can have some biases due to this omission, comparisons between parameters are easier and more appropriate for the aims of this study. To represent the characteristics of each dataset as accurately as possible, the models include interactions and non-linear effects for each dataset. Tables 7a and 7b provides an overview of the parameter estimates and significance levels between the different studies.

The estimation results show some similarities and differences among the study areas. A general comparison highlights the different explanatory power of data hierarchies depending on dataset. Extreme positions can be found for Switzerland on the one hand, where most of the level 1 variables have explanatory power whilst only very few on level 2, and Zurich on the other hand, where the opposite occurs.

Regarding level 2 effects, the influence of egos' age is dissimilar: While distance increases with age in Toronto, Concepción, and Zurich (where a quadratic non-linear influence is observed), the relation is negative in Eindhoven. Another ambiguous effect can be seen for the years a respondent lives at a certain location. Whilst this has a decreasing effect on network distances in Concepción, there is an exponentially increasing effect in Switzerland. In contrast, a non-ambiguous effect results from egos living with a partner or who have children in their households. All these people have smaller distances in their personal network, suggesting that those who care for others prefer to maintain short distance relationships. Certainly, from an ego's perspective this effect can be related to the trade-off between the benefits from emotional contacts and distance costs, since nearby alters can be reached without spending much time and money on travel costs. Such savings in time and money can be invested in emotionally important relations, such as partnership or parenthood. As discussed before, alternatively it might be a more direct effect of having children, who tend to have a local social environment (e.g. school and playground) that enables parents to come into contact with other parents.

The availability of transportation – measured as car availability or season ticket for public transport – and communication means – measured as access to the Internet or a mobile phone – increases network distances. In general, the data show that people with mobility and communication resources maintain longer distance relations. In addition, a higher education level of egos has a positive effect on network distances. This effect is very clear in Eindhoven, Toronto, and Zurich, where network distances increase from egos with mandatory school education to those with a technical background and reporting academic degrees. Considering that institutions of education are good places to meet others and establish relations, this result is not surprising. In addition, it is often reported that people with a higher status than average are often attractive to others (Marin, 2004). Finally, each dataset resulted in additional effects and interaction effects, which are exclusively relevant for one of the five studies and cannot be compared to the other estimates.

Regarding level 1, results show a more diverse picture than for level 2. While relatives live closer to egos than friends in Eindhoven, this relationship is the opposite in Toronto and Concepción. In addition, immediate family members live closer than friends in Zurich and Switzerland, while extended family members live further away. These results suggest that people stay in contact with their family members independently of geographical distances, which could be expected since family members are in general a reliable source of help and support, and may offer kinds of support that other social contacts may not. Differences in effects can be explained with the characteristics of either respondents or survey areas. In case of Toronto, the high share of immigrants among the respondents has an influence. In addition, Canada and Chile are large countries, where

Table 7a
Comparison between effects from each dataset.

Dataset	Eindhoven	Toronto	Concepción	Zurich	Switzerland
Number of respondents	106	67	235	265	426
Number of observations	2452	756	4718	3156	7293
BIC	5888	2535	13,894	8547	18,156
Var[e]	0.552	1.223	0.974	0.745	0.626
Var[u]	0.091	0.356	0.307	0.165	0.124
Corr[v(i,t),v(i,s)]	0.141	0.225	0.239	0.181	0.165
Constant	1.028**	−1.659*	−0.100	2.878**	0.826**
Egos' characteristics (Level 2)					
Female [y/n]				−0.138**	
Age [years]	−0.008**	0.037**	0.007**	−0.023**	
Years in current address [years]			−0.018**		−0.015**
Nr of previous addresses [No]				−0.121**	
Work hours/week [h]					
Network size [social contacts]					
Proportion immediate family					
Proportion extended family					
Proportion males					
Proportion 30–60					
Proportion >60					
Proportion same gender					
Proportion same age					
Living with partner [y/n]	−0.923**			−0.312**	
Children in household [y/n]		−0.332*			
Car available [y/n]	0.076**				
Season ticket [y/n]	0.222**				0.143**
Internet access [y/n]		2.345**	0.393**	0.237**	
Mobile phone access [y/n]					
Medium education [y/n]	0.083	0.578**		0.177	
High education [y/n]	0.188*	0.748**		0.324**	
Medium income [y/n]		−0.412*	0.350**		
High income [y/n]		−0.806**	1.766**		
Level 2 interaction and non-linear effects					
Age [years] * Internet access [y/n]		−0.036**			
High income [y/n] * Internet access [y/n]			−1.289**		
Living with partner [y/n] * Car available [y/n]	0.922**				
(Age)^2				0.001**	
(Years in current address)^2					0.001**

Without flag = not significant ($p > 0.10$).
* Nearly significant ($p < 0.10$).
** Significant ($p < 0.05$).

Table 7b
Comparison between effects from each dataset.

Dataset	Eindhoven	Toronto	Concepción	Zurich	Switzerland
Alters' characteristics (Level 1)					
Female [y/n]			−0.085**		−0.145**
Age (30–60 years) [y/n]	−0.129**	−0.142			−0.019
Age (>60 years) [y/n]	−0.007	0.454**			−0.209**
Immediate family [y/n]	−0.283**	0.659**	0.646**	−0.522**	−0.219**
Extended family [y/n]	−0.004	0.900**	0.303**	2.321**	0.321**
Very close tie (emotionally) [y/n]	0.065**			−0.256**	−0.066**
Duration relationship [years]					0.011**
Known <1 year [y/n]	0.310**				
Known 1–10 years [y/n]	0.553**				
Known <10 years [y/n]				−0.973**	
Same gender as ego [y/n]					
Same age as ego [y/n]	0.075**	0.566**	0.067*		0.222**
Level 1 interaction and non-linear effects					
Age (>60 years) [y/n] * Same age as ego [y/n]		−0.851**			
Age (>60 years) [y/n] * Immediate family [y/n]	−0.240**				
Known <10 years [y/n] * Extended family [y/n]				−2.824**	
Immediate family [y/n] * Duration relationship					0.007**
Alter female [y/n] * Duration relationship					0.006**
Age (30–60 years) [y/n] * Duration relationship					−0.003**
Age (30–60 years) [y/n] * Same age as ego [y/n]					−0.339**
Cross level interaction effects					
Immediate family [y/n] * Children in household [y/n]		0.520**			
Extended family [y/n] * Age [years]			0.008**		
Extended family [y/n] * Internet access [y/n]			−0.192**		
Extended family [y/n] * High education [y/n]	0.169**				
Known 1–10 years [y/n] * Living with partner [y/n]	−0.193**				
Known 1–10 years [y/n] * Season ticket [y/n]	0.177**				
Very close tie (emotionally) [y/n] * Living with partner [y/n]				0.196**	
Known < 10 years [y/n] * Nr of previous addresses [No]				0.145**	
Duration relationship * Years in current address					−0.001**

Without flag = significant ($p < 0.05$).
* Nearly significant ($p < 0.10$).
** Not significant ($p > 0.10$).

movements often cover larger distances than Switzerland and the Netherlands. Another effect, which is visible in all datasets, is the role of homophily in age (similarity between egos and alters in terms of age). This index was calculated using the three available age classes for both egos and alters (<29, 30–60; >60 years). The results from the models echo the influence of homophily on several aspects of social life, as documented by authors such as McPherson et al. (2001). In this case, the models suggest that people tend to maintain long distance relationships with contacts with similar ages. Similarly, ego–alter distances increase the longer their relationship lasts. This result was intuitively expected since, if a relationship has existed for certain time, people take increased efforts to maintain it even if one of them moves and it becomes a long distance relationship. Regarding the relevance of emotional closeness – expressed in terms of tie strengths – distances between strong ties are smaller than average in Switzerland and Zurich whereas they are larger in Eindhoven. As mentioned above, this result may be an effect from the survey instruments since emotional closeness was implied in the name generator of Eindhoven, Toronto and Concepción whilst it was measured with the help of separate questions in the name interpreter for Zurich and Switzerland. Furthermore, the name generator of the Switzerland study highlighted leisure contacts, which are not necessarily identical with emotionally important contacts. The alters' age shows a dissimilar effect with decreasing distances for older egos in Toronto and Switzerland, and a contrary effect in Eindhoven. The alters' sex shows a decreasing effect on distances for females in the models where the coefficient is significant. Finally, there are several significant interaction effects and cross level interactions, which do not overlap between the datasets and that can be found on these individual models.

6. Conclusions

The interest in understanding the role of social networks in travel behaviour has motivated dedicated data collections in several countries that used techniques from sociology and other related fields to elicit the respondents' personal networks. In this context, a key research question is related to the spatial distribution of the individuals' social contacts, which both constitute a relevant portion of their social activity space and is important to understand the destination choices of his social activity travels. Although previous work has analyzed empirical data on this issue from different countries, the aim of this paper was performing a comparative exercise among five datasets from four different countries: Canada (Toronto), Chile (Concepción), Switzerland (Zurich and the whole country), and the Netherlands (Eindhoven). This comparative effort is facilitated by the similar approach of the five data collection efforts, both in terms of the social network eliciting techniques as well as the key explanatory variables explored.

The descriptive analysis highlighted the similarities and differences in the data collected from the different countries, in particular regarding the tie distance distribution. In fact, although all data sets follow a power law distribution, a stronger decay is found in Concepción, and a smoother effect in Toronto, while the European datasets are in between these two cities. These differences highlight the relevance of contextual aspects such as the ratio between wage and transport costs, the availability of mobility tools – such as car ownership and Internet access – and the influence of immigration.

Using multilevel statistical methods, two sets of models were estimated. The first model consisted of a three level structure which used the five data sets simultaneously, capturing the variance at the country, ego-network, and ego–alter levels. The models confirmed the power law distribution results, suggesting further explanations about the differences among cities. In fact, income plays a key role only in Concepción, suggesting that the trade-off between monetary costs and contact maintenance occurs only until a certain level of wealth is achieved, in terms of the relative cost of transport with respect to wages. On the other hand, education – a variable more related to the mobility and residential moving history – is relevant in other cities such as Toronto and Zurich. The respondent's socio-demographic characteristics such as age and gender show a limited influence on the spatial distribution of personal networks in the combined dataset, when controlling for the national differences. In addition, the comparative exercise also illustrates that, although people maintain their family contacts regardless of the geographical distances, distances with respect to strong ties differ throughout the countries.

A second set of models complemented the first by studying the covariates of ego–alter distances in separate models for each dataset. The results show some ambiguous effects regarding sociodemographic characteristics such as age and sex. These contrast with the general tendency to contact family members independently of distance, and to contact emotionally close or similarly aged contacts at farther distances.

Overall, the results suggest that, more than any other country specific variable, the availability of transport and communication relative to income plays a key role in the spatial distribution of contacts. In addition, family and emotionally close members can be spread over different distance ranges, contrasting with other social contacts. Similarly, and despite the context of the country, a proper understanding of the spatial dispersion of social contacts needs to incorporate the characteristics of egos, ties, and the overall personal network.

However, besides these general regularities, the relevance and magnitude of the specific components strongly depends on their city or national context where these networks are embedded. These specificities highlight the need for further research on the key reasons of the different distance patterns between cities, and also developed and developing nations, disentangling to which extent not only costs but also socio-cultural aspects influence these differences. In this way, transport research will gain a more comprehensive understanding of the spatiality of personal networks, incorporating its non-linearities, discontinuities, and overall contextual characteristics.

Finally, future research should also incorporate the relationship between distance and frequency of interaction, increasing our understanding about the social activity-travel patterns in cities.

Acknowledgements

We thank Lilia Smale, Christian Beermann, and Esther Sok for their assistance in the final copyediting of the paper. We also thank the funding agencies that made feasible the collection and analysis of these data bases. In the case of Toronto, we thank the Social Science and Humanities Research Council of Canada, and the GRAND Network of Centres of Excellence. In the case of Concepción, we thank the Chilean National Fund for Sciences and Technology, FONDECYT 11070075, and the Chilean Fund for Research Centers in Prioritary Areas, CONICYT/FONDAP 1511020.

References

Arentze, T.A., Dellaert, B.G.C., Timmermans, H.J.P., 2008. Modeling and measuring individuals' mental representations of complex spatio-temporal decision problems. Environment and Behaviour 40 (6), 843–869.

Axhausen, K.W., 2005. Social networks and travel: some hypotheses. In: Donaghy, K.P., Poppelreuter, S., Rudinger, G. (Eds.), Social Dimensions of Sustainable Transport: Transatlantic perspectives. Ashgate, Aldershot, pp. 90–108.

Banister, D., Bowling, A., 2004. Quality of life for the elderly: the transport dimension. Transport Policy 11, 105–115.

Carrasco, J.A., Miller, E.J., Wellman, B., 2006. Spatial and social networks: the case of travel for social activities. In: 11th International Conference on Travel Behaviour Research, Kyoto, August 16–20, 2006.

Carrasco, J.A., Miller, E.J., Wellman, B., 2008a. How far and with whom do people socialize? Empirical evidence about the distance between social network members. Transportation Research Record: Journal of the Transportation Research Board 2076, 114–122.

Carrasco, J.A., Hogan, B., Wellman, B., Miller, E.J., 2008b. Collecting social network data to study social activity-travel behaviour: an egocentric approach. Environment and Planning B 35 (6), 961–980.

Carrasco, J.A., Cid-Aguayo, B., 2012. Network capital, social networks, and travel: an empirical illustration from Concepción, Chile. Environment and Planning A 44 (5), 1066–1084.

Carrasco, J.A., C. Bustos, and B. Cid-Aguayo (2013), "Affective personal networks versus daily contacts: Analyzing different name generators in a social activity-travel behaviour context", Chapter 22, p. 407–424, in Transport Survey Methods: Best Practice for Decision Making, J. Zmud, M. Lee-Gosselin, M. Munizaga, and J.A. Carrasco (eds.), London: Emerald Press.

Chalasani, V.S., Engebretsen, O., Denstadli, J.M., Axhausen, K.W., 2005. Precision of geocoded locations and network distance estimates. Journal of Transportation and Statistics 8 (2), 1.

Daly, 2010. Cost Damping in Travel Demand Models: Report of a Study for the Department for Transport. Report of a Study for the Department for Transport. RAND TR-717-DFT.

Degenne, A., Forsé, M., 1999. Introducing Social Networks. Sage Publications, London.

Dugundji, E.R., Gulyás, L., 2008. Sociodynamic discrete choice on networks in space: impacts of agent heterogeneity on emergent outcomes. Environment and Planning B: Planning and Design 35 (6), 1028–1054.

Dugundji, E.R., Páez, A., Arentze, T.A., Walker, J.L., Carrasco, J.A., Marchal, F., Nakanishi, H., 2011. Transportation and social interactions. Transportation Research Part A: Policy and Practice 45 (4), 239–247.

Dugundji, E.R., Walker, J.L., 2005. Discrete choice with social and spatial network interdependencies: an empirical example using mixed generalized extreme value models with field and panel effects. Transportation Research Record: Journal of the Transportation Research Board 1921 1, 70–78.

Ettema, D., Arentze, T.A., Timmermans, H.J.P., 2011. Social influences on household location, mobility and activity choice in integrated micro-simulation models. Transportation Research Part A: Policy and Practice 45 (4), 283–295.

Ettema, D., Schwanen, T., 2012. A relational approach to analysing leisure travel. Journal of Transport Geography 24 (3), 173–181.

Frei, A., Axhausen, K.W., 2007. Size and structure of social network geographies. Arbeitsberichte Verkehrs- und Raumplanung 439, IVT, ETH Zürich, Zürich.

Frei, A., Axhausen, K.W., 2011. Collective Location Choice Model, Arbeitsberichte Verkehrs- und Raumplanung, 686. IVT, ETH Zürich, Zürich.

Goldstein, H., 1995. Multilevel Statistical Models. Edward Arnold, London.

Hackney, J., Axhausen, K.W., 2006. An agent model of social network and Travel behavior interdependence. In: 11th International Conference on Travel Behaviour Research. Kyoto.

Han, Q., Arentze, T.A., Timmermans, H.J.P., Janssens, D., Wets, G., 2011. The effects of social networks on choice set dynamics: results of numerical simulations. Transportation Research Part A: Policy and Practice 45 (4), 310–322.

Hanson, S., 1998. Off the road? Reflections on transportation geography in the information age. Journal of Transport Geography 6 (4), 241–249.

Hogan, B., Carrasco, J.A., Wellman, B., 2007. Visualizing personal networks: working with participant-aided sociograms. Field Methods 19 (2), 116–144.

Horton, F., Reynolds, D., 1971. Action space formation: a behavioral approach to predicting urban travel behavior. Highway Research Record 322, 136–148.

Kowald, M., Axhausen, K.W., 2012. Focussing on connected personal networks: selected results from a snowball sample. Environment and Planning A 44 (5), 1085–1100.

Larsen, J., Urry, J., Axhausen, K.W., 2006. Mobilities, Networks, Geographies. Ashgate, Aldershot, UK.

Latané, B., Liu, J.H., Nowak, A., Bonevento, M., Zheng, L., 2005. Personality and Social Psychology Bulletin 21 (8), 795–805.

Marsden, P.V., 2005. Recent developments in network measurement. In: Carrington, P., Scott, J., Wasserman, S. (Eds.), Models and Methods in Social Network Analysis. Cambridge University Press, New York.

Marchal, F., Nagel, K., 2005. Modeling location choice of secondary activities with a social network of cooperative agents. Transportation Research Record: Journal of the Transportation Research Board 1935, 141–146.

Marin, A., 2004. Are respondents more likely to list alters with certain characteristics? Implications for name generator data. Social Networks 26, 289–307.

Marsden, P.V., 1987. Core discussion networks of Americans. American Sociological Review 52 (1), 122–131.

McPherson, M.-L., LynnBrashears, Matthew (2006) Social Isolation in America: Changes in Core Discussion Networks over Two Decades. American Sociological Review 71.

McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: Homophily in social networks. Annual Review of Sociology 27, 415–444.

Miller, H.J., 2006. Social exclusion in space and time. In: Axhausen, K.W. (Ed.), Moving Through Nets: The Social and Physical Aspects of Travel. Elsevier, Amsterdam, pp. 353–380.

Mok, D., Wellman, B., Carrasco, J., 2010. Does Distance Matter in the Age of the Internet? Urban Studies 47, 2747–2783.
Naess, P., 2011. 'New urbanism' or metropolitan-level centralization? A comparison of the influences of metropolitan-level and neighborhood-level urban form characteristics on travel behaviour. Journal of Transport and Land Use 4 (1), 25–44.
Newbold, K., Scott, D., Spinney, J., Kanaroglou, P., Páez, A., 2005. Travel behavior within Canada's older population: a cohort analysis. Journal of Transport Geography 13, 340–351.
Ohnmacht, T., Axhausen, K.W., 2005. Entwicklung des Forschungsdesign und der Erhebungsinstrumente für das Projekt Mobilitätsbiographien, Mobilitätswerkzeuge und soziale Netze, Arbeitsberichte Verkehrs- und Raumplanung, 298. IVT, ETH Zürich, Zürich.
Ohnmacht, T., Gotz, K., Schad, H., 2009. Leisure mobility styles in Swiss conurbations: construction and empirical analysis. Transportation 36, 243–265.
Páez, A., Scott, D.M., 2007. Social influence on travel behavior: a simulation example of the decision to telecommute. Environment and Planning A 39, 647–665.
Páez, A., Scott, D.M., Volz, E., 2008. A discrete choice approach to modeling social influence on individual decision making. Environment and Planning B 35 (6), 1055–1069.
Pinheiro, J.C., Bates, D.M., 1995. Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics* 4 (1), 12–35.
Rietveld, P., Zwart, B., van Wee, B., van den Hoorn, T., 1999. On the relationship between travel time and travel distance of commuters. The Annals of Regional Science 33 (3), 269–287.
Schlich, R., Schönfelder, S., Hanson, S., Axhausen, K.W., 2004. Structures of leisure travel: temporal and spatial variability. Transport Reviews 24 (2), 219–237.
Schönfelder, S., 2006. Urban rhythms: Modelling the rhythms of individual travel behaviour (Doctoral dissertation, ETH Zürich).
Schwanen, T., 2008. Managing uncertain arrival times through sociomaterial Associations. Environment and Planning B 35 (6), 977–1011.
Schwanen, T., Mokhtarian, P., 2005. What if you live in the wrong neighborhood? The impact of residential neighborhood type dissonance on distance travelled. Transportation Research Part D: Transport and Environment 10 (2), 127–151.
Snijders, T.A., Bosker, R.J., 1999. Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modelling. Sage, London.
Snijders, T.A., 2005. Power and Sample Size in Multilevel Linear Models. In: Everitt, B.S., Howell, D.C. (Eds.), Encyclopedia of Statistics in Behavioral Science, vol. 3. Wiley, Chicester, pp. 1570–1573.
Stahel, W., 2009. Statistische Datenanalyse. Eine Einführung für Naturwissenschaftler, Vieweg, Wiesbaden.
Van den Berg, P.E.W., Arentze, T.A., Timmermans, H.J.P., 2009. Size and composition of ego-centered social networks and their effect on travel distance and contact frequency. Transportation Research Record: Journal of the Transportation Research Board 2135, 1–9.
Van der Poel, M.G.M., 1993. Delineating personal support networks. Social Networks 15, 49–70.
Wasserman, S., Faust, K., 1994. Social Network Analysis. Cambridge University Press, New York.