Predictors and Outcomes of Social Network Compositions.

A Compositional Structural Equation Modeling Approach

Tina Kogovšek[*]

Faculty of Arts and Faculty of Social Sciences, University of Ljubljana

Germà Coenders

Department of Economics, University of Girona

Valentina Hlebec

Faculty of Social Sciences, University of Ljubljana

Abstract

Proportions of a total, including social network compositions (proportions of partner, family, friends etc.,) lie in a restricted space, which challenges statistical analysis. Network compositions can be both dependent and explanatory variables and are usually measured with error by survey instruments. Structural equation models make it possible to correct measurement error bias. Coenders et al. (2011) fitted a factor analysis model to transformed network compositions. In this article, we use another transformation called an isometric log-ratio and we extend the model to include predictors and outcomes. The findings and hypotheses in the literature can be reformulated with isometric log-ratios in a more interpretable manner. For instance, we find relationships of gender with partner support, of education and extraversion with friend support, and of family support with tie multiplexity and closeness.

Keywords: network composition, structural equation model, measurement error, log-ratio transformation, compositional data, social support

[*] Corresponding author: University of Ljubljana, Faculty of Arts, Aškerčeva 2, 1000 Ljubljana, Slovenia, Tel.: + 386 1 2411 132, Fax: + 386 1 4259 337 and University of Ljubljana, Faculty of Social Sciences, Kardeljeva pl. 5, 1000 Ljubljana, Slovenia, Tel.: + 386 1 5805 285, Fax: + 386 1 5805 101, E-mail: tina.kogovsek@guest.arnes.si.

Introduction

An important aspect of social network research is network composition. For instance, the composition of egocentered social support networks can be expressed as proportions of family, friends, colleagues and other network members. Network composition can be conceptualized as an independent, explanatory variable that may affect different outcomes (e.g., health, satisfaction with life, satisfaction with relationships). Conversely, network composition can also be a dependent variable. Factors affecting social network composition range from people's demographic (e.g., gender, age, education) to personal characteristics (e.g., personality traits, such as extraversion or neuroticism, social skills) and various other factors (e.g., health, family history, physical environment, organizational membership, psychiatric symptoms).

In this article, we make two methodological contributions to the estimation of relationships involving social network composition.

The first contribution involves the treatment of measurement error in network compositions. It is certainly important to use measurement procedures that provide the most reliable and valid social network data. However, network compositions obtained by means of questionnaires will always be prone to measurement error bias. By means of structural equation models (SEM) and multiple indicators of network compositions we aim to obtain unbiased estimates of substantive effects upon and from network compositions. In this article, we draw from the compositional factor analysis model in Coenders et al. (2011) and construct an extended SEM including predictors and outcomes of social network compositions.

The second methodological contribution refers to the sheer nature of compositional data. Social network compositions are expressed as proportions of a total, whose sum can only be 1. Statistical analysis of compositional data in general, and particularly the estimation of SEM, are challenging tasks because compositional data lie in a restricted space and only convey information regarding the relative size of components to one another. As a result, compositional data have to be transformed by means of logarithms of ratios of components (Aitchison, 1986), which are not always easy to interpret. In this article, we extend the

method in Coenders et al. (2011) with more general and interpretable log-ratio transformations.

Working with log-ratios not only has methodological implications but also substantive ones. Without log-ratios, components are estimated and interpreted separately from each another as if they could vary independently from each another ("all other things being held constant"), which is impossible: the relative importance of one component can only increase if the relative importance of at least one other component decreases. On the other hand, with log-ratios, it is possible to study what happens with one component relative to another component, for instance, what happens to the role of a partner relative to the role of the rest of the family part of the network as a result of the effect of age; or, if we use network composition as an independent variable, how an outcome (e.g., the level or adequacy of social support or tie multiplexity) changes when the ratio between two composition components changes (e.g., percentage of friends increasing relative to the percentage of other non-family network members). Accordingly, instead of seeing network components as independent parts of the network, they are observed in a relationship of one component to another. Of course, log-ratios have to be constructed in a substantively meaningful way, as we believe we have managed to achieve in this article.

We illustrate our approach with a general population study of egocentered support networks. The outcomes we consider in this article are the multiplexity and strength of egocentered network ties. The covariates we consider are gender, age, education and extraversion. All these variables are common in the literature. The main contributions of this article do not lie in the set of studied variables but in the methodological approach.

In the next section, we summarize the original method in Coenders et al. (2011). We then present the extensions made in this article. We subsequently describe the illustration, the data, the log-ratios and rewrite the common theoretical hypotheses in the literature in terms of log-ratios. Finally, we present, interpret and discuss the results.

An approach to measurement error in social network compositions

In this section, we first present a measurement error model. We then discuss the specificities of compositional data and present the Coenders et al. (2011) approach to the use of compositional data in measurement error models.

Correlated uniqueness model for compositional multitrait-multimethod data

Multitrait-multimethod (MTMM) designs (Campbell and Fiske, 1959) are a well-established approach to measurement error in survey data (Saris and Gallhofer, 2007), including measurement error in social network surveys (Coromina and Coenders, 2006; Ferligoj and Hlebec, 1999; Kogovšek, 2006; Kogovšek et al., 2002). These designs consist of multiple measures of at least two factors (traits) with the same set of at least three measurement procedures (methods). Accordingly, these designs include *DM* measures, that is, the number of methods (*M*) times the number of traits (*D*). MTMM designs are usually analyzed by means of SEM (see Schumacker & Lomax 1996 for an overview), with the aim of taking measurement errors into account and estimating unbiased relationships among the traits. A number of SEMs for MTMM data have been formulated and tested in the literature. Coenders and Saris (2000) showed the great flexibility of the so-called correlated uniqueness (CU) model (Marsh, 1989), of which many other MTMM models constitute particular cases.

The CU model is specified as follows. Let x_{idm} be the mean-centered measurement of individual *i*, for trait *d* with method *m*:

$$x_{idm} = \lambda_{dm} t_{id} + e_{idm} \tag{1}$$

where t_{id} is the latent variable score of individual *i* corresponding to trait *d* and e_{idm} is the measurement error term of individual *i*, for trait *d* with method *m*. The model parameters are:

- λ_{dm}: factor loading of x_{idm} on trait t_{id}. The λ_{d1} parameters are constrained to 1 in order to fix the scale of t_{id}. Standardized trait loadings λ_{dm} show the strength of the relationship between the observed scores and latent trait scores and can thus be interpreted in terms of measurement validity and reliability.

- θ_{dm}: measurement error variance of e_{idm};

- $\theta_{dd'm}$: covariance between two measurement error terms sharing a common method e_{idm} and $e_{id'm}$. In an MTMM design it is expected that use of the same method involves common errors. All other error covariances are zero by assumption;

- ϕ_{dd}: error-free variance of the latent trait variable t_{id}; and

- $\phi_{dd'}$: error-free covariance between two latent trait variables t_{id} and $t_{id'}$.

A path diagram of the CU model with three traits and three methods is displayed in Figure 1.

{Figure 1 here}

Compositional or absolute data

Compositional data concern the relative size of *D* components within a total, usually expressed in proportions over 1. This is the case of data on network compositions expressed as proportions of friends, family and other types of members of a personal network. Conversely, absolute data consist of the counts of friends, family and other types of members. There are at least two reasons for using compositional data.

In some cases, network size and absolute data are meaningless. Consider, for instance, the simplified role relation method to measure social networks, which is widely used even in cross-national surveys such as the Generations and Gender Programme, the European Quality of Life Survey or the International Social Survey Programme. Respondents are asked questions about a series of situations which they cannot solve on their own, and have to indicate who they would ask for help (a friend, a family member, and so on). Network size is constant and trivially equal to the number of questions for all respondents. Only the relative importance of the network member types is meaningful.

In some cases, the research question concerns network compositions rather than absolute numbers. Research questions requiring compositions are any in which the relative importance of some network component is believed to be relevant. In some respect, absolute data can be understood as a combination of composition and size. The count of one component can increase both when the network size increases and when the relative importance of the component increases. It may even be useful to build one model for network compositions and a separate model for network size.

Challenges in the analysis of compositional data

The measurement quality of compositional data cannot be studied by simply fitting the proportions to a SEM (e.g., to a CU model). Compared to absolute data, compositional data lie in a constrained space. A D-term composition measured on individual i with a given method m: $x_{i1m}, x_{i2m},...,x_{iDm}$ has the following constraints:

$$0 \leq x_{idm} \leq 1 \text{ and } \sum_{d=1}^{D} x_{idm} = 1 \qquad (2)$$

Aitchison (1986) and Pawlowsky-Glahn and Buccianti (2011) warn against the serious problems that arise when using standard statistical analysis tools on compositional data:

- compositional data are non-normal and heteroskedastic;

- one component can only increase if some other(s) decrease. This results in negative spurious correlations among the components and prevents interpreting effects of linear models in the usual way "keeping everything else constant"; and

- the true dimensionality of a set of compositional variables measured with a given method m is D-1. Analysis of all D dimensions leads to perfect collinearity.

These problems also apply to SEM on compositional data. In the SEM literature, constant sum data are sometimes referred to as ipsative data (Chan, 2003; Cheung, 2004). In the CU model case, an important additional problem occurs (Coenders et al., 2011). The compositional x_{idm} data do not fit the CU model. Intuitively, for a given set of true compositions t_{idm} the observed component x_{i1m} can only increase if some other components decrease. x_{i1m} is thus not only dependent on t_{i1m} but on all t_{idm}. The CU model assuming each observed variable is to load only on a trait is thus misspecified.

Additive log-ratio transformation for compositional data

Even if specialized techniques for compositional data are starting to appear (e.g. Thió-Henestrosa and Martín-Fernández, 2005), compositional data can also be transformed so that they can be subject to standard and well-understood statistical techniques (Aitchison, 1986). This is the approach we take here.

In compositional data the absolute size of components is lost. Only the relative size of some components to the others is maintained. Thus, ratios are the only meaningful way of expressing the data. Several ratio transformations have been suggested in the literature (Egozcue et al., 2003). The additive log-ratio transformation (alr) is the most popular and easiest to compute since it is simply the log-ratio of each component to the last:

$$y_{idm} = \ln(x_{idm}/x_{iDm}) = \ln(x_{idm}) - \ln(x_{iDm}) \text{ with } d=1,2,...,D-1 \qquad (3)$$

The alr transformed y_{idm} variables recover the full unconstrained $-\infty$ to ∞ range. It must be noted that one dimension is lost.

Coenders et al. (2011) suggest simply estimating the CU model on the alr transformed y_{idm} data with conventional methods for SEM. The alr transformed data are appealing because of their simple computation, but still have certain limitations regarding parameter interpretation (Coenders et al., 2011), with the main one being that trait covariances and error covariances are spurious and positive because alr data have a common denominator. In this article, an alternative type of log-ratio transformation will be used, as explained below.

Dealing with zero components

If the x_{idm} variables contain zeroes, then log-ratios cannot be computed. An obvious initial procedure is to amalgamate small and conceptually similar components with many zeroes into larger ones. In social network compositions it may be feasible, for instance, to group several types of family member categories (uncle, cousin, father-in-law etc.) in a more global family member category. In certain instances, some zero components result from individual characteristics (essential zeroes; Aitchison, 1986). For instance, people who have never been employed cannot have co-workers in their social network. In this case, it is advisable to narrow the definition of the target population and drop individuals with essential zeroes.

After the amalgamation of components and redefinition of the population, the remaining (hopefully few) zeroes may be understood as components which are too small to be detected. For instance, in a network with size s_{im}, the smallest detectable component is $1/s_{im}$. Coenders et al. (2011) and Pierotti et al. (2009) recommend replacing values $x_{idm}=0$ with:

$$x'_{idm} = \frac{1}{D(s_{im} + 1)} \qquad (4)$$

Non-zero x_{idm} values have to be reduced in order to preserve the unit sum. As suggested by Martín-Fernández et al. (2003), this implies replacing $x_{idm} > 0$ with:

$$x'_{idm} = x_{idm}\left(1 - \sum_{x_{idm}=0} x'_{idm}\right) \qquad (5)$$

Model estimation

A severe drawback of the standard MTMM approach is that at least three repeated measurements are required from the same respondents. Saris et al. (2004) suggest solving this problem with a split-ballot MTMM design. The split-ballot MTMM design employs various random samples of the same population, but in each of the samples only two methods are used. In the authors' three-group design, the respondent group A uses methods 1 and 2, group B uses methods 2 and 3 and group C uses methods 3 and 1. The only statistical implication is that this design requires estimators which can handle missing data. The ML estimation of SEM with missing data was discussed by Arbuckle (1996). A variant of the ML estimator with missing data described by Yuan and Bentler (2000) and Arminger and Sobel (1990) is robust to non-normality under certain conditions and is the estimation method we use, as implemented in the MLR option of the Mplus6.12 program (Muthén and Muthén 2010). If all three measurements are administered to all respondents, then standard SEM estimation procedures can be used.

Extensions to the method

In this section, we extend the method in Coenders et al. (2011) by considering alternative log-ratio transformations with better interpretational and statistical properties and by adding covariates and outcome variables into the model.

Alternative transformations

In certain cases, other log-ratio transformations are more interpretable than the alr while offering some statistical and geometrical advantages. They are also more flexible in that the

denominator does not have to be the same component in all ratios. A good example is compositions with 4 components which can be meaningfully grouped in pairs. In this article, we will be dealing with 4 network components: x_{i1m}=partner, x_{i2m}=other family members, x_{i3m}=friends and x_{i4m}= other non-family members. The first two constitute a family pair and the second two a non-family pair. A meaningful log-ratio transformation is:

- $y_{i1m} = \ln\left(\dfrac{x_{i1m}}{x_{i2m}}\right) = \ln(x_{i1m}) - \ln(x_{i2m})$ is a ratio of partner over other family members

  within the family pair;

- $y_{i2m} = \ln\left(\dfrac{x_{i3m}}{x_{i4m}}\right) = \ln(x_{i3m}) - \ln(x_{i4m})$ is a ratio of friends over other non-family

  members within the non-family pair;

- $y_{i3m} = \ln\left(\dfrac{\sqrt{x_{i1m}x_{i2m}}}{\sqrt{x_{i3m}x_{i4m}}}\right) = \dfrac{1}{2}\ln(x_{i1m}) + \dfrac{1}{2}\ln(x_{i2m}) - \dfrac{1}{2}\ln(x_{i3m}) - \dfrac{1}{2}\ln(x_{i4m})$ is a ratio of the

  geometric mean of the family pair over the geometric mean of the non-family pair.

This log-ratio transformation happens to be proportional to an isometric log-ratio transformation (ilr; see Egozcue et al. 2003) because it can be expressed as:

$$\mathbf{y} = \mathbf{\Psi}\ln(\mathbf{x}) =$$

$$\begin{pmatrix} y_{i1m} \\ y_{i2m} \\ y_{i3m} \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} \ln(x_{i1m}) \\ \ln(x_{i2m}) \\ \ln(x_{i3m}) \\ \ln(x_{i4m}) \end{pmatrix} \tag{6}$$

where the $\mathbf{y}$ vector has D-1 dimensions, and the rows of the $\mathbf{\Psi}$ matrix have zero means and scalar products.[1]

---

[1] In a genuine ilr transformation, the rows of $\mathbf{\Psi}$ also have unit sums of squares. Thus, $\mathbf{\Psi}$ defines an orthonormal coordinate base and preserves distances. In order to achieve this, the genuine ilr transformation applies a scale change to the log-ratios which is called balance (Egozcue & Pawlowsky-Glahn, 2005). For instance, the first balance in our example is $y_{i1m} = \dfrac{1}{\sqrt{2}}\ln\left(\dfrac{x_{i1m}}{x_{i2m}}\right) = \dfrac{1}{\sqrt{2}}\ln(x_{i1m}) - \dfrac{1}{\sqrt{2}}\ln(x_{i2m})$ because $\left(\dfrac{1}{\sqrt{2}}\right)^2 + \left(\dfrac{1}{\sqrt{2}}\right)^2 = 1$. Since the SEMs used in this article are scale invariant (Cudeck, 1989), we choose not to use the scale change to build balances but to use the log-ratios as they are. This makes them more easily interpretable.

The advantages of ilr transformations are discussed in Egozcue et al. (2003) and Egozcue and Pawlowsky-Glahn (2005; 2006). Among others, ilr transformations do away with spurious all-positive or all-negative covariances among components. However, ilr covariances cannot be interpreted as covariances among components but have yet to be interpreted according to the definition of the ratios as combinations of components.

Other examples of interpretable transformations for network compositions which are proportional to ilr transformations are given in the appendix.

Inclusion of covariates and outcome variables

As described in the Introduction, certain individual background and attitudinal variables are known to have an effect on social network compositions and can be included in the standard way in SEM models as covariates (see the path diagram in Figure 2). Effects of these covariates cannot be understood on components but on log-ratios of components, hence the importance of building interpretable log-ratios.

Social network compositions are also known to affect certain outcome variables. The interpretation of effects on outcomes is slightly more complicated. One would be tempted to interpret the effects of increasing a component while keeping all other components constant. To begin with, this type of interpretation is impossible for compositional data because a component can only increase if some other(s) decrease. However, it can be feasible to increase one log-ratio while keeping the remaining log-ratios constant. For this purpose, great care is needed in interpreting the results according to the manner in which log-ratios are built. In our example:

- The effect of y_{i1m} on an outcome variable refers to the effect of increasing x_{i1m} (partner) and reducing x_{i2m} (remaining family) in the same proportion so that the product $x_{i1m} x_{i2m}$ remains constant, as well as both x_{i3m} and x_{i4m}.

- The effect of y_{i2m} on an outcome variable refers to the effect of increasing x_{i3m} (friends) and reducing x_{i4m} (other non-family members) in the same proportion so that the product $x_{i3m} x_{i4m}$ remains constant, as well as both x_{i1m} and x_{i2m}.

- The effect of y_{i3m} on an outcome variable refers to the effect of increasing both x_{i1m} and x_{i2m} in the same proportion while reducing both x_{i3m} and x_{i4m} in the same proportion.

It is important to note that the estimates of these effects will be free of measurement error bias since measurement error variances and covariances are accounted for by the θ parameters.

It is of course possible to add direct effects from the predictor to the outcome variables (not represented in Figure 2 for simplicity).

{Figure 2 here}

Illustration

In this section, we present an illustration of the proposed method. We first describe the data, type of network, components and measurement methods. Then we describe a few appropriate covariates and outcomes which are common in the literature and show how to reformulate the findings in the literature in terms of log-ratios. We then present the results and show how they should be interpreted.

Data

The focus of this study is compositions of egocentered networks in the city of Ljubljana collected by Kogovšek et al. (2002). Networks were defined as personal social support networks understood as a multidimensional construct. Kogovšek et al. (2002) used the typology of Cohen and Wills (1985), which distinguishes instrumental support, informational support, emotional support and social companionship. The names of alters were obtained with so-called name generator questions, asking whom the respondent would ask for help in a number of situations (small jobs in or around the house, advice about a major change in life, talk about personal matters, borrowing money, and the like). Respondents could mention as many names as they wished.

Once the names had been collected, several additional questions (name interpreters) were posed to find out about the type of relationship between the ego and alters (family

members, friends etc.) and an evaluation of the relationship (perceived importance of each alter to the ego). The characteristics measured through the name interpreters can be used to build network components and certain outcome variables. The questionnaire also included questions on covariates related to the respondent's background and attitudes.

Kogovšek et al. (2002), Kogovšek and Ferligoj (2005) and Coenders et al. (2011) already studied the measurement quality of these data, from non-compositional and compositional perspectives. The methodological focus of the Kogovšek et al. (2002) study offers a limited number of potential outcomes and covariates of social network composition, although we think they are enough for the purpose of illustrating the methodology.

The network components used in this article are the same as in the example in the Extensions to the Method section (partner, other family members, friends, other non-family members). The methods used in the data set of Kogovšek et al. (2002) and in this article are:

- 1: A face-to-face survey with name interpreters ordered by alters (respondents answer all questions about an alter before moving to the next alter).
- 2: A telephone survey with name interpreters ordered by alters.
- 3: A telephone survey with name interpreters ordered by name interpreters (respondents answer a question for all alters before moving to the next question).

Kogovšek et al. (2002) designed their study as a split-ballot MTMM design (Table 1).

{Table 1 here}

The data were collected between March and June 2000 by computer-assisted telephone interviews (CATI) and computer-assisted personal interviews (CAPI) for a representative sample of 1,033 inhabitants of the city of Ljubljana, Slovenia. 59% were women and 41% men. With regard to education, 9% of respondents had primary school or less, 13% had professional training, 43% had a 4-year secondary school and 35% had a higher education.

Covariates

The purpose of this illustration is to relate network composition to some background and attitudinal covariates and outcomes. The following covariates were included in the model:

- Gender was introduced as a dummy coded variable (0=male; 1=female). According to the literature, a partner tends to be a more important source of support for men than for women, who tend to seek support, especially emotional, more with friends (e.g., Edwards et al., 1998). Women also tend to have more kin relations in their networks (Fischer, 1982; Iglič, 1988; Marsden, 1987; van der Poel, 1993).

- Education was introduced as a dummy coded variable (0=professional or lower; 1=secondary or higher). According to the literature, the percentage of friends tends to increase with education (e.g., Cornwell et al., 2009; Fischer, 1982; Iglič, 1988; Marsden, 1987; van der Poel, 1993).

- Age was expressed in decades in order to obtain reasonable-sized effects. The non-linear effect was considered by also including age squared (age was mean-centered in order to reduce collinearity with age squared). In general, the literature reports the percentage of family members within a network to increase with age, although the relationship is not necessarily linear. It is often relatively large in the youngest and oldest age groups and may be relatively low in middle age groups, especially on account of friends and co-workers who retire or die in later stages (e.g., Ajrouch et al., 2001; Cornwell et al., 2009; Dremelj, 2003; Hoffmeyer-Zlotnik, 1990; Iglič, 1988; Marsden, 1987; van der Poel, 1993). Other ties gain importance with age, especially neighbors and distant family.

- Extraversion was measured with a summated scale (nine Likert-type items ranging from 1 to 5[2]). The literature shows that extraverts have larger networks with diverse compositions and a smaller family component (e.g., Doeven-Eggens et al., 2008; Newcomb and Keefe, 1997; Ying, 2002).

Following the major findings in the literature, we formulate the hypotheses below. Since compositional data only convey information on component size relative to other

---

[2] The scale is part of the larger 50-item Big Five Factor Markers taken from the International Personality Item Pool webpage (http://ipip.ori.org/; also see Goldberg, 1992). The extraversion dimension in the original scale had 10 items. The item »Don't like to draw attention to myself« was left out of our analyses owing to the extremely low correlations with all other variables (communality in a factor analysis was only 0.072). Cronbach's alpha was 0.773; Heise and Bohrnstedt's (1970) omega was 0.781 (omega is a reliability measure based on the factor analysis model and assumes congeneric items, which is a less strict assumption than tau-equivalent items, as implied by alpha).

components (in other words, ratios), meaningful hypotheses have to be reformulated with ratios in mind, in other words, in the context of a relationship among the network components:

- For women, family is more prominent within the network in relation to a partner (H1a) and family is more prominent in relation to other non-family relationships (H1b).

- For more educated respondents, friends are more prominent within the network in relation to other non-family members (H2).

- For older respondents, other family members are more prominent within the network than a partner (H3a) and other non-family members are more prominent in relation to friends (H3b). For both the youngest and the oldest respondents, family members are more prominent in relation to non-family members (H3c).

- In the networks of extraverted respondents, a partner is more prominent in relation to the family part of the network (H4a) and friends are more important in relation to other non-family members (H4b).

Outcome variables

The following outcome variables were included in the model:

- Multiplexity of ties: in how many name generators each alter is mentioned (average for all alters related to the ego). This can be understood as the average number of support functions played by all alters within an ego's network. According to the literature, close network members more often provide different types of social support (i.e., are multiplex, Dremelj, 2003; Laireiter and Baumann, 1992; Wellman and Wortley, 1990). Family relationships (especially the closest kin like parents and children) tend to be very important for providing multidimensional social support. Friends are also strong providers of all types of support (Dremelj, 2003; Norris and Tindale, 1994).

- Strength of ties: average perception of importance for all alters related to the ego[3]. For obvious reasons, network components differ in tie strength. Close or strong ties usually include a partner, close friends and close family, while less close or weak ties usually include co-workers, neighbors, distant family and the like. The social network composition thus affects the tie strength.

From the literature, the following hypotheses were reformulated in terms of ratios with regard to the outcome variables:

- Respondents where the partner is more prominent in relation to other family members (H5a) and where friends are more prominent in relation to other non-family members (H5b) and where family is more prominent in relation to non-family (H5c) tend on average to have a higher level of multiplexity in the network.

- Respondents where the partner is more prominent in relation to family (H6a), where friends are more prominent in relation to other non-family members (H6b) and where family members are more prominent in relation to non-family members (H6c) on average have a greater degree of closeness to their network members.

Results

Table 2 shows the basic descriptive statistics of the included variables. The x'_{idm} scores are percentages of network components before any log-ratio transformation is applied. The y_{idm} scores are log-ratios computed as in the example in the Extensions to the Method section. As shown by the x'_{idm} scores, other family members and friends are on average the biggest components for all methods. As a result, the first log-ratio (partner/other family members) is on average negative for all methods and the second (friends/other non-family members) is positive. The family and non-family parts of the network are quite balanced and the third log-ratio is close to zero for all methods. As is often the case, the x'_{idm} scores have quite large skewness and kurtosis, which is much reduced for log-ratios.

{Table 2 here}

---

[3] Average response to the name interpreter "How important is this person in your life? Please describe how close you feel on a scale from 1 to 5, where 1 means not important and 5 means very important".

The model was built as in Figure 2. Since some significant direct effects of covariates on outcomes were found, they are included in the model for control purposes, even if they are not specified in any of the hypotheses. After including these direct effects, the fit of the model was excellent. The Chi-Square test of model fit was 63.573 with 57 degrees of freedom (p-value=0.2561). The 90 Percent C.I. for the root mean square error of approximation was (0.000; 0.023). The CFI was 0.997 and the TLI 0.993. The standardized root mean square residual was 0.020. As recommended by Saris et al. (2009), even if the goodness-of-fit measures were acceptable, we explored alternative ways of improving the model by adding parameters. The expected standardized changes for candidate parameters to be added were either small or nonsensical, which led us to maintain the model.

A complete treatment of the interpretation of the measurement part of the model can be found in Coenders et al. (2011), therefore we will not interpret it extensively here. As in that article, the best measurement quality of network composition (here indicated by the standardized trait loadings in Table 3) is found in method 1 (Face to face) for all traits. Within the telephone method, no ordering seems to be consistently better. The second trait (ratio of friends over other members) has the lowest measurement quality for all methods. Error variances are omitted because they do not provide any new information (standardized error variances are equal to one minus the squared standardized trait loadings). Error covariances are included only with methodological purposes (Coenders et al., 2011) and are not interpreted.

{Table 3 here}

The part of the model which is most relevant to this article is summarized in Table 4: the estimates of the effects of covariates on network components and the effects of network components on outcome variables. The two coefficients relating age and age squared are best jointly interpreted later on in the context of Figure 3, which plots the non-linear relationship of age on all three log-ratios.

With regard to the partner over family ratio (t_1, proportion of explained variance R^2=0.131), it can be seen that the effects of extraversion, age and gender are statistically significant. Extraverted respondents tend to have support networks in which the partner is relatively

more prominent than the remainder of the family (H4a confirmed). For women, the family part of the network is more prominent than the partner (H1a confirmed).

Regarding the ratio of friends over other non-family members (t_2, $R^2=0.229$), the effects of extraversion, age and education are statistically significant. More extraverted respondents tend to have more friends compared to other non-family members in their networks (H4b confirmed). Respondents with a higher education tend to have more friends than other non-family members in their networks (H2 confirmed).

Regarding the prediction of the family over non-family ratio (t_3, $R^2=0.013$), only the effect of age is statistically significant. Gender had no statistically significant effect (H1b not confirmed).

{Table 4 here}

With regard to the effects of network composition on tie multiplexity ($R^2=0.123$), the partner over other family ratio and the family over non-family ratio are statistically significant. Respondents, where the proportion of the partner is increasing on account of other family members decreasing, have a higher level of multiplexity of ties (H5a confirmed). Respondents, with an increasing proportion of family on account of a decreasing proportion of non-family members, also have higher multiplexity (H5c confirmed). The effect of the ratio of friends over other non-family was not statistically significant (H5b not confirmed).

Regarding the effect of the network composition on the importance of ties within the support network ($R^2=0.222$), only the effect of the family over non-family ratio is statistically significant – respondents where the proportion of family increases on account of a decreasing proportion of non-family have a higher level of importance of their alters (H6c confirmed). The effects of the partner over other family ratio (H6a) and of the friends over other non-family ratio (H6b) were not statistically significant. One possible explanation for failing to confirm H5b, H6a and H6b is that we did not distinguish between close and distant friends or between close and distant family, while only the close parts are expected to provide multiple support functions and to be perceived as important.

{Figure 3 here}

Regarding the effect of age on the ratio between the partner and family, the non-linearity of the effect is obvious (Figure 3). The partner is the most important in relation to other family members in middle age categories and decreases toward both the younger and older age categories. Accordingly, H3a cannot be confirmed univocally. With regard to the ratio of friends over other non-family members, it can be seen that the proportion of friends tends to decrease with age (H3b confirmed), although friends still remain the more important component of the two (the log-ratio remaining positive). The ratio of family over non-family tends to be stable with a slight tendency to increase in the youngest and oldest age categories, therefore giving support for H3c.

Discussion

From the substantive point of view, the limitation of this article is that the aim of the Kogovšek et al. (2002) survey was methodological (testing the reliability and validity of social network measurements) and not substantive. We were therefore somewhat limited in finding the most appropriate covariates and outcomes. The results have to be understood as simply an illustration of the proposed methodology. We believe we have shown it is possible to successfully build models that include testing measurement as well as estimating unbiased substantive effects on key concepts, while treating compositional data in a proper, simple and interpretable manner. It is then not surprising that most of our results are in line with previous research on the topic, which provides support for the appropriateness of the method.

With regard to the statistical model itself, we benchmarked it with a variety of log-ratio transformations (including the old alr used in Coenders et al., 2011) and alternative sets of covariates and outcome variables, with equivalent results with varying interpretational difficulty. The results regarding measurement quality were always stable and coincident with the results in Coenders et al. (2011), namely, the face-to-face method had the highest quality estimates of network composition. We also see a substantive advantage in using the ilr instead of the alr log-ratio transformation. It makes more sense to build log-ratios of substantively compatible network components (partner over family, friends over non-family, family over non-family) than to compare all components in relation to the last. This is shown by the fact that the effects described in the literature review were very easy to translate into

testable hypotheses involving ilr log-ratios. Ilr log-ratio transformations are not any more difficult to construct than alr, as shown in the appendix.

The advantage of using a SEM with multiple indicators is not their ability to deal with compositional data but their ability to deal with measurement error attenuation bias. The basic method for dealing with network compositions outlined in this paper can also be applied to simpler statistical techniques than SEM. This method includes the way of dealing with zeroes in the components, and the way of transforming the original D components into a set of D-1 interpretable log-ratios. The latter can be achieved both by means of ilr and alr, with the key reason for choosing one over the other being the interpretability of the obtained log-ratios according to the conceptual meaning of the components. Once zeroes have been replaced and log-ratios computed, the researcher can proceed as in standard statistical analysis. If the researcher's aim is to relate network compositions to a set of outcomes and covariates, in the usual case in which only one measurement of network composition is available, linear regressions are an appropriate tool. The researcher can thus fit the linear regression models of each log-ratio on the covariate set and the linear regression models of each outcome variable on the log-ratio set. Standard estimation methods for linear regression models can be used. Even ordinary least squares can be used whenever the usual diagnostic tools show that the method's assumptions hold, which occurs often with log-ratios and never with raw proportions (Aitchison, 1986). The up to recently standard practice of treating raw proportions of components with standard statistical tools (e.g., Ajrouch et al., 2001; Cornwell et al., 2009) leads at the least to wrong statistical tests and confidence intervals because of non-normality, of limited variable distribution, and of heteroskedasticity.

Acknowledgements

Appendix. Other interpretable log-ratio transformations which are proportional to ilr

An interpretable transformation is very easy to build whenever components can be ordered. If we have the following components in a study of personal support networks: x_{i1m}=lives in the same neighborhood, x_{i2m}=lives in the same city, x_{i3m}=lives further away, then a meaningful log-ratio transformation is:

- The first log-ratio $y_{i1m} = \ln\left(\dfrac{x_{i2m}}{x_{i1m}}\right) = -\ln(x_{i1m}) + \ln(x_{i2m})$ is a ratio of the second category (same city) over the first (same neighborhood).

- Each following log-ratio $y_{i2m} = \ln\left(\dfrac{x_{i3m}}{\sqrt{x_{i1m}x_{i2m}}}\right) = -\dfrac{1}{2}\ln(x_{i1m}) - \dfrac{1}{2}\ln(x_{i2m}) + \ln(x_{i3m})$ is a ratio of the next ordered category over the geometric mean of all lower categories.

It is proportional to an ilr transformation because the **Ψ** matrix has D-1 rows with zero means and scalar products:

$$\Psi = \begin{pmatrix} -1 & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{pmatrix}$$

In general, an interpretable log-ratio transformation which is proportional to an ilr transformation is easy to compute whenever there is an interpretable sequential binary partition of components into pairs of groups of components according to the conceptual similarity of the components. These partitions start by dividing components into two clusters and then continue by subdividing one of the clusters into two until each component constitutes its own cluster. D components always involve D-1 partitions. These partitions are best understood as a partition tree or dendrogram (Pawlowsky-Glahn & Egozcue, 2011). If we have the following components in a study of personal support networks: x_{i1m}=close friends, x_{i2m}=non close friends, x_{i3m}=neighbors, x_{i4m}=co-workers, x_{i5m}=co-members of associations, then a meaningful sequential partition is given in the conceptual classification tree below.

{Figure 4 here}

A meaningful log-ratio transformation takes ratios of the geometric means of the two component clusters at each partition. Numerators and denominators are interchangeable. For instance, we may have:

- $y_{i1m} = \ln\left(\dfrac{x_{i2m}}{x_{i1m}}\right)$ is a ratio of other friends over close friends.

- $y_{i2m} = \ln\left(\dfrac{x_{i4m}}{x_{i5m}}\right)$ is a ratio of co-workers over co-members of associations.

- $y_{i3m} = \ln\left(\dfrac{\sqrt{x_{i4m}x_{i5m}}}{x_{i3m}}\right)$ is a ratio of colleagues over neighbors.

- $y_{i4m} = \ln\left(\dfrac{\sqrt{x_{i1m}x_{i2m}}}{\sqrt[3]{x_{i3m}x_{i4m}x_{i5m}}}\right)$ is a ratio of friends over non-friends.

It is proportional to an ilr transformation because the $\mathbf{\Psi}$ matrix has D-1 rows with zero means and scalar products:

$$
\mathbf{\Psi} = \begin{pmatrix}
-1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & -1 \\
0 & 0 & -1 & \frac{1}{2} & \frac{1}{2} \\
\frac{1}{2} & \frac{1}{2} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3}
\end{pmatrix}
$$

References

Aitchison, J., 1986. The Statistical Analysis of Compositional Data. Monographs on Statistics and Applied Probability. Chapman & Hall, London. (Reprinted in 2003 with additional material by The Blackburn Press).

Ajrouch, K. J., Antonucci, T. C., Janevic, M. R., 2001. Social networks among blacks and whites: The interaction between race and age. Journals of Gerontology 56B, S112–S118.

Arbuckle, J. L., 1996. Full information estimation in the presence of incomplete data. In: Marcoulides, G. A., Schumacker, R. E., (Eds.), Advanced Structural Equation Modeling. Erlbaum, Mahwah, NJ, pp. 243–277.

21

Arminger, G., Sobel, M. E., 1990. Pseudo maximum likelihood estimation of mean and covariance structures with missing data. Journal of the American Statistical Association 85, 195–203.

Campbell, D. T., Fiske, D. W., 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin 56, 81–105.

Chan, W., 2003. Analyzing ipsative data in psychological research. Behaviormetrika 30, 99–121.

Cheung, M. W. L., 2004. A direct estimation method on analyzing ipsative data with Chan and Bentler's (1993) method. Structural Equation Modeling 11, 217–243.

Coenders, G., Hlebec, V., Kogovšek, T., 2011. Measurement quality in indicators of compositions. A compositional multitrait-multimethod approach. Survey Research Methods 5, 63–74.

Coenders, G., Saris, W. E., 2000. Testing nested additive, multiplicative and general multitrait-multimethod models. Structural Equation Modeling 7, 219–250.

Cohen, S., Wills, T. A., 1985. Stress, social support, and the buffering hypothesis. Psychological Bulletin 98, 310–357.

Cornwell, B., Schumm, L. P., Laumann, E. O., Graber, J., 2009. Social networks in the NSHAP study: rationale, measurement, and preliminary findings. Journals of Gerontology 64B, i47–i55.

Coromina, L., Coenders, G., 2006. Reliability and validity of egocentered network data collected via Web. A meta-analysis of multilevel multitrait multimethod studies. Social Networks 28, 209–231.

Cudeck. R., 1989. Analysis of correlation matrices using covariance structure models. Psychological Bulletin 105, 317–327.

Doeven-Eggens, L., De Fruyt, F., Jolijn Hendriks, A. A., Bosker, R. J., Van der Werf, M. P. C., 2008. Personality and personal network type. Personality and Individual Differences 45, 689–693.

Dremelj, P., 2003. Sorodstvene vezi kot vir socialne opore posameznikov. Družboslovne razprave 43, 149–170.

Edwards A. C., Nazroo, J. Y., Brown, G. W., 1998. Gender differences in marital support following a shared life event. Social Science & Medicine 46, 1077–1085.

Egozcue, J. J., Pawlowsky-Glahn, V., 2005. Groups of parts and their balances in compositional data analysis. Mathematical Geology 37, 795–828.

Egozcue, J. J., Pawlowsky-Glahn, V., 2006. Simplicial geometry for compositional data. In: Buccianti, A., Mateu-Figueras, G., Pawlowsky-Glahn, V. (Eds.), Compositional Data Analysis in the Geosciences: From Theory to Practice. Geological Society, London, pp. 145–159.

Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., Barceló-Vidal, C., 2003. Isometric logratio transformations for compositional data analysis. Mathematical Geology 35, 279–300.

Ferligoj, A., Hlebec, V., 1999. Evaluation of social network measurement instruments. Social Networks 21, 111–130.

Fischer, C. S., 1982. To Dwell among Friends. University of Chicago Press, Chicago.

Goldberg, L. R., 1992. The development of markers for the Big-Five factor structure. Psychological Assessment 4, 26–42.

Heise, D.R., Bohrnstedt, G.W., 1970. Validity, invalidity and reliability. In: Borgatta, E.F., Bohrnstedt, G.W. (Eds). Sociological Methodology. Jossey-Bass, San Francisco pp. 104–129.

Hoffmeyer-Zlotnik, J. H. P., 1990. The Mannheim comparative network research. In: Weesie, J., Flap, H. (Eds.). Social Networks Through Time. ISOR, Utrecht, pp. 265–279.

Iglič, H., 1988. Ego-centrične socialne mreže. Družboslovne razprave 5, 82–93.

Kogovšek, T., 2006. Reliability and validity of measuring social support networks by web and telephone. Metodološki zvezki 3, 239–252.

Kogovšek, T., Ferligoj, A., 2005. Effects on reliability and validity of egocentered network measurements. Social Networks 27, 205–229.

Kogovšek, T., Ferligoj, A., Saris, W. E., Coenders, G., 2002. Estimating the reliability and validity of personal support measures: full information ML estimation with planned incomplete data. Social Networks 24, 1–20.

Laireiter, A., Baumann, U., 1992. Network structures and support functions – theoretical and empirical analyses. In: Veiel, H. O. F., Baumann, U. (Eds.). The Meaning and Measurement of Social Support. Hemisphere, New York pp. 33–55.

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

Marsh, H. W., 1989. Confirmatory factor analysis of multitrait-multimethod data: Many problems and few solutions. Applied Psychological Measurement 13, 335–361.

Martín-Fernández, J. A., Barceló-Vidal, C., Pawlowsky-Glahn, V., 2003. Dealing with zeros and missing values in compositional data sets using non-parametric imputation. Mathematical Geology 35, 253–278.

Muthén, L. K., Muthén, B., 2010. Mplus User's Guide. Sixth Edition. Muthén and Muthén, Los Angeles, CA.

Newcomb, M. D., Keefe, K., 1997. Social support, self-esteem, social conformity, and gregariousness: Developmental patterns across 12 years. In: Pierce, G. R., Lakey, B., Sarason, I. G., Sarason, B. R. (Eds.). Sourcebook of Social Support and Personality. Plenum Press, New York pp. 303–333.

Norris, J. E., Tindale, J. A., 1994. Among Generations: The Cycle of Adult Relationships. W.H. Freeman and Company, New York.

Pawlowsky-Glahn, V., Buccianti, A., (eds.), 2011. Compositional Data Analysis. Theory and Applications. Wiley, New York.

Pawlowsky-Glahn, V., Egozcue, J. J., 2011. Exploring compositional data with the CoDa-dendrogram. Austrian Journal of Statistics 40, 103–113.

Pierotti, M. E. R., Martín-Fernández, J. A., Seehausen, O., 2009. Mapping individual variation in male mating preference space: multiple choice in a colour polymorphic cichlid fish. Evolution 63, 2372–2388.

Poel, M. G. M. van der., 1993. Personal Networks. Swets and Zeitlinger, Lisse.

Saris, W. E., Gallhofer, I. N., 2007. Design, Evaluation, and Analysis of Questionnaires for Survey Research. Wiley, New York.

Saris, W. E., Satorra, A., Coenders, G., 2004. A new approach to evaluating the quality of measurement instruments: the split-ballot MTMM design. Sociological Methodology 34, 311–347.

Saris, W.E., Satorra, A., van der Veld, W.M., 2009. Testing structural equation models or detection of misspecifications. Structural Equation Modeling 16, 561–582.

Schumacker, R. E., Lomax, R. G., 1996. A Beginner's Guide to Structural Equation Modeling. Lawrence Erlbaum, Mahwah, NJ.

Thió-Henestrosa, S., Martín-Fernández, J. A., 2005. Dealing with compositional data: the freeware codapack. Mathematical Geology 37, 773–793.

Wellman, B., Wortley, S., 1990. Different strokes from different folks. American Journal of Sociology 96, 558–588.

Ying., Y. W., 2002. Formation of cross-cultural relationships of Taiwanese international students in the United States. Journal of Community Psychology 30, 45–55.

Yuan, K. H., Bentler, P. M., 2000. Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. Sociological Methodology 30, 165–200.

Table 1: The split-ballot MTMM design of the study*

Group	*n*	First measurement with method:	Second measurement with method:
A	320	Face to face ordered by alters	Telephone ordered by alters
B	402	Telephone ordered by alters	Telephone ordered by name interpreters
C	311	Face to face ordered by alters	Telephone ordered by name interpreters

*Original sample sizes in Kogovšek et al. (2002) are shown. Seven cases with network size equal to 1 or with missing data for the covariates were deleted for the analyses in this article

Table 2: Descriptive statistics

	Min	Max	mean	st.dev.	skewness	kurtosis
x'_{idm} scores (compositions with zeroes replaced %)						
x_{11} (partner, method 1)	1.32	87.50	12.41	11.62	1.95	6.91
x_{21} (other family, method 1)	1.47	90.63	32.94	21.02	0.42	-0.39
x_{31} (friends, method 1)	1.47	93.75	40.14	23.57	0.09	-0.82
x_{41} (other non-family, method 1)	1.47	91.67	14.51	17.05	1.94	3.94
x_{12} (partner, method 2)	1.04	45.83	11.17	9.28	1.13	0.93
x_{22} (other family, method 2)	1.25	91.67	34.10	21.45	0.43	-0.37
x_{32} (friends, method 2)	1.56	94.64	40.54	23.68	0.04	-0.82
x_{42} (other non-family, method 2)	1.47	91.67	14.20	16.56	1.88	3.58
x_{13} (partner, method 3)	1.04	85.00	11.20	9.87	1.64	4.98
x_{23} (other family, method 3)	1.25	93.75	33.72	20.32	0.53	-0.07
x_{33} (friends, method 3)	1.79	95.31	41.50	22.25	-0.02	-0.64
x_{43} (other non-family, method 3)	1.04	85.00	13.58	14.87	1.65	2.51
y_{idm} logratios						
y_{11} (partner/family, method 1)	-3,84	3,04	-1,07	1,40	0.49	-0.07
y_{21} (friends/other, method 1)	-3,54	3,82	1,29	1,68	-0.80	-0.03
y_{31} (family/non-family, method 1)	-3,01	3,22	-,10	1,08	-0.01	-0.01
y_{12} (partner/family, method 2)	-3,62	2,76	-1,17	1,29	0.40	-0.23
y_{22} (friends/other, method 2)	-3,92	3,98	1,31	1,64	-0.80	0.07
y_{32} (family/non-family, method 2)	-3,48	2,77	-,10	1,12	0.00	-0.05
y_{13} (partner/family, method 3)	-3,81	2,83	-1,21	1,26	0.41	-0.17
y_{23} (friends/other, method 3)	-2,92	4,11	1,40	1,53	-0.66	-0.20
y_{33} (family/non-family, method 3)	-3,30	3,30	-,12	1,09	0.01	0.23
covariates						
age	18.00	91.00	43.98	17.93	0.42	-0.86
extraversion Likert scale	1.00	5.00	3.55	0.83	-0.43	-0.31
outcome variables						
average number of support ties of alters	0.40	4.00	1.50	0.52	0.96	1.36
average importance of alters	1.67	5.00	4.04	0.56	-0.34	0.14

	Estimate	std. err.	t-value	p-value	Standardized estimate
Loadings					
λ_{11}†	1.000*				0.867
λ_{21}	1.000*				0.722
λ_{31}	1.000*				0.868
λ_{12}	0.885	0.059	15.008	0.000	0.828
λ_{22}	0.808	0.103	7.815	0.000	0.594
λ_{32}	0.918	0.066	13.867	0.000	0.797
λ_{13}	0.779	0.064	12.249	0.000	0.754
λ_{23}	0.906	0.108	8.377	0.000	0.712
λ_{33}	0.879	0.065	13.588	0.000	0.769
Error covariances‡					
θ_{211}†	-0.049	0.067	-0.731	0.465	-0.060
θ_{311}	-0.084	0.038	-2.243	0.025	-0.221
θ_{321}	0.101	0.051	1.981	0.048	0.160
θ_{212}	0.168	0.057	2.966	0.003	0.175
θ_{312}	-0.101	0.036	-2.802	0.005	-0.210
θ_{322}	0.010	0.047	0.218	0.827	0.012
θ_{213}	0.050	0.053	0.947	0.344	0.057
θ_{313}	-0.024	0.034	-0.696	0.486	-0.041
θ_{323}	0.082	0.043	1.922	0.055	0.109

* Constrained to 1 to fix the scale of the latent trait scores

†The last parameter subindex shows the method. The first subindex shows the trait (loadings) or the pair of traits (covariances).

‡Each error term also has a variance parameter, which is not reported for the sake of simplicity. Standardized error variances equal one minus the squared standardized loading.

	Estimate	std. err.	t-value	p-value	Standardized estimate*

Table 4: Model estimates. Relationship part

Equation predicting t_1 (partner over remaining family members) R^2=0.131

	Estimate	std. err.	t-value	p-value	Standardized estimate*
Extraversion (H4a)	0.112	0.049	2.280	0.023	0.077
Age (H3a)	0.044	0.026	1.738	0.082	0.066
age^2	-0.102	0.012	-8.348	0.000	-0.288
female (H1a)	-0.585	0.085	-6.890	0.000	-0.237
secondary-higher	-0.047	0.100	-0.476	0.634	-0.016

Equation predicting t_2 (friends over remaining non-family members) R^2=0.229

extraversion (H4b)	0.187	0.059	3.136	0.002	0.129
age (H3b)	-0.253	0.030	-8.297	0.000	-0.375
age^2	0.027	0.013	2.088	0.037	0.077
female	-0.010	0.091	-0.109	0.913	-0.004
secondary-higher (H2)	0.666	0.119	5.619	0.000	0.231

Equation predicting t_3 (family over non-family) R^2=0.013

extraversion	-0.071	0.040	-1.769	0.077	-0.062
age	0.002	0.019	0.118	0.906	0.004
age^2 (H3c)	0.022	0.011	2.025	0.043	0.080
female (H1b)	0.021	0.067	0.310	0.757	0.011
secondary-higher	-0.074	0.083	-0.892	0.372	-0.032

Equation predicting average number of support ties of alters (multiplexity) R^2=0.123

t_1 (H5a)	0.039	0.016	2.399	0.016	0.090
t_2 (H5b)	0.002	0.020	0.087	0.930	0.004
t_3 (H5c)	0.078	0.018	4.336	0.000	0.143
extraversion	0.016	0.019	0.829	0.407	0.026
age	-0.084	0.011	-7.459	0.000	-0.289
age^2	0.003	0.005	0.567	0.571	0.019
female	0.166	0.033	5.023	0.000	0.157
secondary-higher	-0.014	0.041	-0.340	0.734	-0.011

Equation predicting average importance of alters (strength) R^2=0.222

t_1 (H6a)	-0.022	0.016	-1.345	0.179	-0.048
t_2 (H6b)	0.004	0.020	0.196	0.845	0.009
t_3 (H6c)	0.216	0.022	9.674	0.000	0.372
extraversion	0.082	0.020	4.221	0.000	0.124
age	-0.003	0.011	-0.258	0.796	-0.009
age^2	0.007	0.005	1.269	0.204	0.043
female	0.219	0.034	6.476	0.000	0.194
secondary-higher	-0.181	0.040	-4.471	0.000	-0.136

Disturbance covariances†

t_1-t_2	0.011	0.062	0.173	0.863	0.009
t_1-t_3	-0.094	0.041	-2.286	0.022	-0.087
t_2-t_3	-0.072	0.049	-1.478	0.139	-0.072
number of ties-importance	0.067	0.008	8.404	0.000	0.279

* Standardized estimates of age and the dummy variables on education and gender cannot be interpreted in terms of standard deviation increases but only as measures of effect size.

† Each disturbance term also has a variance parameter, which is not reported for the sake of simplicity. Standardized disturbance variances equal one minus the R^2.
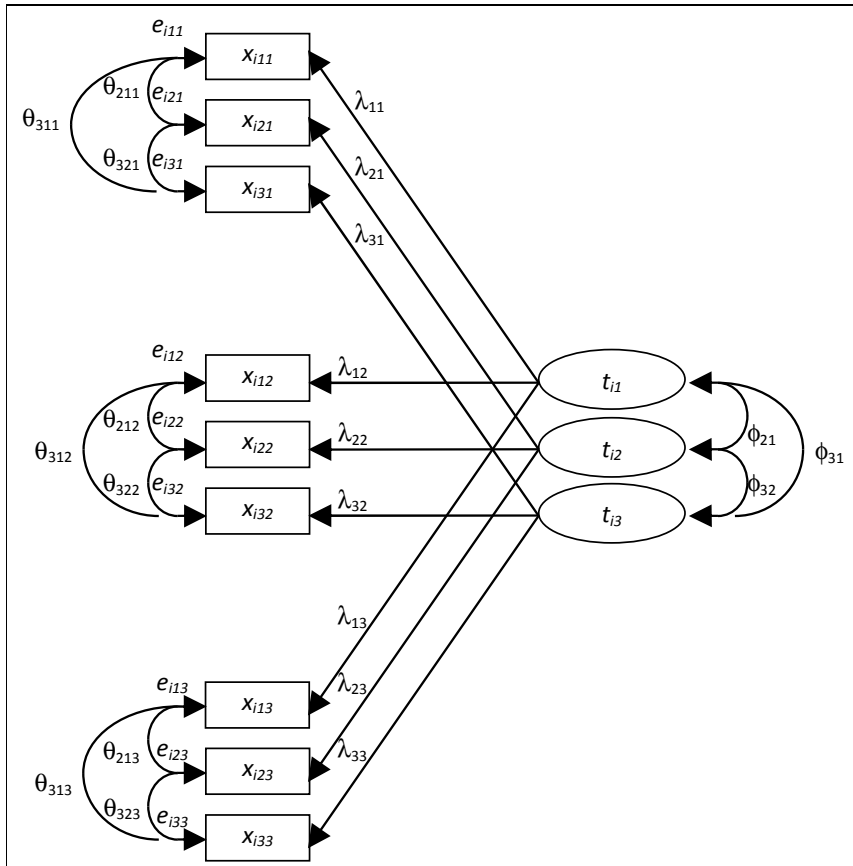
Figure 1

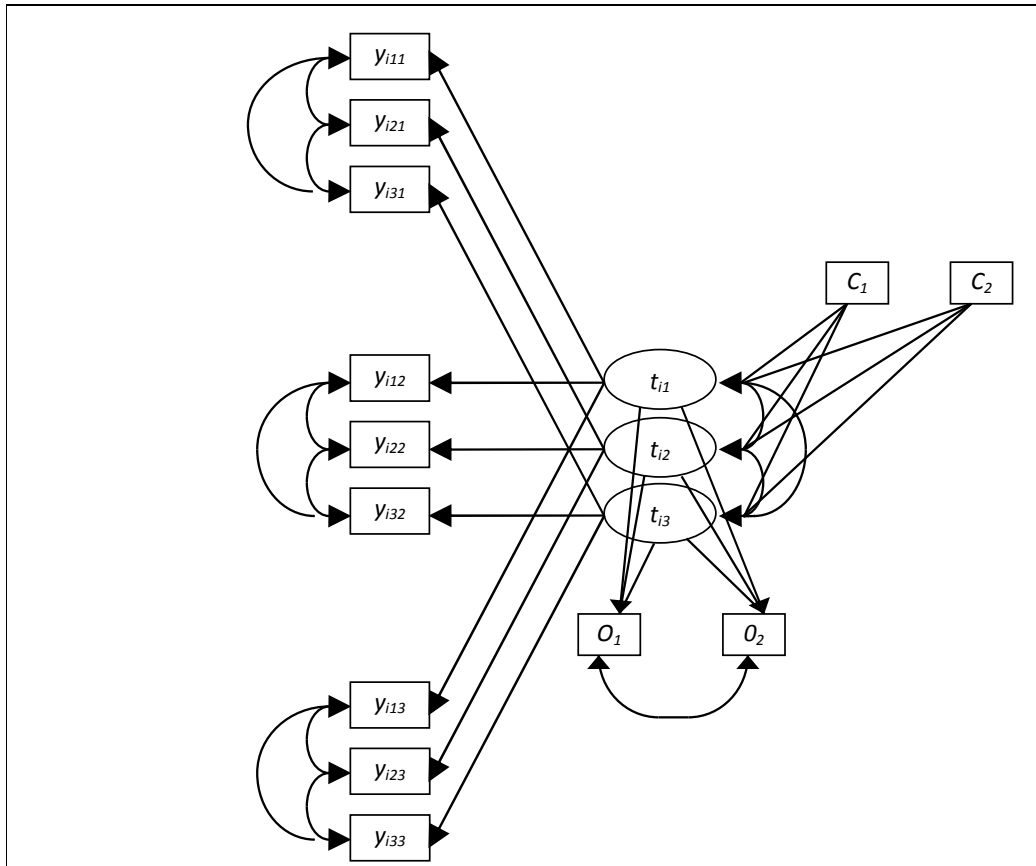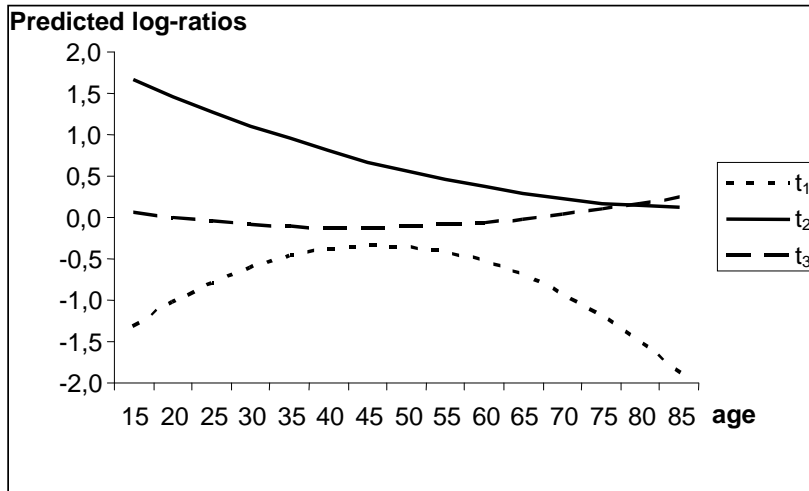**Figure 2**

Figure 3

Figure 4

Global partition:
friends vs. non friends

Partition non friends:
neighbors vs. colleagues

Partition friends:
close vs. non close

Partition colleagues:
co-workers vs. co-members

x_1 close friends

x_2 non close friends

x_3 neighbors

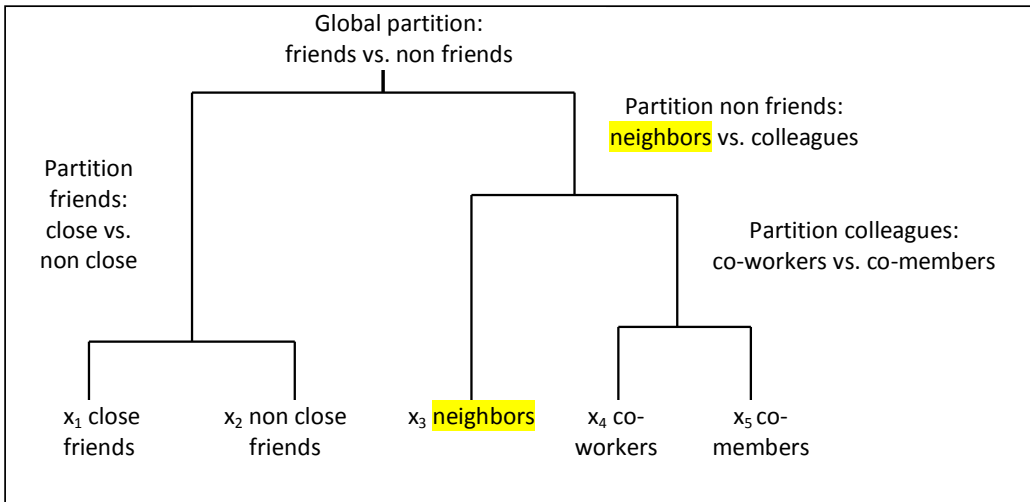x_4 co-workers

x_5 co-members

Figure captions

Figure 1: Path diagram of the CU model for MTMM designs. 3 traits and 3 methods

Figure 2: Path diagram of the extended CU model for compositional MTMM designs. 3 traits, 3 methods, 2 covariates (c_1, c_2), 2 outcomes (o_1, o_2)

Figure 3: Non linear effect of age on the log-ratios (t_1: partner over remaining family; t_2: friends over remaining non-family; t_3: family over non-family). Predicted log-ratios for males with professional or lower education with average extraversion

Figure 4: Example classification tree for a sequential binary partition