Social Networks 48 (2017) 46–56

Contents lists available at ScienceDirect

Social Networks

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

# Actor non-response in valued social networks: The impact of different non-response treatments on the stability of blockmodels

Anja Žnidaršič[a,*], Anuška Ferligoj[b], Patrick Doreian[b,c]

[a] University of Maribor, Faculty of Organizational Sciences, Kidričeva 55a, 4000 Kranj, Slovenia
[b] University of Ljubljana, Faculty of Social Sciences, Kardeljeva ploščad 5, 1000 Ljubljana, Slovenia
[c] University of Pittsburgh, Department of Sociology, USA

ARTICLE INFO

*Article history:*
Available online 2 September 2016

*Keywords:*
Valued network
Missing data
Actor non-response
Actor non-response treatment
Blockmodeling

ABSTRACT

Social network data usually contain different types of errors. One of them is missing data due to actor non-response. This can seriously jeopardize the results of analyses if not appropriately treated. The impact of missing data may be more severe in valued networks where not only the presence of a tie is recorded, but also its magnitude or strength. Blockmodeling is a technique for delineating network structure. We focus on an indirect approach suitable for valued networks. Little is known about the sensitivity of valued networks to different types of measurement errors. As it is reasonable to expect that blockmodeling, with its positional outcomes, could be vulnerable to the presence of non-respondents, such errors require treatment. We examine the impacts of seven actor non-response treatments on the positions obtained when indirect blockmodeling is used. The start point for our simulation are networks whose structure is known. Three structures were considered: cohesive subgroups, core-periphery, and hierarchy. The results show that the number of non-respondents, the type of underlying blockmodel structure, and the employed treatment all have an impact on the determined partitions of actors in complex ways. Recommendations for best practices are provided.



## 1. Introduction

A key advantage of valued relations in network data, where the strength, intensity, weight, or frequency is recorded instead of only the simplified presence of ties, is a better description of the real world relational data they are trying to capture. However, the recorded tie values are prone to having measurement errors. Not only the misspecified presence or absence of a tie possible (Holland and Leinhardt, 1973), also incorrect tie values can be recorded. Here, we will focus on one specific type of error where one or more actors provide no information regarding all other network members, i.e. actor non-response. Patterns of ties are important in revealing both macro and micro network structure. Misspecification of tie values could severely affect the obtained clusters of actors. To examine this, we investigated the stability of partitions of actors obtained from indirect blockmodeling of valued networks after seven actor non-response treatments are applied.

The paper is organized as follows: Section 2 discusses valued networks. Section 3 focuses on actor non-response. Section 4 presents suitable treatments for actor non-response in valued networks. The basic concept of indirect blockmodeling is presented in Section 5. Section 6 presents the simulation study with regard to the overall design of the simulations, three types of blockmodels used in simulation of network data, and provide the numerical summary of the simulation study. Results are presented in Section 7 by graphical and model representations. Section 8 presents conclusions with an emphasis on recommendations for researchers.

## 2. Valued networks

Valued network data have their ties measured in terms of magnitudes rather than only the presence or absence of a tie (Wasserman and Faust, 1998; Scott, 2013).[1]

For social networks, vertices represent social actors over which many social relations can be defined. In most settings, relations

* Corresponding author.
*E-mail address:* anja.znidarsic@fov.uni-mb.si (A. Žnidaršič).

[1] Some authors refer to those networks as *weighted networks*, but we regard value as a broader, more general, concept than a weight. For example, Horvath (2011) defined a weight as a real number between 0 and 1. Here the term '*valued*' is used instead of '*weighted*' when referring to networks, unless we cite from an original source.

http://dx.doi.org/10.1016/j.socnet.2016.06.001

can be operationalized to include values. Doing this is not always straightforward. Consider, for example, friendship'. Its meaning can range from 'acquaintance' to 'best friend' (Holland and Leinhardt, 1973). Operationalizing magnitude can be done in many ways. Girard et al. (2015) investigated factors influencing formation of a new social network in a new environment by gathering data from university freshmen on 4-point Likert type scale from 'university-acquaintance' to 'a close friend', while Van De Bunt et al. (1999) used a 6-point scale of friendship from 'best friend' to 'troubled relationship'. Grund (2012) studied interactions among soccer players in a longitudinal study. He emphasized the intensity of interactions, and not only the interaction itself, as being especially relevant for small teams where all team members are linked to each other.

A range of values for some relations can be defined unequivocally, for example, trade between countries and rail passengers traveling between cities. Other examples include the co-occurrence of keywords or collaboration of authors in bibliographic analysis of papers, and other networks which are calculated from 2-mode networks. More often, the range of values is established during the preparatory phase of data collection and not only by the question wording. A scale's level of measurement has to be precisely defined. For example, the frequency of collaboration between individuals or departments could be measured by the *number* of e-mails or face-face meetings. Alternatively, it could be estimated on a scale of frequencies ranging from 'never' to 'daily' or 'more often'. Potter et al. (2015) measured contacts between coworkers on 4-point ordinal scale of duration of daily contact (from up to 5 min to at least an hour to eight hours). Hlebec and Ferligoj (2002) used 5-point scales (from 0 (not at all) to 4 (certainly)) for three different sociometric questions: how likely would you borrow study materials from a classmate, likelihood of asking classmates for information about important study assignment, and how likely would you invite a classmate to a birthday party.

Apart from noting that the measurement of valued ties can vary greatly with many operationalizations being available, our intent is not to enter these debates about which measurement approach is the most appropriate for specific empirical situations. Instead, we assume that researchers can and do collect valued network data. These values can vary greatly in magnitude and range. We focus primarily on social relations among individuals. Operationally, our choice was to limit the tie values in our simulations of valued network data in Section 6 to the range 0 to 5, where 0s indicate absence of ties. In principle, this can be extended to other values and larger networks than those considered here.

Investigating the sensitivity of network properties in valued networks with introduced errors has been limited. Some attempts to evaluate regular blockmodeling structures on valued networks with *random* errors were carried out by Žiberna (2009). Páez et al. (2008) investigated impacts of erroneously omitting relevant ties and erroneously including irrelevant *subsets of* ties in the weight matrices for social influence analysis. They emphasized both situations as resulting in biased parameter estimates in network autocorrelation models. Given our concern with *actor non-response*, this study tackles a related problem for valued networks regarding blockmodeling (in Section 5).

## 3. Non-response in social networks

Actor non-response in social networks is one source of errors in network data (Žnidaršič et al., 2012). Each non-respondent in a network with $n$ actors implies $(n-1)$ missing ties. While all outgoing ties are missing for each non-respondent, the incoming ties are still observed. Fig. 1(a) presents a demonstration network with 15 actors. Suppose it was a real network to be measured. Suppose, further, three actors ($A_2$, $A_{10}$, and $A_{14}$ denoted with horizontal gray rectangles) had refused to respond. The actor response rate is 80% – the same as the overall relational response rate reported by Stork and Richards (1992). Fig. 1(b) represents the same network reorganized to have respondents in the upper rows and non-respondents at the bottom. The columns have been reorganized in the corresponding fashion. Missing ties consist of: (i) absent ties between non-respondents and respondents (bottom left part placed in a larger gray rectangle in Fig. 1(b)) and (ii) absent missing ties between non-respondents (right bottom part contained in the larger white square). This distinction for missing ties is important under the different treatments presented in Section 4.

While many studies report response rates the subsequent analyses most often deal only with the data from respondents about other respondents. In effect, this is a data collection imposition raising the well known boundary specification problem (Laumann et al., 1983) because the effective network boundary excludes non-respondents. Studies dealing seriously with boundary problems for networks remain quite rare. Doreian and Woodard (1994) discussed another variant where an 'official' list of the relevant organizations for an inter-organizational study left out many relevant organizations, organizations that were included subsequently by an expanding selection strategy. In general, omitting units is consequential. See also Kossinets (2006), Wang et al. (2012). The obvious question is: does this matter as far as the results of the subsequent analyses?

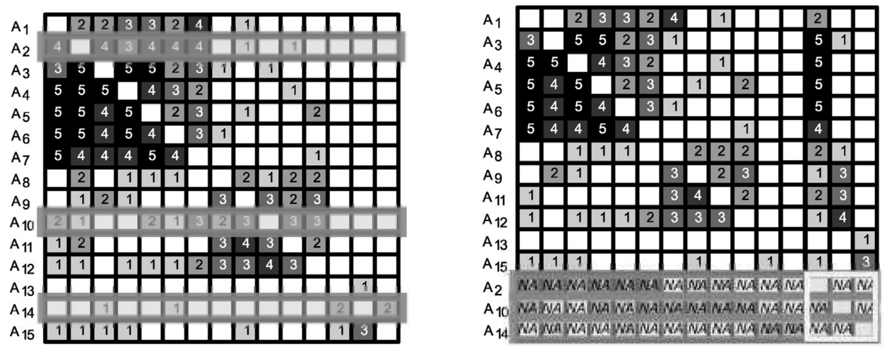
When the actor response rates are reported, literature reviews reveal a broad range in the number of reported non-respondents. Based on sample of 59 networks, Costenbader and Valente (2003) reported response rates between 51% and 100%.[2] Stork and Richards (1992) reported response rates varying from 65% to 90% of actors. Johnson et al. (2012) reported a 57% overall response rate in a sociometric survey (on friendship, advice, and information flow networks) among employees in Central European bank before investigating only three departments with the highest response rates varying from 63% to 71%. Ellwardt et al. (2012) reported on three waves of a longitudinal study of gossip and friendship relations among employees in organizations with response rates between 85% and 87%. Scherer and Cho (2003) studied risk perception among individuals involved in a community environmental conflict over a hazardous waste site cleanup and they reported 49.5% response rate.[3]

Clearly, actor non-response is a prevalent problem in studying social networks. Having non-respondents in a network be around half may appear to be an extreme case. But it is not so rare in empirical sociometric research that it can be ignored. As a result, we took this notion into account when we included in the simulations seemingly extreme rates of non-response (see Section 6).

Effects of actor non-response on different network properties in binary networks such as network density, average vertex degree, outdegree, indegree, clustering coefficients, transitivity, assortativity, mean inverse geodesic distance and blockmodel structures have been examined previously (Stork and Richards, 1992; Costenbader and Valente, 2003; Borgatti et al., 2006; Kossinets, 2006; Huisman, 2009; Wang et al., 2012; Žnidaršič et al., 2012; Niu et al., 2015). Some of these studies delete the non-respondents and compare the results of analyses but other studies impose different actor non-response treatments (Stork and Richards, 1992; Huisman, 2009; Žnidaršič et al., 2012). It was necessary to establish a set of

[2] Four networks were excluded from their analysis as more than 50% of the actors were non-respondents. This exclusion *might not* have been necessary given some of the results reported below.

[3] The results they reported were based on omitting respondents with any missing data.



(a) A demonstration network with 15 actors and three non-respondents ($A_2$, $A_{10}$ and $A_{14}$)

(b) A reorganized demonstration network according to respondents and non-respondents

**Fig. 1.** Demonstration network with 15 actors.

treatments to be applicable also for valued networks. This is done in Section 4.

## 4. Actor non-response treatments

Stork and Richards (1992) claim the problem of non-respondents can be treated in three different ways: (i) using a complete-case analysis; (ii) using an available-case analysis; or (iii) by imputing data values as replacements of the missing data. Since the complete-case analysis (or 'listwise' deletion) removes non-respondents from the network, the result is a smaller network. The specification of network boundaries is compromised, sometimes severely. This, alone, provides a rationale for not considering this approach. A second reason for omitting the complete-case approach from our simulations is that, when blockmodeling, we are interested in the micro positions of *all* network actors including non-respondents.

We extend and modify existing actor non-response treatments (Stork and Richards, 1992; Huisman, 2009; Žnidaršič et al., 2012) to be suitable for valued networks. The result was seven different actor non-response data treatments for valued networks: (i) reconstruction; (ii) imputations based on modal values of incoming ties; (iii) a combination of reconstruction and imputation based on modal values; (iv) imputations of the mean values of incoming ties; (v) imputations of the total mean; (vi) null tie imputations and (vii) the median of the 3-nearest neighbours based on incoming ties. Without surprise, different imputation methods will produce different imputed values. This raises the issue of assessing which imputation method(s) is (are) better.

### 4.1. Reconstruction

In the reconstruction procedure, missing outgoing ties of non-respondents are replaced by the observed incoming ties to them (Stork and Richards, 1992; Huisman, 2009). As a result, the ties between non-respondents and respondents are made symmetric. The reconstruction procedure can be viewed in two different ways according to the direction of ties: (i) for undirected networks it is an 'available case approach': the relationship between two individuals is measured by using the one report of the tie (see Stork and Richards, 1992) or (ii) in case of directed networks, it is an imputation because the missing tie is estimated from the incoming tie (as used by Huisman, 2009).

However, for two non-respondents the reconstruction of ties between them is not possible. Some additional imputations are required. In the demonstration network of Fig. 1(b), six ties (located in the white square in the right lower part of the figure) between pairs of non-respondents required additional treatment. In the simplest case, the ties between pairs of non-respondents are imputed as zeroes (a treatment labeled 'reconstruction' in the following sections).

Fig. 2 shows the data for the designated non-respondents and the results for the seven treatment methods applied to the demonstration network of Fig. 1. It has eight panels for each of the three non-respondents, the first being the actual data for the designated non-respondents. The remaining panels are the imputations based on the seven treatments considered here. As described above, for reconstruction, the ties between non-respondents are replaced with zeroes (the white elements for three non-respondents $A_2$, $A_{10}$, and $A_{14}$ which are marked at the top of figure) in the second panel for each non-respondent.

### 4.2. Imputations of modal values of incoming ties

In this treatment for each missing outgoing tie $v_{ij} (i \neq j)$ of the non-respondent $i$, the modal value of values on all available incoming ties of all actors $j$ is imputed. Imputations of the modal values for three non-respondents in the demonstration network produced the following ties: for ties from non-respondents to the first two actors, 5s were imputed; for missing ties to $A_4$, 1s were imputed, and 0s for all other missing ties (see Fig. 2, third panel for each non-respondent).

### 4.3. Reconstruction and imputations based on modal values of incoming ties

This treatment combines reconstruction and imputation based on the modal values of the incoming ties (see Section 4.2). Consider the tie $x_{10,2}$ between non-respondents $A_{10}$ and $A_2$. Under the simple reconstruction procedure this tie was imputed as 0. When the combination of reconstruction and imputations of modal values of incoming ties was used, the imputed value is 5. This and other imputed values for the three non-respondents are shown in the fourth panel (for each non-respondent) of Fig. 2.

### 4.4. Null tie imputations

A frequently used actor non-response treatment is null tie imputation where zeroes are imputed for all missing ties. For the binary networks it is the worst treatment (Žnidaršič et al., 2012) for revealing both micro (position membership) and macro level (blockmodel structure) of the network, we include it here in our

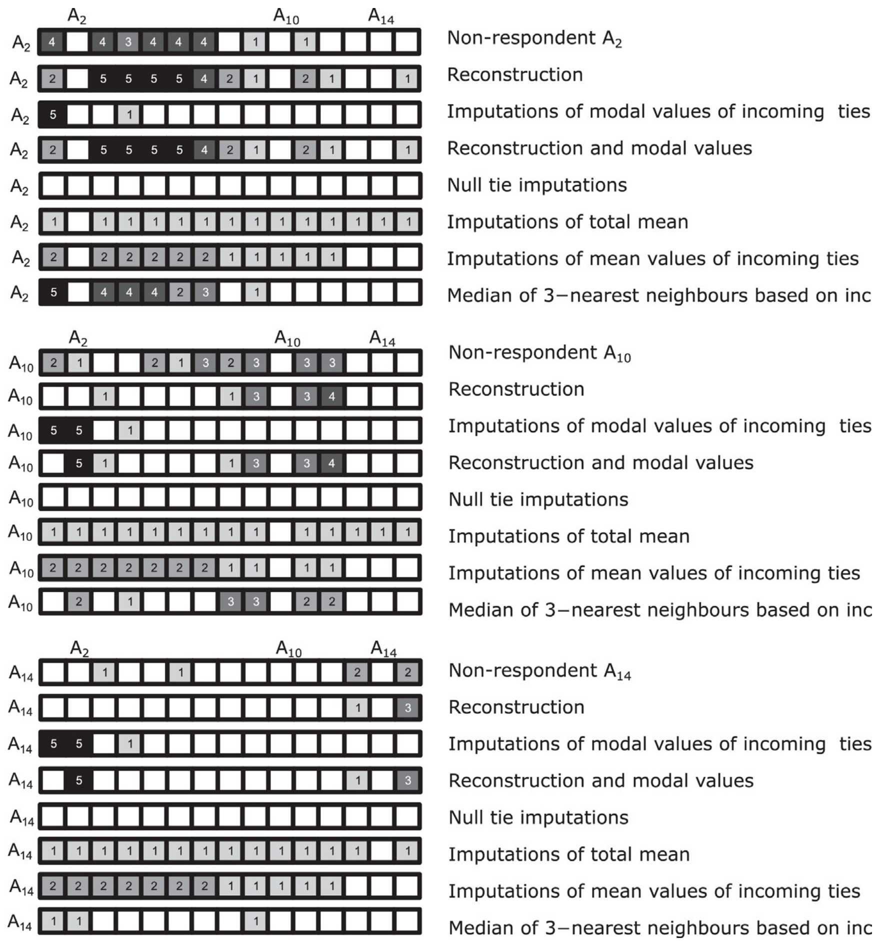



**Fig. 2.** Seven actor non-response treatments used for the demonstration network.

comparisons for valued networks. This imputation for the demonstration network is shown for each non-respondent in the fifth panel of Fig. 2.

### 4.5. *Imputations of the total mean*

For valued networks, the total mean is calculated by using all of the available tie values.[4] For the demonstration network (Fig. 1(b)), the sum of the available ties is 220. This was divided by $12 \times 14$, since each of the 12 respondents provide 14 tie values. The resulting total mean is 1.3 and we imputed the rounded value of 1 for all outgoing ties of each non-respondent. These imputations are shown for each non-respondent in the sixth panel of Fig. 2.

### 4.6. *Imputations of mean values of incoming ties*

Huisman (2009) emphasized for social networks, unconditional means can be computed in three ways: (i) the average number of ties in the network (see Section 4.5), (ii) the average number of incoming ties or 'item mean', and (iii) the average number of outgoing ties or 'person mean'. For actor non-respondents the third option cannot be used. The first two options – with some modifications – see Section 4.2 – have potential value.

The average number of incoming ties in valued networks can be the mean of the values on incoming ties of actors. More precisely, for each missing outgoing tie $v_{ij}(i \neq j)$ of the non-respondent $i$, the (rounded) mean value of values on all available incoming ties of actor $j$ is imputed. In the demonstration network, for outgoing ties from non-respondents to actors $A_1$ to $A_7$ the value 2 was imputed, for ties to actors $A_8$ to $A_{12}$ the value 1 was imputed, and 0 is the imputed value for incoming ties to actors $A_{13}$ to $A_{15}$ (twice). The values are shown for each non-respondent in the seventh panel of Fig. 2.

### 4.7. *Imputations of median of 3-nearest neighbours based on incoming ties*

For this treatment, the Euclidean distances between actors is computed based on the values of their incoming ties. Then, for each non-respondent, $i$, the three nearest neighbours are selected (e.g. $p$, $r$, $q$) according to the smallest calculated (Euclidean) distances. Finally, the median value of ties values $v_{pj}$, $v_{rj}$, and $v_{qj}$ is imputed as the outgoing tie $v_{ij}$. The values for the demonstration network are shown for each non-respondent in the final panel of Fig. 2. The rationale for this imputation is that the non-respondents are treated individually and not as a group. The comparison of the tie values for the three non-respondents and imputed values with this treatment reveal a quite good correspondence.

## 5. Indirect blockmodeling

The results of blockmodeling procedures are partitions of the actors into clusters (called positions), and, simultaneously, partitions of the ties into blocks which are determined by the clusters of actors in positions (Doreian et al., 2005).

[4] For binary networks this method uses the average number of ties in the network, that is, the network density (Huisman, 2009): for sparse networks 0s are imputed for missing ties and 1s are imputed for dense networks.

Batagelj et al. (1992) distinguish indirect and direct blockmodeling approaches. The direct approach considers only the network data by searching for a best-fitting partition given a selected equivalence type (defined by a set of permitted block types). A compatible criterion function is used to assess the agreement between the ideal blocks and the empirically delineated blocks.[5]

The direct approach is computationally burdensome, especially when networks are large. As a result, we focused on the indirect approach involving two steps. The first is computing some measure of (dis)similarity between each pair of units based on a selected equivalence. The second step features the use of classical clustering methods (e.g. hierarchical clustering, a relocation algorithm, or the leader algorithm) to identify clusters of units (Doreian et al., 2005).

Here, we consider only structural equivalence[6] because of its frequent use. Also, as pointed out by Doreian et al. (2005, p. 178) "although the definition of structural equivalence is local (in the sense of being connected to some other units), it has global implications because both location and position in networks are defined in terms of *all* other units in a network."

Dissimilarities compatible with structural equivalence include corrected Euclidean distance, corrected Manhattan distance, and the corrected dissimilarity (see Batagelj et al., 1992). Here, we used corrected Euclidean distance for a dissimilarity measure and Ward's agglomerative clustering algorithm applied to these dissimilarities.

### 5.1. Comparison of two partitions of actors

The basic ideas underlying our simulation study are: (i) start with known (designed) networks having a known partition structure; (ii) impose various amounts of non-response on them; (iii) treat the non-response with the seven treatments described in Section 4; (iv) partition the treated networks via indirect blockmodeling and (v) compare the resulting partition with known partition structure. The Adjusted Rand Index (*ARI*) was used for assessing the extent to which paritions matched (or not). Its definition is based on the Rand Index (Hubert and Arabie, 1985) measuring the concordance between two partitions and corrected for chance (see Yeung et al., 2001; Steinley, 2004). The expected value of *ARI* is 0 and its maximal value is 1. Steinley (2004), based on extensive simulations, presented some general guidelines for interpreting the *ARI* values: (i) $ARI \geq 0.9$ indicates excellent agreement; (ii) $0.9 > ARI \geq 0.8$ suggests good agreement; (iii) $0.8 > ARI \geq 0.65$ indicates moderate agreement; and iv) $ARI \leq 0.65$ indicates poor agreement. Our primary focus for whether two partitions match was restricted to the first two of these criteria with a preference for excellent agreement.

## 6. The design of the simulation study

Simulation provides a sound approach for investigating the stability of partitions of actors from indirect blockmodeling having different numbers of non-respondents and to determine the best treatment(s) of missing data. As described in Section 3 with actor non-response *all* outgoing ties of at least one actor are missing. We use the following terms: (i) a *whole network* is a known network (here, a starting network constructed to have a known partition); (ii) a *measured network* is obtained from the whole network by removing all outgoing ties for some actors; and (iii) a *treated* network obtained by employing a treatment of actor non-response to impute the missing data.

The simulations were run using R in combination with the Žiberna (2010) blockmodeling package. The code for implementing one of the seven non-response treatments in R is available under supplementary materials. The detailed code for networks simulations due to its complexity is available on request from the authors.

Section 6.1 describes the overall design of the simulations with Section 6.2 describing three distinct *types* of whole networks in more detail.

### 6.1. A basic scheme for simulations

The basic scheme of our simulation study is straightforward:

1. Generate a whole valued network using a set of three design criteria:
   (a) there are three starting (designed) structures (see Section 6.2.1);
   (b) the number of clusters (positions); and
   (c) the level of weighted reciprocity.
2. For each whole network establish a partition under indirect blockmodeling with the Corrected Euclidean distance and using the Ward's clustering method.
3. For each simulated whole network do the following:
   (a) Construct the data with non-respondents (the measured networks) by selecting some number of actors to become non-respondents and delete their outgoing ties.
   (b) Treat the measured network by imputing values for the missing data with each selected non-response data treatment (as discussed in Section 4).
   (c) Establish a partition of each treated network, again based on indirect blockmodeling with the Corrected Euclidean distance and using Ward's clustering method. We used the known number of clusters (positions) for selecting the cut points in the resulting dendrogram.
   (d) Compare the partition results of the whole and treated networks using the Adjusted Rand Index (*ARI*).
4. Investigate the impact of actor non-response and various non-response data treatments in terms of the mean values of *ARI* – denoted as *mARI* for each treatment regime. By having the same partitioning method for the whole and treated networks, the comparisons focus solely on the treatment methods.

### 6.2. Simulated whole valued networks

The whole (starting) valued networks were generated with three parameters: (i) a known starting model, (ii) the number of clusters, (iii) and varying levels for the distribution of magnitude of ties inside blocks measured with weighted reciprocity. Details are presented in the following subsections.

#### 6.2.1. Starting models

The starting whole networks were constructed based on a specified blockmodel structures: a cohesive subgroups (CS), a core-periphery (CP) model and a hierarchical (H) model, each with different numbers of positions.

All simulated networks had 75 actors with the distribution of actors inside clusters as follows (to allow for clusters of different sizes):

- Networks with 3 clusters: 35 actors in a first cluster, 25 in a second cluster, and 15 actors in a third cluster,

[5] For direct approaches to blockmodeling of binary networks see Doreian et al. (2005), while for the direct approach to valued networks see Žiberna (2007).

[6] Actors are structurally equivalent if they are connected to the rest of the network in identical ways (Lorrain and White (1971), for mathematical notation see, e.g. Batagelj et al. (1992), Doreian et al. (2005)).

- Networks with 4 clusters: 30 actors in a first cluster, 20 in a second cluster, 15 actors in a third cluster, and 10 actors in a fourth cluster, and
- Networks with 5 clusters: 25 actors in a first cluster, 20 in a second cluster, 15 actors in a third cluster, 10 actors in the a fourth cluster, and 5 actors in a fifth cluster.

An example with three cohesive subgroups is presented in Fig. 3(a). Cohesive groups have more and larger values for their internal ties compared to ties going elsewhere. Put differently, diagonal blocks have greater densities than off-diagonal blocks. We included in the design a feature where the upper triangular part of the diagonal blocks have slightly lower tie values compared to the lower triangular part (see Section 6.2.2 for further explanations). Since real word networks rarely have an ideal blockmodel structure consistent with some equivalence, ties with lower values were generated also in off-diagonal blocks.

A (single core) core-periphery model consist of one core with many internal ties but connected also with some actors in all other positions. Other positions are termed peripheral, relative to the core. They tend to be connected to the core, but with few internal connections (Doreian et al., 2005). Borgatti and Everett (1999) emphasized that a valued network has a core-periphery structure if the difference in means across blocks is relatively large compared to the variation within blocks.

Core-periphery structures have been identified through network studies in different fields. These include: a study of an urban social movement consisting of 62 civil-society organization engaged in protection of Stockholm National Urban Park (Ernstson et al., 2008); an analysis of a knowledge network addressing the problem of providing technical assistance for a set of wine producers (Giuliani, 2005); patterns of export between European countries (Rašković et al., 2015); and investigating how well the various images of the world political system drawn from the literature characterize the actual world-political network (Beckfield, 2008).

An example of a designed core-periphery model for valued network is presented in Fig. 3(b). It has one core and three peripheral positions that are connected to the core in different ways. The first peripheral cluster is more connected with core, the second much less so with the third peripheral cluster having the weakest ties to the core. The first peripheral cluster was constructed in a way that its vertices are also intraconnected with line values much lower than inside core position.[7] The third and fourth peripheral clusters have no internal ties with no ties between these positions.

Hierarchies are pervasive in human societies, taking many forms. The simplest has ties on a single path from the lowest to the highest unit. This can be extended to positions forming such a hierarchy. A network with considerable transitivity is less simple as it permits ties from lower positions to multiple higher positions (Doreian et al., 2005). Fig. 3(c) shows a designed hierarchical network with five positions having these properties. There is one main path from the bottom position to the top with varying densities. The between positions in the top right part of Fig. 3(c) are other ties flowing upwards. Consistent with examples where there are downward flowing ties (see, for example, e-mail communication inside a corporation (Kolli and Narayanaswamy, 2013) or terrorist networks (Memon et al., 2008)) some downward ties were included.

Further details on the construction of the whole models forming the foundations for our simulations are shown in Fig. 4 (and explained in the following subsection).

[7] For real world examples, the identified blockmodels can be more sophisticated with multiple cores and so-called bridging cores (Kronegger et al., 2011; Chinchilla-Rodríguez et al., 2012). Such structures were not examined here.

*6.2.2. Simulation of tie values*

We required a way of generating tie values in a consistent fashion for each network used in the simulations. These values were generated randomly using a normal distribution for the values of ties. The details vary by the blocks. More precisely:

(i) Values in the lower triangular part (for the *CS* and *CP* models) of the network were generated according to random generation of ties following the normal distribution with selected means and standard deviations. The specific parameters of normal distributions used in simulations are presented in Fig. 4. For each block the mean is provided with the standard deviation in parenthesis.
(ii) Values on ties were simulated in the upper part of the matrix with slightly moderated mean and standard deviation values.[8]
(iii) For the *CS* and *CP* models the networks with symmetrical lower and upper triangular part were generated by requiring the weighted reciprocity (*recW*) equal to 1 (see the definition below).
(iv) Since tie values were generated randomly as specified in Fig. 4, we required these values to be integers bounded by 5 and 0. Values below this range were discarded, non-integer values inside it were rounded, and generated values higher than five were set to 5.

Using the above broad strategy, different levels of symmetry of network structure were obtained. This was measured by the weighted reciprocity (*recW*) (Squartini et al., 2013):

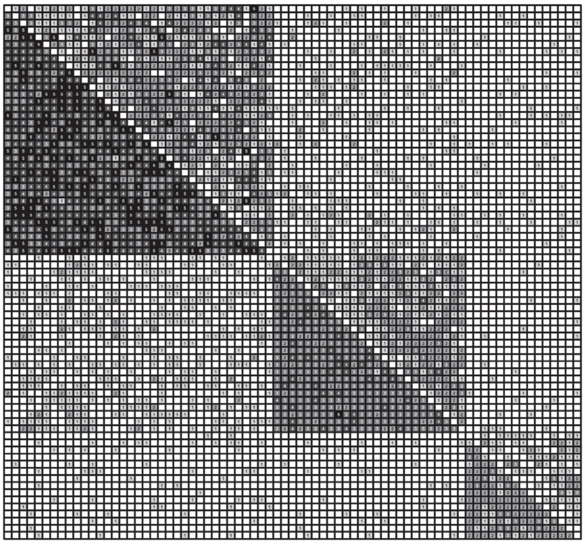
$$recW = \frac{\sum_i \sum_{j \neq i} \min(v_{ij}, v_{ji})}{\sum_i \sum_{j \neq i} v_{ij}} \tag{1}$$

The reason for including various levels of reciprocity stems from the work of Squartini et al. (2013). They argued that in directed networks, reciprocal links have dramatic effects on dynamical processes, network growth, and other structures such as triadic motifs and community structures. In networks aggregating temporal information such as e-mail or phone-call networks, reciprocity provides a measure of the simplest feed-back process occurring in the network, i.e. the tendency of a vertex to respond to another vertex stimulus. Levels of reciprocity are important features of social networks.

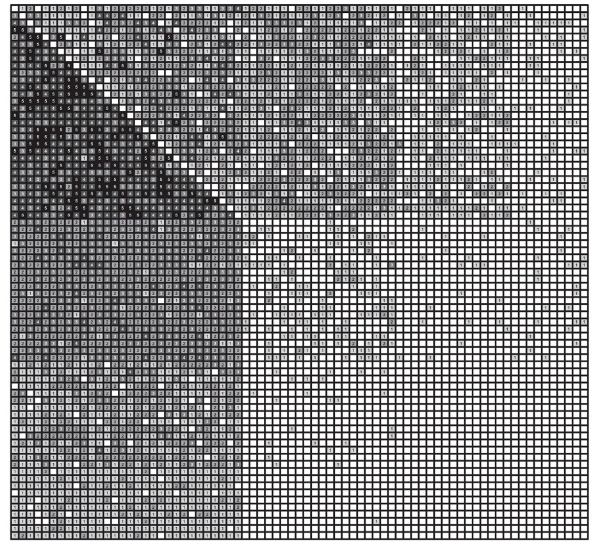
### 6.3. Simulation of non-respondents and treatments used

Some studies, for example Potter et al. (2015), show actor attributes affect who do not respond while others, for example (Johnson et al., 2012), suggest actor attributes do not affect on which actors provide no data regarding their network ties. Our primary concern here is the impact of different levels of non-response on blockmodeling results. To have a less complex study design, our simulations did not include actor attributes. Instead, actors were selected at random for our simulated networks as the non-respondents. The number of simulated non-respondents for the simulated whole networks ranged from 1 to 40 (with actor response rates ranging from 47% to 99%). The generation of incomplete data was repeated 75 times for networks with one missing actor, 100 times for all combinations of two or more non-respondents. The measured networks were then treated with the seven actor non-response treatments described in Section 4 before conducting blockmodeling of the resulting data.

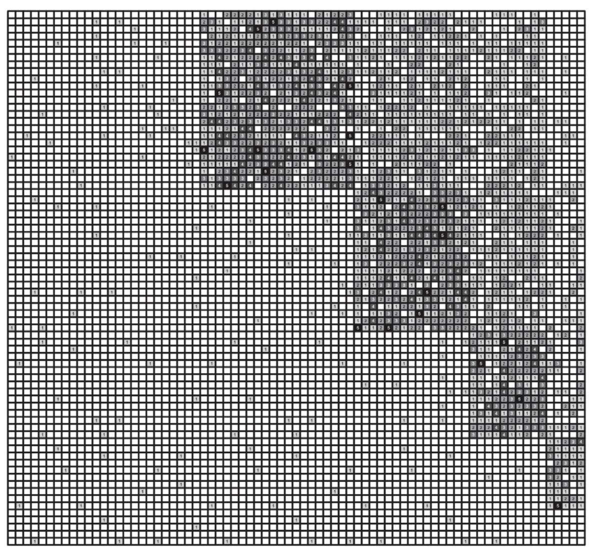
[8] The mean values from the normal distribution used in lower triangular part were multiplied by symmetry factor (*sR*) lower than 1, while the standard deviation was increased by multiplying it by $(2 - sR)$.



(a) A cohesive subgroups model with 3 clusters



(b) A core-periphery model model with 4 clusters



(c) A hierarchical model model with 5 clusters

**Fig. 3.** Examples of three starting blockmodel structures for networks with 75 actors.

### 6.4. The total number of (simulated) networks

Using the basic simulation scheme described in Section 6.1 with varying parameters, the following number of starting whole networks, generated measured networks, and treated networks were constructed:

- Number of whole starting networks:
  - Eleven networks with different distributions of tie values according to levels of weighted reciprocity for the cohesive subgroups model. There were eleven more for the core-periphery model and ten for the hierarchical model (where complete symmetry is not possible). The number of such generated networks is 32 $(2 \cdot 11 + 10)$.
  - For blockmodels with three to five clusters and five repetitions for each combination of level of reciprocity and number of clusters there were 15 $(3 \cdot 5)$ constructions.
  - The total number of whole starting networks for the simulations was 480 $(32 \cdot 15)$.
- The number of measured networks for each whole network, with different numbers of non-respondent is 2275.
- The total number of treated networks is 7,644,000 $(480 \cdot 2275 \cdot 7)$ for 480 whole networks, 2275 simulated non-response regimes and seven non-response treatments. The results reported below are based on a very high number of treated networks.

## 7. Results

Our results are presented for three outcomes. The first results, for the *mARI* values, are presented in Section 7.1. These are about the adequacy (or not) for different treatments where the criterion is the known blockmodel structure. The effects of design parameters on the stability of the blockmodel partitions of actors are examined in Section 7.2 using the analysis of covariance

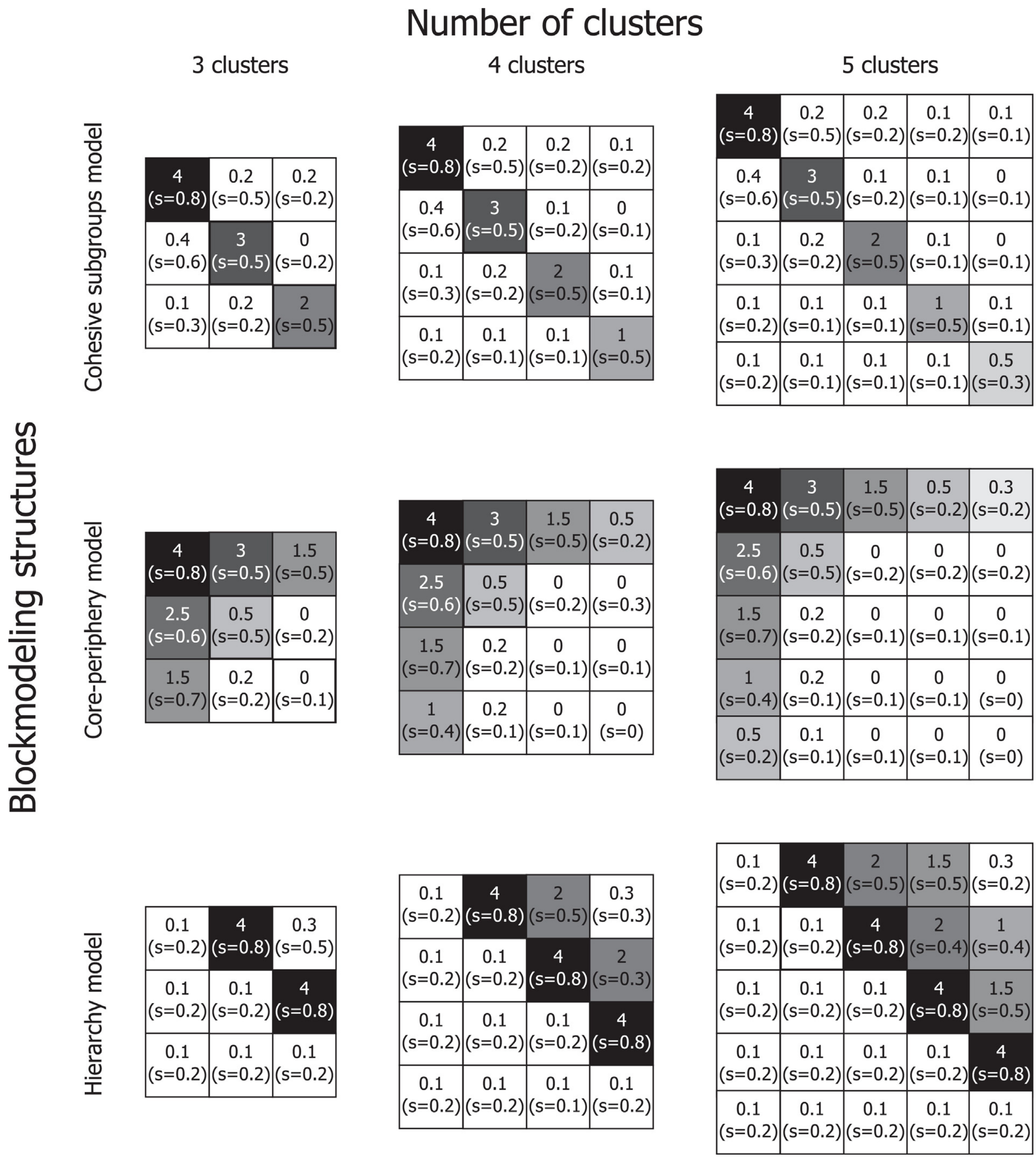



**Fig. 4.** The scheme for simulating blockmodels according to the number of clusters with parameters of normal distribution used in simulations of values in whole networks.

procedure. Finally, the best actor non-response treatment(s) for each of the three blockmodel structures are shown in Section 7.3.

### 7.1. The overall results

The overall results for the agreement between the original designed partition and treated partitions, as measured by the Adjusted Rand Index, for more than seven and half million of treated networks are presented in Fig. 5. The *x*-axis is the number of non-respondents, while on the *y*-axis are the mean values of *ARI* (*mARI*). The trajectories for the seven treatments are shown. Without surprise, the values of *mARI* decline as the amount of non-response increases. The minimal criterion for accepting that two partitions correspond is 0.8. Examining the separate trajectories for the seven treatments is instructive as there are clear differences between them.

The trajectory for the treatment using the median of the 3-nearest neighbours based on incoming ties shows this treatment is far superior for dealing with non-response. Throughout the entire range of non-response, the trajectory remains well above 0.8. Indeed, it is above 0.9 of the first third of the non-response range. Regardless of the type of blockmodel, symmetry of a network and/or number of clusters this treatment works extremely well. Using it allows the recovery of a blockmodel even when we have less than half respondents in a network. Based on Fig. 5, our recommendation is to always use this treatment. Of course, using it requires more effort.

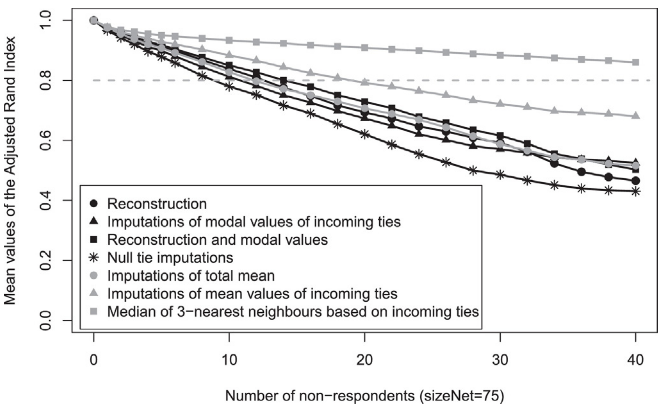



**Fig. 5.** Overall results of the simulation study for the Mean of Adjusted Rand Index.

The second best treatment for non-response is using imputations of mean values of the incoming ties to non-respondents. Its trajectory is above 0.8 until the number of non-respondents reaches 20. The worst treatment, according to its *mARI* trajectory is the null tie imputation. Even so, when there were eight or fewer non-respondents, the *mARI* values were above 0.8 for all seven treatments. The differences between reconstruction, imputation of modal values of incoming ties, reconstruction in combination with imputation of modal values, and imputation of the total mean are modest in this low range of non-respondents. The values of *mARI* for those four treatments are above 0.8 also for up to 10 non-respondents. Thereafter, while they perform in a similar fashion across the whole range of non-respondents, the trajectories drop below 0.8.

### 7.2. The ANCOVA model

Even though the superiority of the median of the 3-nearest neighbours based on incoming ties was determined according to the overall results of *mARI* values, the whole set of results is more complex. To explore this further, ANCOVA was used to investigate the effects of: (i) the number of non-respondents (labeled as *nNR*); (ii) treatments of non-response data (*T*); (iii) the number of clusters (*nClu*); (iv) the type blockmodel structure of the network (*S*), and (v) weighted reciprocity (*recW*).

Table 1 contains the ACNOVA results for the mean values of the Adjusted Rand Index. The main effects and all interactions (two, three-way, four-way, and five-way) are ordered according to their partial $\eta^2$ values. Without surprise, the number of non-respondents in a network has the highest effect on the mean Adjusted Rand Index (partial $\eta^2 = 0.0901$).

From Fig. 5 it was clear that the larger the number of non-respondents, the lower was the mean Adjusted Rand Index (*mARI*). The interaction of the type of non-response treatment and number of non-respondents had the second highest effect (partial $\eta^2 = 0.0171$) where the null tie imputations performed the worst and the median of 3-nearest neighbours based on incoming ties is the best. The third largest effect on *mARI* is the interaction between treatment, blockmodel structure of a network and number of non-respondents (partial $\eta^2 = 0.0119$). The additional attention this merits is provided in Section 7.3 where the results of *mARI* for the three blockmodels structures are examined separately.

The weighted reciprocity has its highest effect in its interaction with the blockmodel structure, the treatment and the number of non-respondents (partial $\eta^2 = 0.0091$) and in its interaction with the blockmodel structure, the number of clusters and the number of non-respondents (partial $\eta^2 = 0.0079$).

**Table 1**
Analysis of covariance for the mean values of the Adjusted Rand Index.

| Effect | $Df_1$ | $F$ | Partial $\eta^2$ |
|---|---|---|---|
| nNR | 1 | 757,232.6 | 0.0901 |
| T * nNR | 6 | 22,185.6 | 0.0171 |
| S * T * nNR | 12 | 7644.7 | 0.0119 |
| S * nClu * nNR | 4 | 19,116.8 | 0.0099 |
| S * T * recW * nNR | 12 | 5837.4 | 0.0091 |
| S * nClu * recW * nNR | 4 | 15,208.5 | 0.0079 |
| S * nClu * T * nNR | 24 | 2349.2 | 0.0073 |
| T * recW * nNR | 6 | 7801.7 | 0.0061 |
| S * nNR | 2 | 23,299.5 | 0.0061 |
| S * nClu * T * recW * nNR | 24 | 1745.5 | 0.0055 |
| nClu * T * nNR | 12 | 3465.5 | 0.0054 |
| recW | 1 | 39,783.1 | 0.0052 |
| S * recW | 2 | 19,832.3 | 0.0052 |
| S * recW * nNR | 2 | 18,562.6 | 0.0048 |
| recW * nNR | 1 | 35,000.4 | 0.0046 |
| S * nClu | 4 | 6794.9 | 0.0035 |
| nClu * nNR | 2 | 11,724.8 | 0.0031 |
| S * nClu * recW | 4 | 5643.7 | 0.0029 |
| nClu * T * recW * nNR | 12 | 1819.3 | 0.0028 |
| S | 2 | 10,430.8 | 0.0027 |
| S * T * recW | 12 | 1449.4 | 0.0023 |
| nClu * recW | 2 | 8156.4 | 0.0021 |
| S * nClu * T * recW | 24 | 532.6 | 0.0017 |
| S * T | 12 | 990.1 | 0.0016 |
| T | 6 | 1831.8 | 0.0014 |
| S * nClu * T | 24 | 453.2 | 0.0014 |
| nClu * recW * nNR | 2 | 4665.6 | 0.0012 |
| nClu * T | 12 | 582.7 | 0.0009 |
| nClu * T * recW | 12 | 577.4 | 0.0009 |
| T * recW | 6 | 1133.5 | 0.0009 |
| nClu | 2 | 54.0 | 0.0000 |
| Residual degrees of freedom, $Df_2 = 7, 643, 748$ | | | |
| $R^2 = 0.733$ | | | |

nNR – number of non-respondents; T – treatment; nClu – number of clusters in network; S – structure of a network; recW – weighted reciprocity.

As a main effect, the number of clusters alone has the lowest effect (partial $\eta^2 < 0.0000$). It has a significant effect only in combination with blockmodel structure and number of non-respondents (partial $\eta^2 = 0.0099$).

### 7.3. Separate results for the three blockmodel structures

We turn now to consider the separate results for cohesive subgroups, the core-periphery model and the hierarchical model. The nature of the 'true' structural model does make a difference.

#### 7.3.1. The cohesive subgroups models

Fig. 6 presents the results for the mean Adjusted Rand Index for the range of non-respondents for the cohesive subgroups model. The median of 3-nearest neighbours based on incoming ties is the best treatment: the values of *mARI* are even above 0.9 throughout the *entire* range of non-response levels. This is remarkable and indicates excellent agreement between original and treated partitions. To the extent that blockmodeling is used for identifying cohesive subgroups, it is hard to avoid the recommendation of this being by far the *best* treatment option. Both reconstruction treatments are practically interchangeable for 32 non-respondents or less. They are the second best option in non-response treatment selection since values of *mARI* are above 0.8 over most, but not all of, the range of non-response studied here. The imputations of the total mean and the imputations of mean values of incoming ties perform well for up to 24 non-respondents. However, for larger numbers of non-respondents they are both unacceptable. The imputations using the modal values of incoming ties are practically identical for cohesive subgroups to using null ties imputations. Although, the *mARI* values are above 0.8 for up to 12 non-respondents, they are the worst treatments compared to others. Therefore, neither of

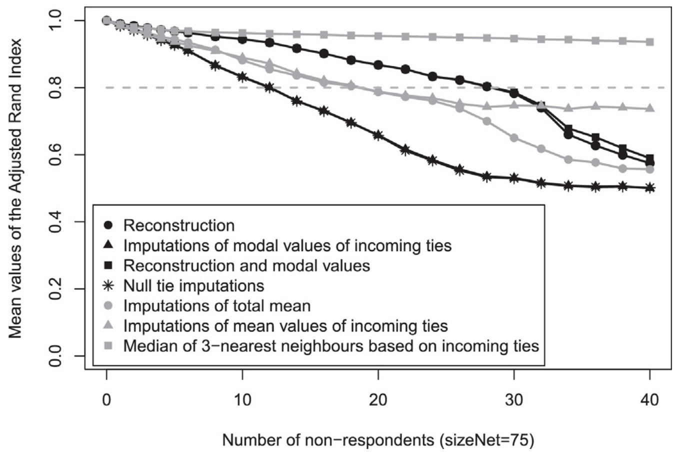



**Fig. 6.** Results of the simulation study for the Mean of Adjusted Rand Index for the cohesive subgroups model.

these two treatments can be recommended for larger numbers of non-respondents.

*7.3.2. The core-periphery model*

The corresponding results for the core-periphery model are shown in Fig. 7. Compared to the cohesive subgroups model, there is a larger diversity across the non-response treatments. The best treatment, again, is the median of the 3-nearest neighbours where the *mARI* values are above 0.8 for the whole range of simulated non-respondents. The null tie imputations are the worst: this treatment is not worthwhile. While the reconstruction in combination with modal values and the imputations of mean values of incoming ties are acceptable treatments for low levels of non-response, they have little value when non-response levels are higher.

*7.3.3. The hierarchical model*

Results for the hierarchical model are presented in Fig. 8. Yet again, the best non-response treatment is the median of 3-nearest neighbours based on incoming ties. The values of *mARI* are above 0.8 for the whole range of simulated non-respondents. For up to 12 non-respondents, the imputation using the mean values of incoming ties is a credible rival. Indeed, in this range, it is even better than the median of 3-nearest neighbours. Thereafter, its performance is worse but it remains the second best treatment. Only for low values of non-response are the other treatments are acceptable (but still inferior).

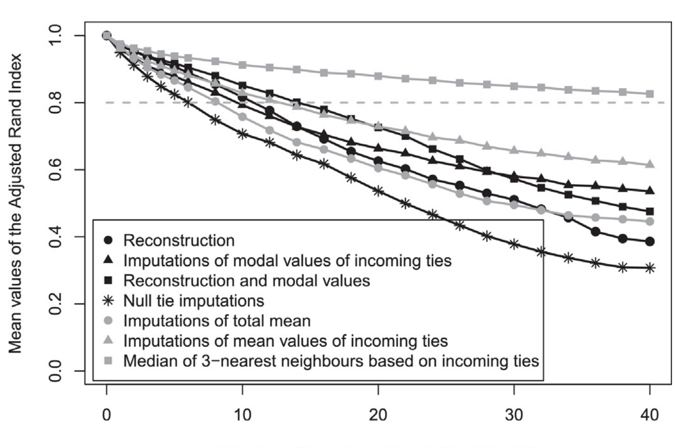



**Fig. 7.** Results of the simulation study for the Mean of Adjusted Rand Index for core periphery model.

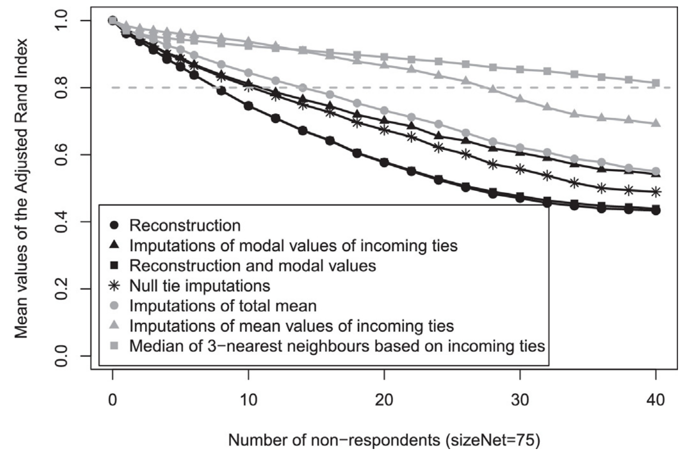



**Fig. 8.** Results of the simulation study for the Mean of Adjusted Rand Index for the hierarchy model.

Both reconstruction procedures perform the worst since the blockmodel structure is highly non-symmetrical. The *mARI* values are above 0.8 only for 6 non-respondents or less. They cannot be recommended for studying this type of blockmodel structure.

## 8. Conclusions and recommendations

Non-response when studying the overall structure of social networks is a major problem. This problem is *not* solved by simply accepting or ignoring non-response. This paper presents an extensive set of simulations of actor non-response for three well known blockmodel structures: cohesive subgroups, core-periphery and hierarchical models. The impact of seven actor non-response treatments was examined on the partitions obtained when using indirect blockmodeling.

Based on this simulation study plus the ANCOVA model and graphical representations in the previous Section, the following recommendations can be given:

- At a minimum, report the percentage (or number) of actor(s) who did not provide responses together with the size of the network.
- When there are non-respondents, do *not* discard the incoming ties to them from actors who did respond. Instead, consider using some non-response treatment. But note that some treatments are much better than others.
- Regardless the (hypothesized) blockmodel structure of the network the preferable actor non-response treatment is using the median of k-nearest neighbours[9] based on incoming ties.
- If the weighted reciprocity of the network is low and if the percentage of non-respondents is below 15% the imputations of the total mean could be an appropriate treatment.
- Never replace missing ties with zeroes because null tie imputation is the *worst* treatment when attempting to reveal position membership of non-respondents in valued networks. See also (Žnidaršič et al., 2012) for similar results for binary networks.

The main limitation of the study is the limited selection of starting networks especially according to their size. The design of simulations was limited due to computational constraints. Due to the greater complexity of valued networks according to both the patterns of ties and values on those ties, further simulations have

[9] In the simulations (only) three nearest neighbours was determined and the results were superb.

to be performed where a broader set of tie magnitudes will be taken into account.

In addition, since the magnitude of values is a great advantage of valued networks compared to (simplified) binary networks, the changes of values on ties have to be examined closely in terms of measurement errors, tie non-response or other errors (Žnidaršič et al., 2012) in the research designs of social networks.

While the results reported here pertain to blockmodel structures, it seems likely that non-response issues will affect most other indices of network structure at both global and local levels. *Simply ignoring non-response by pretending it does not matter is not an option.*

Clearly, this kind of measurement error has implications for other network features including centrality measures. Some results of a simulation study on stability of several centrality measures (e.g. weighted outdegree and indegree (Opsahl et al., 2010), and weighted betweenness and closeness centrality (Opsahl et al., 2010)) together with density of a valued network (Wasserman and Faust, 1998) using the same set of valued networks as presented here can be found in (Žnidaršič et al., submitted for publication). Further simulations based on extended set of centrality measures (e.g. eigenvector centrality (Bonacich, 1987), information centrality (Stephenson and Zelen, 1989), load centrality (Goh et al., 2001), stress centrality (Shimbel, 1953), diffusion centrality (Banerjee et al., 2014), page rank (Brin and Page, 1998), hub and authority centrality scores (Kleinberg, 1999) and fragment centrality (Borgatti, 2006)) are currently under investigation.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.socnet.2016.06.001.

## References

Banerjee, A., Chandrasekhar, A.G., Duflo, E., Jackson, M.O., 2014. Gossip: Identifying Central Individuals in a Social Network. Tech. Rep.
Batagelj, V., Ferligoj, A., Doreian, P., 1992. Direct and indirect methods for structural equivalence. Social Netw. 14 (1–2), 63–90.
Beckfield, J., 2008. The dual world polity: fragmentation and integration in the network of intergovernmental organizations. Social Probl. 55 (3), 419–442.
Bonacich, P., 1987. Power and centrality: a family of measures. Am. J. Sociol. 92 (5), 1170–1182.
Borgatti, S.P., 2006. Identifying sets of key players in a social network. Comput. Math. Organ. Theory 12 (1), 21–34.
Borgatti, S.P., Carley, K.M., Krackhardt, D., 2006. On the robustness of centrality measures under conditions of imperfect data. Social Netw. 28 (2), 124–136.
Borgatti, S.P., Everett, M.G., 1999. Models of core/periphery structure. Social Netw. 21, 375–395.
Brin, S., Page, L., 1998. The anatomy of a large-scale hypertextual web search engine. In: Proceedings of the Seventh International Conference on World Wide Web 7. WWW7, pp. 107–117 http://dl.acm.org/citation.cfm?id=297805.297827.
Chinchilla-Rodríguez, Z., Ferligoj, A., Miguel, S., Kronegger, L., de Moya-Anegón, F., 2012. Blockmodeling of co-authorship networks in library and information science in Argentina: a case study. Scientometrics 93 (3), 699–717.
Costenbader, E., Valente, T.W., 2003. The stability of centrality measures when networks are sampled. Social Netw. 25 (4), 283–307.
Doreian, P., Batagelj, V., Ferligoj, A., 2005. Generalized Blockmodeling. Cambridge University Press, New York, NY, USA.
Doreian, P., Woodard, K.L., 1994. Defining and locating cores and boundaries of social networks. Social Netw. 16 (4), 267–293.
Ellwardt, L., Steglich, C., Wittek, R., 2012. The co-evolution of gossip and friendship in workplace social networks. Social Netw. 34 (4), 623–633.
Ernstson, H., Sörlin, S., Elmqvist, T., 2008. Social movements and ecosystem services – the role of social network structure in protecting and managing urban green areas in Stockholm. Ecol. Soc. 13.
Girard, Y., Hett, F., Schunk, D., 2015. How individual characteristics shape the structure of social networks. J. Econ. Behav. Organ. 115, 197–216, Behavioral Economics of Education.
Giuliani, E., 2005. The structure of cluster knowledge networks: uneven and selective, not pervasive and collective. In: Applied Evolutionary Economics and Economic Geography, pp. 5–11.
Goh, K.-I., Kahng, B., Kim, D., 2001. Universal behavior of load distribution in scale-free networks. Phys. Rev. Lett. 87, 1–4.
Grund, T.U., 2012. Network structure and team performance: the case of English premier league soccer teams. Social Netw. 34 (4), 682–690.
Hlebec, V., Ferligoj, A., 2002. Reliability of social network measurement instruments. Field Method 14, 288–306.
Holland, P.W., Leinhardt, S., 1973. The structural implications of measurement error in sociometry. J. Math. Sociol. 3 (1), 85–11.
Horvath, S., 2011. Weighted Network Analysis: Applications in Genomics and Systems Biology. Springer Science & Business Media, New York, NY, USA.
Hubert, L., Arabie, P., 1985. Comparing partitions. J. Classif. 2, 193–218.
Huisman, M., 2009. Effects of missing data in social networks. J. Social Struct. 10 (1) http://www.cmu.edu/joss/content/articles/volume10/huisman.pdf.
Johnson, R., Kovács, B., Vicsek, A., 2012. A comparison of email networks and off-line social networks: a study of a medium-sized bank. Social Netw. 34 (4), 462–469.
Kleinberg, J.M., 1999 Sep. Authoritative sources in a hyperlinked environment. J. ACM 46 (5), 604–632, http://dx.doi.org/10.1145/324133.324140.
Kolli, N., Narayanaswamy, B., 2013. Analysis of e-mail communication using a social network framework for crisis detection in an organization. Proc. – Social Behav. Sci. 100, 57–67, Applications of Social Network Analysis 2011.
Kossinets, G., 2006. Effects of missing data in social networks. Social Netw. 28 (3), 247–268.
Kronegger, L., Ferligoj, A., Doreian, P., 2011. On the dynamics of national scientific systems. Qual. Quant. 45 (5), 989–1015.
Laumann, E., Marsden, P., Prensky, D., 1983. The boundary specification problem in network analysis. In: Burt, R., Minor, M. (Eds.), Applied Network Analysis: A Methodological Introduction. Sage Publications, London, pp. 18–34.
Lorrain, F., White, H.C., 1971. Structural equivalence of individuals in social networks. J. Math. Sociol. 1, 49–80.
Memon, N., Larsen, H., Hicks, D., Harkiolakis, N., 2008. Retracted: detecting hidden hierarchy in terrorist networks: some case studies. In: Yang, C., Chen, H., Chau, M., Chang, K., Lang, S.-D., Chen, P., Hsieh, R., Zeng, D., Wang, F.-Y., Carley, K., Mao, W., Zhan, J. (Eds.), Intelligence and Security Informatics. Vol. 5075 of Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 477–489.
Niu, Q., Zeng, A., Fan, Y., Di, Z., 2015. Robustness of centrality measures against network manipulation. Physica A: Stat. Mech. Appl. 438, 124–131.
Opsahl, T., Agneessens, F.F., Skvoretz, J., 2010. Node centrality in weighted networks: generalizing degree and shortest paths. Social Netw. 32, 245–251.
Páez, A., Scott, D.M., Volz, E., 2008. Weight matrices for social influence analysis: an investigation of measurement errors and their effect on model identification and estimation quality. Social Netw. 30 (4), 309–317.
Potter, G.E., Smieszek, T., Sailer, K., 2015. Modeling workplace contact networks: the effects of organizational structure, architecture, and reporting errors on epidemic predictions. Netw. Sci. 3, 298–325.
Rašković, M., Udovič, B., Žnidaršič, A., 2015. Network analysis of inter-country export patterns in the EU: implications for small states. Teorija in Praksa LII (1-2), 150–174.
Scherer, C.W., Cho, H., 2003. A social network contagion theory of risk perception. Risk Anal. 23 (2), 261–267.
Scott, J., 2013. Social Network Analysis, 3rd ed. Sage Publications, London, UK.
Shimbel, A., 1953. Structural parameters of communication networks. Bull. Math. Biophys. 15 (4), 501–507.
Squartini, T., Picciolo, F., Ruzzenenti, F., Garlaschelli, D., 2013. Reciprocity of weighted networks. Sci. Rep. 3.
Steinley, D., 2004. Properties of the Hubert–Arabie Adjusted Rand Index. Psychol. Methods 9 (3), 386–396.
Stephenson, K., Zelen, M., 1989. Rethinking centrality: methods and applications. Social Netw. 92 (11), 1–37.
Stork, D., Richards, W.D., 1992. Nonrespondents in communication network studies: problems and possibilities. Group Organ. Manage. 17, 193–209.
Van De Bunt, G.G., Van Duijn, Marijtje, A.J., Snijders, T.A., 1999. Friendship networks through time: an actor-oriented dynamic statistical network model. Comput. Math. Organ. Theory 5 (2), 167–192.
Žiberna, A., 2007. Generalized blockmodeling of valued networks. Social Netw. 29 (1), 105–126.
Žiberna, A., 2009. Evaluation of direct and indirect blockmodeling of regular equivalence in valued networks by simulations. Metodološki zvezki 6, 99–134.
Žiberna, A., 2010. Blockmodeling 0.1.8: An R Package for Generalized and Classical Blockmodeling of Valued Networks. https://cran.r-project.org/web/packages/blockmodeling/blockmodeling.pdf.
Žnidaršič, A., Ferligoj, A., Doreian, P., 2012. Non-response in social networks: the impact of different non-response treatments on the stability of blockmodels. Social Netw. 34, 438–450.
Žnidaršič, A., Ferligoj, A., Doreian, P., 2016. Stability of centrality measures in valued networks regarding different actor non-response treatments and macro-network structures (submitted for publication).
Wang, D.J., Shi, X., McFarland, D.A., Leskovec, J., 2012. Measurement error in network data: a re-classification. Social Netw. 34 (4), 396–409.
Wasserman, S., Faust, K., 1998. Social Network Analysis: Methods and Applications, 2nd ed. Cambridge University Press, Cambridge, USA.
Yeung, K.Y., Yeung, K.Y., Ruzzo, W.L., Ruzzo, W.L., 2001. An empirical study on principal component analysis for clustering gene expression data. Bioinformatics 17, 763–774.