Social Networks 44 (2016) 160–178

Contents lists available at ScienceDirect

Social Networks

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

A deviational approach to blockmodeling of valued networks

Carl Nordlund*

Center for Network Science, Department of Political Science (shared affiliation), Central European University, Budapest, Hungary

ARTICLE INFO

Article history:
Available online 14 September 2015

Keywords:
Valued networks
Valued blockmodeling
Structural equivalence

ABSTRACT

This article proposes a novel approach to blockmodeling of valued (one-mode) networks where the identification of (binary) block patterns in the valued relations differ from existing approaches. Rather than looking at the absolute values of relations, or examining valued ties on a per-actor basis (cf. Nordlund, 2007), the approach identifies prominent (binary) ties on the basis of deviations from expected values. By comparing the distribution of each actor's valued relations to its alters with the macro-level distributions of total in- and outdegrees, prominent (1) and non-prominent (0) ties are determined both on a per-actor-to-actor and a per-actor-from-actor basis. This allows for a direct interpretation of the underlying functional anatomy of a non-dichotomized valued network using the standard set of ideal blocks as found in generalized blockmodeling of binary networks.

In addition to its applicability for direct blockmodeling, the article also suggests a novel indirect measure of deviational structural equivalence on the basis of such deviations from expected values.

Exemplified with the note-sharing data in Žiberna (2007a), citations among social work journals (Baker, 1992), and total commodity trade among EU/EFTA countries as of 2010, both the direct and indirect approach produce results that are more sensitive to variations at the dyadic level than existing approaches. This is particularly evident in the case of the EU/EFTA trade network, where the indirect approach yields partitions and blockmodels in support of theories of regional trade, despite the significantly skewed valued degree distribution of the dataset.

1. Introduction

As the quintessential tool in the study of social roles, blockmodeling allows for the analysis of role-equivalence in networks and the different roles actors have with each other (e.g. Ferligoj et al., 2011, p. 434). Building on the foundational paper of Lorrain and White (1971), the technique stems from a series of articles in the mid-1970s (Breiger et al., 1975; White, 1974a, 1974b; White et al., 1976; see Galtung, 1966 for an independent precursor). Refined further by the Ljubljana school of network scholars, generalized blockmodeling (e.g. Doreian et al., 2005) constitutes a formal and integrated approach for the study of the underlying functional anatomies of virtually any set of relational data.

An open problem with contemporary blockmodeling is that it is primarily designed for binary, rather than valued, networks (Doreian, 2006, p. 127). This is particularly apparent in the set of ideal blocks used for interpreting blockmodels, ideal role-relational templates whose characteristic tie patterns are binary and thus not readily comparable with valued empirical blocks. Apart from the dichotomization of valued data and the loss of information this entails, less crude alternatives for comparing ideal and empirical blocks have been suggested (Žiberna, 2007a; Nordlund, 2007). These alternatives are more sensitive to variations in valued data, but they nevertheless conflate the prominence of valued ties with their absolute strengths.

This article proposes a novel approach to generalized blockmodeling of valued directional (one-mode) networks where the underlying conceptualization of "prominence" of ties differs from existing approaches. By analyzing how observed tie values deviate from expected values based on the distribution of total in- and outdegrees for each actor, the suggested approach determines whether a tie is prominent (a 1-cell) or not (a 0-cell) on a dyadic basis. Allowing for direct comparisons between empirical valued blocks and the ideal binary blocks used in generalized blockmodeling, the approach, by default non-parametric, excludes dichotomization, pruning or otherwise modification of the original data, where the identification of prominent ties only partially depends on their absolute values. By adapting how block penalty scores and overall goodness-of-fit measures are calculated in generalized blockmodeling, the suggested approach allows for either finding optimal partitions and/or block images, or for evaluating

* Correspondence to: Central European University, Center for Network Science, Nador utca 9, 1051 Budapest, Hungary. Tel.: +36 70 592 8442.
E-mail addresses: nordlundc@ceu.edu, carl.nordlund71@gmail.com

http://dx.doi.org/10.1016/j.socnet.2015.08.004
0378-8733/

hypothetical blockmodels, i.e. similar to how this is done in generalized blockmodeling.

Whereas the approach is non-parametric by default, this being the recommended setting in the general case, an optional parameter determines the prominence of valued ties as a minimum percentual deviation between observed and expected flows. A "two-sided" application of this parameter, equally optional, introduces the possibility of non-determined ties, i.e. valued relations that are deemed neither prominent (1) nor non-prominent (0). Occurrences of such ties motivate the modification of how inconsistencies between empirical and ideal blocks are calculated. Such ties also introduce measures of interpretational certainty, specific to each ideal block. The maximum two-sided deviation threshold that holds the aggregate uncertainty score at zero (or near-zero levels) is suggested as a measure of interpretational certainty for valued blockmodels, in effect transforming the optional parameter into an outgoing stat.

In addition to using the suggested approach in direct blockmodeling, this article also proposes an alternative indirect route for determining structurally equivalent positions in valued networks. Using the deviations between observed and expected values as input, a modified formula for calculating Pearson correlation coefficients is proposed as a measure of deviational structural equivalence. Having partitioned the network based on cluster analysis of the resulting equivalence coefficients, the subsequent identification of prominent ties makes the interpretation of such deviational structural blockmodels more straight-forward than what is the case for other existing indirect approaches.

Implementing a local-optimization search algorithm for finding optimal partitions and/or blockimages, the provided software client[1] also implements the modified Pearson correlation formula. The direct blockmodeling approach is tested on the note-sharing data analyzed by Žiberna (2007a) and citations among social work journals (Baker, 1992). Using a pre-specified partition into geographical regions, a dataset on total commodity trade among European (EU plus EFTA) countries as of 2010 is analyzed for would-be spatial trade-gravity effects. The proposed formula for deviational structural equivalence is exemplified on both the journal citation data and the European trade datasets. Taken together, these examples demonstrate how the proposed approach seemingly produces viable and intuitive results for valued datasets, irrespectively of how skewed their valued in- and outdegree distributions are.

The remainder of this article is divided into four parts. The first section provides an introduction to blockmodeling, with a focus on valued networks. Discussing existing approaches to partition and interpreting valued blockmodels, this section highlights how these approaches tend to emphasize tie strengths when empirical valued blocks are compared with ideal binary blocks. Addressing this concern, the second part introduces and specifies the suggested deviational approach to generalized blockmodeling of valued network. This part also specifies the indirect measure of deviational structural equivalence. The third part applies direct and indirect deviational approaches to blockmodeling on the three example datasets, comparing these results with those obtained when using existing (conventional) direct and indirect approaches. Concluding this article, the final part discusses the findings, outlines future research and the possibilities to transpose the proposed deviational approach to other binary-oriented network-analytical metrics.

[1] A demonstrational Windows software client (KrishKrosh) with example datasets available at http://cnslabs.ceu.hu.

2. Blockmodeling of binary and valued networks[2]

A blockmodel is created by partitioning actors into subsets ('positions' or 'clusters' in blockmodel terminology) based on a meaningful definition of equivalence, subsequently sorting the original sociomatrix by these positions. It is interpreted by comparing intra- and inter-positional 'blocks' with a given set of ideal block types, the latter corresponding to different types of ideal role-relational patterns. A blockmodel can be reduced to a block image and a corresponding image graph that ideally capture the underlying role-structure – the functional anatomy – of the network.

The partitioning of actors is either done indirectly, typically using a suitable algorithm or formula intended to capture and measure the degree of equivalence (of some type) among actors, or directly by permutating actors among positions in the blockmodel in search of partitions that minimize the number of inconsistencies between observed and ideal blocks. A search algorithm could be used to determine an optimal partition, possibly with a predetermined hypothetical block image or only with a preset number of positions. A pre-specified partition could also be examined by comparing the resulting blocks with a set of ideal blocks. In indirect approaches, dyadic measures of equivalence are typically used in conjunction with hierarchical clustering and/or multidimensional scaling to identify suitable positions of equivalent actors, positions that specify a blockmodel for subsequent interpretation. In all these approaches, the total inconsistencies between observed and ideal blocks are calculated for each hypothetical blockmodel as measures of goodness-of-fit.

The ideal block types are templates that represent different types of role-relations. In structural blockmodeling, where actor equivalence implies having identical ties to alters, the ideal block types are either complete (oneblocks) or null (zeroblocks). In regular blockmodeling, where equivalence between two actors implies having similar ties to actors that in turn are regularly equivalent, the ideal regular block is used, consisting of at least one tie in each row and column, respectively, of the block. In generalized blockmodeling, these traditional ideal blocks are supplemented with additional ideal blocks – see Table 1.

The binary nature of the ideal block types makes them less than ideal for comparison with valued data. To allow for such comparisons, valued blockmodels are often dichotomized, where a statistically, theoretically or arbitrarily determined threshold value determines the prominence (1) and non-prominence (0) of each valued relation.

In addition to the inevitable loss of information and the dilemma of choosing a suitable cutoff value, the feasibility of a dichotomization hinges on an implicit assumption of equal relational capacities among actors. Although a viable assumption in many situations, e.g. when mapping playtime among school children during a 45-min lunch break, the relationship between absolute values and local perceptions of importance could differ substantially among actors in a network. Colombia indeed deems its flow of exports to USA as highly significant, a valued directional tie that pales in comparison with other bilateral imports to USA.[3] In this and other situations of highly skewed valued degree distributions, the dichotomization of valued blockmodels inevitably puts the emphasis on relative tie strengths at the macro-level, rather than how a dyad is perceived as prominent or not by its two actors.

Žiberna has suggested two alternative approaches to valued blockmodeling that exclude dichotomization of the raw data. In

[2] An introduction to blockmodeling and role-analysis is given in Ferligoj et al. (2011) – more thorough descriptions are found in Doreian et al. (2005) and Wasserman and Faust (1994, chapter 9–10).

[3] Total commodity trade from Colombia to USA in 2010 was valued at 17.1 billion USD. Representing 43% of Colombia's total exports for that year, it only corresponds to 0.9% of US total imports in 2010.

Table 1
Set of ideal block types used in generalized blockmodeling (Doreian et al., 2005, pp. 14ff, 212).[a]

Name	Ideal block condition(s)	Examples (non-diagonal blocks)	Examples (diagonal blocks)
Complete	A tie between all pairs of actors	$\begin{bmatrix} 1&1&1&1&1\\ 1&1&1&1&1\\ 1&1&1&1&1\\ 1&1&1&1&1 \end{bmatrix}$	$\begin{bmatrix} \backslash&1&1&1\\ 1&\backslash&1&1\\ 1&1&\backslash&1\\ 1&1&1&\backslash \end{bmatrix}$
Null	No ties between any pairs of actors	$\begin{bmatrix} 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0 \end{bmatrix}$	$\begin{bmatrix} \backslash&0&0&0\\ 0&\backslash&0&0\\ 0&0&\backslash&0\\ 0&0&0&\backslash \end{bmatrix}$
Regular	At least one tie in each row and column, respectively	$\begin{bmatrix} 1&0&1&0&0\\ 0&0&0&1&0\\ 0&1&0&0&1\\ 0&0&1&0&0 \end{bmatrix}$	$\begin{bmatrix} \backslash&1&1&0\\ 1&\backslash&1&0\\ 0&0&\backslash&1\\ 0&1&0&\backslash \end{bmatrix}$
Row-regular	At least one tie in each row	$\begin{bmatrix} 1&0&0&1&0\\ 0&1&1&0&0\\ 0&0&0&1&0\\ 0&0&1&0&0 \end{bmatrix}$	$\begin{bmatrix} \backslash&1&0&1\\ 1&\backslash&0&0\\ 0&0&\backslash&1\\ 0&1&0&\backslash \end{bmatrix}$
Column-regular	At least one tie in each column	$\begin{bmatrix} 1&1&0&0&0\\ 0&0&0&0&0\\ 0&0&1&0&1\\ 1&0&0&1&0 \end{bmatrix}$	$\begin{bmatrix} \backslash&0&1&0\\ 1&\backslash&0&1\\ 0&0&\backslash&0\\ 0&1&1&\backslash \end{bmatrix}$
Row-functional	Exactly one tie in each row	$\begin{bmatrix} 1&0&0&0&0\\ 0&1&0&0&0\\ 0&0&0&1&0\\ 0&0&1&0&0 \end{bmatrix}$	$\begin{bmatrix} \backslash&0&1&0\\ 0&\backslash&1&0\\ 1&0&\backslash&0\\ 0&1&0&\backslash \end{bmatrix}$
Column-functional	Exactly one tie in each column	$\begin{bmatrix} 1&0&0&1&0\\ 0&0&1&0&1\\ 0&1&0&0&0\\ 0&0&0&0&0 \end{bmatrix}$	$\begin{bmatrix} \backslash&0&0&1\\ 1&\backslash&1&0\\ 0&0&\backslash&0\\ 0&1&0&\backslash \end{bmatrix}$

[a] Excluded from this table, the set of ideal blocks in generalized blockmodeling also includes row-dominant and column-dominant ideal blocks (e.g. Doreian et al., 2005, p. 212, 224; Žiberna, 2007a, p. 108).

the first approach, the ideal blocks used in generalized blockmodeling are modified by introducing parametric vector functions in their definitions (Žiberna, 2007a, pp. 108, 111). For instance, for row(-f)-regular blocks (Žiberna, 2007a, p. 108) where $f(V) = max(V)$, at least one value in each block row vector V is equal to or above a predetermined network-wide parametric value m. If the function instead is the sum of values, row-f-regularity implies that the sum of all values in each block row is equal to or above m. Possibly combined with pruning exceptionally large values (Žiberna, 2007a, p. 125), this approach nevertheless assumes equal relational capacities among actors by determining the prominence of ties by comparison with the network-wide absolute parameter m.

Žiberna's second suggestion is to use homogeneity blockmodeling (2007a, p. 115), i.e. where the blockmodel is optimized to minimize intra-block variance (see also Borgatti and Everett, 1992; Hartigan, 1972). Proposing ways to identify ideal blocks in such homogenized blocks, homogeneity blockmodeling is by definition a partitioning of actors based on tie value similarities within blocks, i.e. explicitly focusing on strengths, rather than pattern variations, of ties.

In the heuristic proposed by Nordlund (2007), the valued matrix is marginal-normalized by rows and columns, respectively, resulting in an outbound (ORB) and an inbound (IRB) regular blockmodel. Each row in ORB contains the shares of the total valued outdegree for each row actor to potential alters, and columns in IRB similarly reflect the distribution of inflows to each column actor from potential alters. Dichotomizing these normalized matrices with a cutoff value that specifies tie prominence as a minimum share of total in- and outflows, respectively, occurrences of regular blocks are identified by simultaneously measuring criteria-fulfillment for row-regular blocks in ORB and column-regular blocks in IRB.

Whereas Nordlund's (2007) suggestion captures the prominence of valued relations on a per-actor basis, there are two drawbacks with his approach. First, the procedure discards variations within row- and column vectors that could be of relevance.

For example, when choosing a significance cutoff of 1/5, an outbound tie representing 20% of an actor's total outdegree would be deemed equally significant to ties corresponding to 30, 50 or 100% of an actor's total outflows.

Most importantly, although the suggestion in Nordlund (2007) indeed captures prominent ties on a per-actor basis, this operationalization of tie prominence still builds on an assumption of equal relational capacities among actors. By only determining tie prominence on the basis of one of the two actors in each dyad, Nordlund's (2007) heuristic ignores whether the relational capacities of *both* actors differ to such a degree that a particularly high (or low) share of an actor's out- or inbound tie constitutes an anomaly from what could be expected. This implies that ties identified as prominent typically appear in conjunction with actors having large in- and/or outdegrees. That is, a valued tie deemed significant in Nordlund (2007) might indeed be so for either of the actors in the dyad, but it does not necessarily constitute an anomaly from what could be expected on the basis of the overall valued degree distribution of the network.

Focusing explicitly on deviations from expected values, the approach proposed in this article differs from existing approaches in how the prominence of ties is perceived and operationalized. First, rather than applying an arbitrary threshold value at the network-wide or per-actor level (cf. Nordlund, 2007), the herein suggested approach conceptualizes prominence (and non-prominence) of valued ties in terms of unexpected anomalies at the dyadic level, determining such for each tie on both a per-actor-to-actor and a per-actor-from-actor basis. Secondly, whereas Nordlund (2007) only examined and measured occurrences of regular blocks in pre-partitioned networks, the herein suggested approach incorporates additional ideal block types, whether examining pre-partitioned networks, finding an optimal partition for a pre-given block image, or for simultaneously finding optimal partitions and block images. Thirdly, the proposed approach is by default non-parametric. Although an optional deviation threshold can be set when determining the prominence (and non-prominence) of ties, this parameter constitutes a percentual deviation from expected flows on a dyadic basis, independent of network size, valued degree distributions, and tie value spans. Related to this, fourthly, depending on how the optional deviation threshold parameter is applied, the heuristic could result in certain ties categorized as non-determined, i.e. being neither prominent (1) nor non-prominent (0). To cater for this, the proposed approach operationalizes a measure of interpretational uncertainty of blockmodels that, when used in conjunction with a two-sided deviation threshold, results in a measure of interpretational certainty and/or stability of a given blockmodel. Finally, although the suggested approach is primarily developed for direct blockmodeling, this article also proposes an indirect correlation-based measure of structural equivalence in valued networks that is based on the deviations between observed and expected values.

3. Mathematical derivation/specification of suggested approach

For a sociomatrix X, representing a directional valued one-mode network without self-ties, $x_{i,j}$ is the directed valued relation from i to j (where $x_{i,i}=0$ for all i), the valued in- and outdegree of actor k is $indeg(k)$ and $outdeg(k)$, respectively, and sum is the total sum of values in X. A particular blockmodel is specified by a partition P consisting of non-overlapping subsets (positions/clusters) of all actors, i.e. $P=\{C_A, C_B, C_C\}$. A block is the submatrix specified by two positions, where $B(C_A, C_B)$ is the set of ties from actors in C_A to actors in C_B. Diagonal (quadratic) blocks, e.g. $B(C_A, C_A)$, lack ties in their diagonals.

Identical to the ORB and IRB matrices in Nordlund (2007), two normalized versions of matrix X are created, marginal-normalized with respect to rows (RN) and columns (CN), respectively:

$$rn_{a,b} = \begin{cases} 0, & \text{if } outdeg(a) = 0; \\ \dfrac{x_{a,b}}{outdeg(a)}, & \text{if } outdeg(a) > 0; \end{cases} \tag{1a}$$

$$cn_{a,b} = \begin{cases} 0, & \text{if } indeg(b) = 0; \\ \dfrac{x_{a,b}}{indeg(b)}, & \text{if } indeg(b) > 0; \end{cases} \tag{1b}$$

Whereas Nordlund (2007) determined prominence and non-prominence of ties on the basis of RN and CN alone, the herein suggested approach compares these distributions with expected values based on the distribution of total in- and outflow. Prominent and non-prominent valued relations are those that deviate, respectively, positively and negatively, from what could be expected on the basis of aggregate in- and outflows of the actors for each dyad.

In valued networks with possible self-ties, the expected distribution of outbound and inbound valued ties for all actors would correspond to the normalized in- and outdegree distributions. For networks that lack self-ties, the expected distributions of values for outbound (RE) and inbound (CE) flows exclude the indegree and outdegree contributions from, respectively, the source and the destination of each dyad:

$$re_{a,b} = \frac{indeg(b)}{sum - indeg(a)};\ a \neq b \tag{2a}$$

$$ce_{a,b} = \frac{outdeg(a)}{sum - outdeg(b)};\ a \neq b \tag{2b}$$

Combining the above formulas, the calculated percentual differences between actual and expected flows end up in the two matrices RD and CD:

$$rd_{a,b} = \begin{cases} 0, & \text{if } indeg(b) = 0 \text{ or } outdeg(a) = 0; \\ \left(\dfrac{rn_{a,b}}{re_{a,b}}\right) - 1 = x_{a,b}\left(\dfrac{sum - indeg(a)}{outdeg(a)\cdot indeg(b)}\right) - 1, & \text{otherwise}; \end{cases} \quad a \neq b \tag{3a}$$

$$cd_{a,b} = \begin{cases} 0, & \text{if } outdeg(a) = 0 \text{ or } indeg(b) = 0; \\ \left(\dfrac{cn_{a,b}}{ce_{a,b}}\right) - 1 = x_{a,b}\left(\dfrac{sum - outdeg(b)}{indeg(b)\cdot outdeg(a)}\right) - 1, & \text{otherwise}; \end{cases} \quad a \neq b \tag{3b}$$

Positive (negative) values in RD indicate that the valued tie from i to j is $rd_{i,j}$ percent higher (lower) than expected from the point of view of i, and positive (negative) values in CD indicate that the flow to j from i is $cd_{i,j}$ percent higher (lower) than expected from j's point of view.[4] Ranging from negative one and upwards, a value of zero implies no deviation from what is expected.

Based on the deviations in RD and CD, two binary matrices – RB and CB – are created with values indicating prominent (1) and non-prominent (0) valued relations. In its default (non-parametric) form, the values in RB and CB used for subsequent comparisons with ideal block types correspond to positive and negative deviations in, respectively, RD and CD:

$$rb_{a,b} = \begin{cases} 1, & \text{if } rd_{a,b} > 0; \\ 0, & \text{if } rd_{a,b} \leq 0; \end{cases}, \tag{4a}$$

[4] As evident in formulas (3a) and (3b), the sign and magnitude of the deviations depend on how dyadic values relate to total in- and outdegrees of respective actor in each dyad. Thus, although deviations in RD and CD do depend on the actual values in X, it is indeed possible that the largest tie value actually could have corresponding negative values in RD and CD, just as a very small (non-zero) valued relation could be substantially larger than what could be expected based on corresponding in- and outdegrees, resulting in very large deviations in RD and CD. Occurrences of both these phenomena are demonstrated in the example on EU/EFTA trade below.

$$cb_{a,b} = \begin{cases} 1, & \text{if } cd_{a,b} > 0; \\ 0, & \text{if } cd_{a,b} \leq 0; \end{cases} \tag{4b}$$

As in generalized blockmodeling, the number of discrepancies between ideal and observed blocks constitute a penalty score for each block. However, whereas penalty scores in conventional blockmodeling are calculated on the basis of a singular sociomatrix, the simultaneous analysis of *RB* and *CB* implies modifications to how such penalties are calculated – see the *p* formulas in Table 2. Row-regular blocks are identified by examining *RB*, column-regular blocks are determined by examining *CB*, and their combined occurrences in corresponding blocks indicate a regular block. Two types of penalties for regular blocks (including derivatives) are provided here: the heavy-penalty types (*regg*, *rreg*, *creg*, *rfng*, *cfng*) recommended for generalized blockmodeling (see Doreian et al., 2005, p. 225), where the number of null rows and columns are multiplied by, respectively, the number of columns and rows in the block, and the light-penalty types (*regr*, *rrer*, *crer*, *rfnr*, *cfnr*) that exclude such block penalty weighting (Doreian et al., 2005, p. 187; Nordlund, 2007, p. 62). As complete and null blocks are determined by simultaneously inspecting both *RB* and *CB*, it is suggested to separate these two structural block types into strong and weak varieties, differing by the criteria for identifying such blocks. Strong null blocks (complete blocks) are identified by occurrences of 0-cells (1-cells) in the corresponding blocks in *both RB* and *CB*, whereas weak null blocks (complete blocks) are identified by occurrences of 0-cells (1-cells) in *either RB* or *CB*.[5] The penalty table also includes light-penalty weak and strong null blocks (nulwr, nulsr), reflecting the null block penalty recommended for regular equivalence blockmodeling (Doreian et al., 2005, p. 187). Aggregating the penalty scores for all blocks results in an overall goodness-of-fit measure that either describes a given partition or that can be used as a fitting function for finding an optimal partition.

4. The optional deviational threshold: measuring interpretational uncertainty

An optional (one-sided) percentual threshold parameter can be applied, where only those values in *RD* and *CD* equal to or above this threshold are coded as ties in *RB* and *CB*, and values whose *RD* and *CD* values below the deviation threshold are coded as zero (see Formulas (5a) and (5b)). A two-sided alternative only identifies non-prominent ties as deviations below the *negative* threshold value (see Formulas (6a) and (b)). For instance, with a two-sided 5% deviation threshold, deviations above 0.05 and below −0.05 are, respectively, deemed prominent (1) and non-prominent (0). For observed deviations within this range, ties are neither prominent nor non-prominent, but instead non-determined.

$$rb_{a,b} = \begin{cases} 1, & \text{if } rd_{a,b} \geq \text{cutoff}; \\ 0, & \text{if } rd_{a,b} < \text{cutoff}; \end{cases}, \tag{5a}$$

$$cb_{a,b} = \begin{cases} 1, & \text{if } cd_{a,b} \geq \text{cutoff}; \\ 0, & \text{if } cd_{a,b} < \text{cutoff}; \end{cases} \tag{5b}$$

$$rb_{a,b} = \begin{cases} 1, & \text{if } rd_{a,b} \geq \text{cutoff}; \\ 0, & \text{if } rd_{a,b} \leq -\text{cutoff}; \\ \text{undefined}, & \text{otherwise}; \end{cases}, \tag{6a}$$

$$cb_{a,b} = \begin{cases} 1, & \text{if } cd_{a,b} \geq \text{cutoff}; \\ 0, & \text{if } cd_{a,b} \leq -\text{cutoff}; \\ \text{undefined}, & \text{otherwise}; \end{cases} \tag{6b}$$

With the content of *RB* and *CB* depending on whether the default or parametric conversion is used, possibly resulting in different optimal blockmodel configurations in the direct blockmodeling approach, the non-parametric identification of prominent and non-prominent ties (Formulas (4a) and (4b)) is suggested in the general case. However, there might be both theoretical and substantial reasons for applying a one- or two-sided deviation threshold to identify prominent and non-prominent ties. Although no general guidelines can be given whether, and at what value, to apply a one- or two-sided deviational threshold, a particular research question or theory could for example dictate that non-prominent ties must deviate to a certain percentage below what is expected. Also, if the dataset is suspected of containing a certain amount of measurement errors, the researcher could choose a suitable two-sided deviation threshold to reduce occurrences of would-be "white noise".

The recommended usage of deviational thresholds is however not as an inbound parameter but rather as a measure of interpretational uncertainty. Whereas the penalty scores for the various ideal blocks are unaffected by would-be non-determined ties (see Table 2), occurrences of such ties could however lead to interpretational uncertainties. For instance, if an empirical block in *RB* is row-regular with the exception of a single row lacking prominent ties, instead having one, many, or only non-determined cells, we do not know for certain that this is an ideal row-regular block. For each block, a measure of interpretational uncertainty is proposed that reflects how certain we are when we state that a particular empirical valued block conforms to a particular ideal block to a certain degree. The interpretational uncertainty indices for each ideal block type are given by the *u* formulas in Table 2.

By aggregating the block-wise uncertainty indices for all blocks in a blockmodel, we arrive at a measure of interpretational uncertainty for blockmodels where non-determined cells are a possibility. Although these uncertainty scores can be interpreted as is, it is instead proposed to combine such with the deviation threshold. As an increase in the two-sided deviation threshold monotonically increases occurrences of non-determined ties and, by extension, the interpretational uncertainty of blocks and the blockmodel as a whole, we can determine the maximum deviation threshold that keeps the uncertainty at a zero (or near-zero) value for a given blockmodel. For instance, if a deviation threshold larger than *m* results in a non-zero uncertainty value, we could state that the actual blockmodel interpretation, with its particular penalty score, is certain up to the deviation threshold of *m*. Doing so, the two-sided deviation threshold switches from being an inbound parameter to instead being an indicator of interpretational certainty.

5. An indirect measure of deviational structural equivalence

A well-established indirect measure of structural equivalence of valued networks is the correlation coefficients of row and column vectors for each pair of actors in the sociomatrix. Whereas this conventional approach calculates such coefficients using the original dataset, the herein suggested approach instead uses deviations from expected values as input. For each pair of actors, the correlation between their respective row vectors in *RD* and column vectors in *CD* is proposed as a measure of deviational structural equivalence. As these are calculated using data from *two* matrices, a modified formula for Pearson cross-product correlation (ignoring

[5] When only fitting (weak and strong) complete and null blocks in a blockmodel, a translation table allows for converting the *RD* and *CD* matrices directly into a single binary matrix – see Table 3.

Table 2
Penalty and uncertainty scores for different types of ideal blocks (Iverson square brackets, i.e. true → 1, false → 0; *AND*, *OR*: Boolean operators; P_A: set of actors in position A; $card(P_A)$: number of actors in position A; ? = non-determined; a ≠ b in all functions).

Block type	Notation	Penalty (p) and uncertainty (u) functions
Complete (strong)	coms	$p_{\text{coms}} = \sum_{a\in P_R}\sum_{b\in P_C}([(rb_{a,b}=0)\,\text{OR}\,(cb_{a,b}=0)])$
		$u_{\text{coms}} = \sum_{a\in P_R}\sum_{b\in P_C}([(rb_{a,b}=?)\,\text{OR}\,(cb_{a,b}=?)])$
Complete (weak)	comw	$p_{\text{comw}} = \sum_{a\in P_R}\sum_{b\in P_C}([(rb_{a,b}=0)\,\text{AND}\,(cb_{a,b}=0)])$
		$u_{\text{comw}} = \sum_{a\in P_R}\sum_{b\in P_C}([(rb_{a,b}=?\ \text{AND}\ cb_{a,b}=0)\,\text{OR}\,(rb_{a,b}=0\ \text{AND}\ cb_{a,b}=?)])$
Null (strong)	nuls/nulsr	$p_{\text{nuls}} = \sum_{a\in P_R}\sum_{b\in P_C}([(rb_{a,b}=1)\,\text{OR}\,(cb_{a,b}=1)])$
		$p_{\text{nulsr}} = \sum_{a\in P_R}\left(\left[\sum_{b\in P_C}(rb_{a,b}+cb_{a,b})>0\right]\right)+\sum_{b\in P_C}\left(\left[\sum_{b\in P_R}(rb_{a,b}+cb_{a,b})>0\right]\right)$
		$u_{\text{nuls}} = \sum_{a\in P_R}\sum_{b\in P_C}([(rb_{a,b}=?)\,\text{OR}\,(cb_{a,b}=?)])$
Null (weak)	nulw/nulwr	$p_{\text{nulw}} = \sum_{a\in P_R}\sum_{b\in P_C}([(rb_{a,b}=1)\,\text{AND}\,(cb_{a,b}=1)])$
		$p_{\text{nulwr}} = \sum_{a\in P_R}\left(\left[\sum_{b\in P_C}(rb_{a,b}\cdot cb_{a,b})>0\right]\right)+\sum_{b\in P_C}\left(\left[\sum_{a\in P_R}(rb_{a,b}\cdot cb_{a,b})>0\right]\right)$
		$u_{\text{nulw}} = \sum_{a\in P_R}\sum_{b\in P_C}([(rb_{a,b}=?\ \text{AND}\ cb_{a,b}=1)\,\text{OR}\,(rb_{a,b}=1\ \text{AND}\ cb_{a,b}=?)])$
Row-regular	rrer/rreg	$p_{rrer} = \sum_{a\in P_R}\left(\left[\left(\sum_{b\in P_C} rb_{a,b}\right)=0\right]\right)$; $p_{rreg} = p_{rrer}\times card(P_C)$
		$u_{rre} = \sum_{a\in P_R}\left(\left[\left(\left(\sum_{b\in P_C} rb_{a,b}\right)=0\right)\ \text{AND}\ \left(\left(\sum_{b\in P_C}[rb_{a,b}=?]\right)>0\right)\right]\right)$
Column-regular	crer/creg	$p_{crer} = \sum_{a\in P_C}\left(\left[\left(\sum_{b\in P_R} cb_{a,b}\right)=0\right]\right)$; $p_{creg} = p_{crer}\times card(P_R)$
		$u_{cre} = \sum_{b\in P_C}\left(\left[\left(\left(\sum_{a\in P_R} cb_{a,b}\right)=0\right)\ \text{AND}\ \left(\left(\sum_{a\in P_R}[cb_{a,b}=?]\right)>0\right)\right]\right)$
Regular	*regr/regg*	$p_{regr} = p_{rrer} + p_{crer}$; $p_{regg} = p_{rreg} + p_{creg}$
		$u_{reg} = u_{rre} + u_{cre}$
Row-functional	*rfnr/rfng*	$p_{rfnr} = \sum_{a\in P_R}\left(\left[\left(\sum_{b\in P_C} rb_{a,b}\right)\neq 1\right]\right)$; $p_{rfng} = p_{rfnr}\times card(P_C)$
		$u_{rfn} = \sum_{a\in P_R}\left(\left[\left(\left(\sum_{b\in P_C} rb_{a,b}\right)\neq 1\right)\ \text{AND}\ \left(\left(\sum_{b\in P_C}[rb_{a,b}=?]\right)>0\right)\right]\right)$
Column-functional	*cfnr/cfng*	$p_{cfnr} = \sum_{a\in P_C}\left(\left[\left(\sum_{b\in P_R} cb_{a,b}\right)\neq 1\right]\right)$; $p_{cfng} = p_{cfnr}\times card(P_R)$
		$u_{cfn} = \sum_{b\in P_C}\left(\left[\left(\left(\sum_{a\in P_R} cb_{a,b}\right)\neq 1\right)\ \text{AND}\ \left(\left(\sum_{a\in P_R}[cb_{a,b}=?]\right)>0\right)\right]\right)$

diagonal elements), with support functions for means and standard deviations, are given below.

$$\mu_a = \frac{\sum_{i\neq a, i\neq b}(rd_{a,i}+cd_{i,a})}{2\cdot(N-2)} \tag{7}$$

$$\sigma_a = \sqrt{\frac{\sum_{i\neq a, i\neq b}((rd_{a,i}-\mu_a)^2+(cd_{i,a}-\mu_a)^2)}{2\cdot(N-2)}} \tag{8}$$

$$devSE_{a,b} = \frac{1/(2\cdot(N-2))\cdot\sum_{i\neq a, i\neq b}(rd_{a,i}\cdot rd_{b,i}+cd_{i,a}\cdot cd_{i,b})-(\mu_a\cdot\mu_b)}{\sigma_a\cdot\sigma_b} \tag{9}$$

Similar to the conventional correlation measure of structural equivalence, suitable partitions can be obtained through cluster analysis of the *devSE* matrix. The resulting blockmodel can then be interpreted using *RD* and *CD* matrices as guidance. This indirect approach differs from the direct approach by using the full range of values in *RD* and *CD*: rather than determining the optimal partition on the basis of the binary patterns in *RB* and *CB*, the categorization

Table 3
1/0-cell translation table for (weak and strong) null and complete blocks.

	rd<-cutoff	-cutoff<=rd<=cutoff	rd>cutoff
cd<-cutoff	0	nulw: 0 nuls: non-determined	Contradictory
-cutoff<=cd<=cutoff	nulw: 0 nuls: non-determined	Non-determined	comw: 1 coms: non-determined
cd>cutoff	Contradictory	comw: 1 coms: non-determined	1

of prominent and non-prominent ties, possibly using a one- or two-sided deviation threshold, is done a posteriori in the indirect approach.

Examining *RB* and *CB*, possibly created with a deviation threshold (that possibly yields non-determined ties), structural blockmodels based on the indirect formula above are easier to interpret than what is the case when only having access to the raw valued data. As these blockmodels only contains the complete and null blocks, the translation table (Table 3) allows for converting (either using the weak or strong criteria for these blocks) the deviational values in RD and CD directly into a combined binary matrix with possible contradictory and non-determined ties.[6]

In what follows, the direct and indirect deviational approaches will be demonstrated using three datasets – the note-sharing dataset (Žiberna, 2007a; Hlebec, 1996), the social work journal citation dataset (Baker, 1992; see also Borgatti and Everett, 1999; Doreian et al., 2005, p. 265), and a dataset on total commodity trade between EU and EFTA countries for 2010 – comparing results with those obtained from more conventional approaches. In the note-sharing example, the direct (non-parametric) deviational approach finds an almost identical partition to Žiberna's "best fit" solution for the optimal blockimage that he finds, in addition also providing a measure of interpretational certainty. Without pre-specifying a particular blockimage, the direct approach finds different optimal blockimages and partitions than that found by Žiberna, solutions that also depend on whether the light or heavy penalty for regular blocks is used. For the journal citation data, the direct deviational blockmodeling approach confirms the core–periphery structure as identified in previous studies, although with a partition that differs slightly from the previous studies. Applying the indirect measure of deviational structural equivalence to the citation data yields clusters of journals that, seemingly, better reflect the various themes and focus areas of the journals than those obtained in Baker's original study. The final example on EU/EFTA commodity trade also underlines the fundamental difference between the deviational and conventional indirect blockmodeling approaches. Whereas a conventional indirect structural blockmodel of this data resemblances a traditional core–periphery block image, the deviational indirect approach yields a typical cohesive subgroup block image, indicating strong regional tendencies between countries in the European economy.

6. Applications/examples

6.1. Example dataset 1: note-sharing among students

Using data collected by Hlebec (1996) on the sharing of notes among 13 students, Žiberna (2007a) demonstrates his proposals for blockmodeling of valued networks. Representing individual estimates on how often each student has borrowed notes from alters,[7] the values of the directional ties ranges from 1 to 19. With 71 (non-zero) ties (with mean and median at 5.9 and 3, respectively), the topological (non-zero) density is 0.41. The valued (one-mode[8]) blockmodeling approach suggested by Žiberna, fitting null and sum-regular blocks with the *m* parameter set to 10 (Žiberna, 2007a, p. 119), results in the optimal 3-positional blockmodel and blockimage as given in Fig. 1. Arguing *f*-regular blocks using sum functions to be the most appropriate in this particular context (2007a, p. 117), Žiberna also finds that ideal max-regular blocks (with *m* set to 5) yield an identical blockmodel as well as a similar one where student 13 is in the second group.

Applying the deviational heuristic to the note-sharing data, pre-specifying the blockimage as suggested by Žiberna above (fitting null and heavy-penalty regular (*regg*) blocks), an exhaustive search for the optimal partition results in the blockmodel in Fig. 2. Noticeably, the only difference to Žiberna's findings is with regards to student 13, here placed in the second position. The aggregate penalty score for this optimal blockmodel is 9: finding five simultaneously occurring null block[9] discrepancies in *RB* and *CB*, an additional penalty of 4 is incurred by the empty column In block *B*(3,3).

Approximately a third of all (non-zero) ties, with values ranging from 1 to 8 (mean and standard deviation of 2.7 and 2.1, respectively), are categorized as non-prominent, where the values of the remaining prominent ties cover the whole range of values (with a mean and standard deviation of 7.7 and 5.7, respectively). *RB* and *CB* are here identical, with one exception: the measure of notes student 10 borrows from student 4 ($rd_{10,4}$ and $cd_{10,4}$ are, respectively, +1.9 and −2.6%). Despite the non-determined state of this particular flow, the conditions for column-regularity of block *B*(2,3) are however satisfied by the positive deviation, i.e. identified prominence, of $cd_{5,4}$. The $cd_{5,4}$ value constitutes the lowest absolute deviation in *RD* and *CD*: although corresponding to a valued tie of 16, the value of $cd_{5,4}$ is only +0.012. As the significance of this particular flow is crucial for the column-regularity of block *B*(2,3), a would-be two-sided deviation threshold above this value would thus add an uncertainty in the interpretation of this block. We can thus state that the certainty of this particular blockmodel interpretation is valid up to a two-sided deviational threshold of 0.012

Without a pre-specified blockimage, instead searching for the particular block image and partition that yields the lowest penalty score, the direct deviational approach yields different solutions than those above. Following Žiberna (2007a, p. 119), null and (heavy-penalty) regular blocks were fitted to a 3-positional (non-parametric) deviational blockmodel, resulting in the two rather different solutions (penalty = 5) in Fig. 3a and b.[10]

[6] Whereas non-determined ties only appear at a non-zero deviation threshold, contradictory ties can (as the examples will demonstrate) indeed appear in the default (non-parametric) calculations.

[7] Rows correspond to note-borrowers and columns to note-lenders, i.e. the actual notes and the information they contain goes from columns to rows.

[8] As pointed out by a reviewer of a previous version of this manuscript, it is possible that a two-mode blockmodeling approach (e.g. Doreian et al., 2005, p. 247ff) might be preferable than the one-mode approaches in Žiberna (2007a,b).

[9] As these null block discrepancies appear in both *RB* and *CB*, the penalties would be the same for both strong (*nuls*) and weak (*nulw*) null blocks (see Table 2).

[10] At the 4-positional partition, a single optimal solution was found (penalty = 7) that differs from the two 3-positional solutions in Fig. 3a and b.

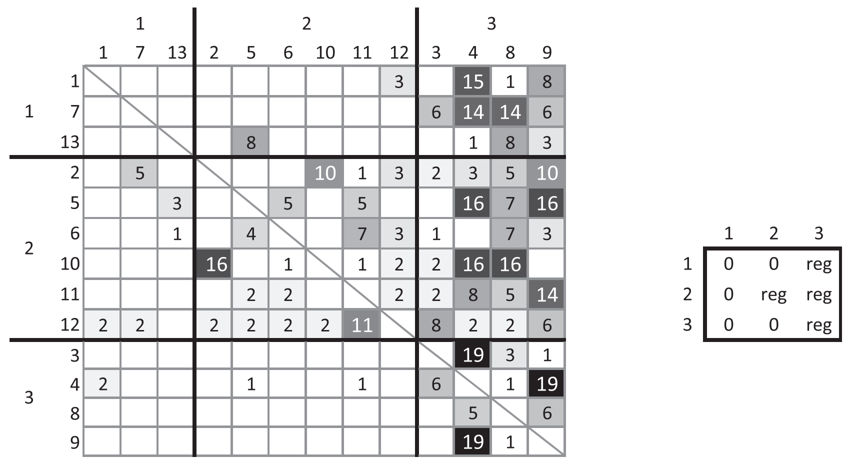

Fig. 1. "Best-fit" regular blockmodel and blockimage of note-sharing data using Žiberna's valued network approach (Žiberna, 2007a, p. 120).

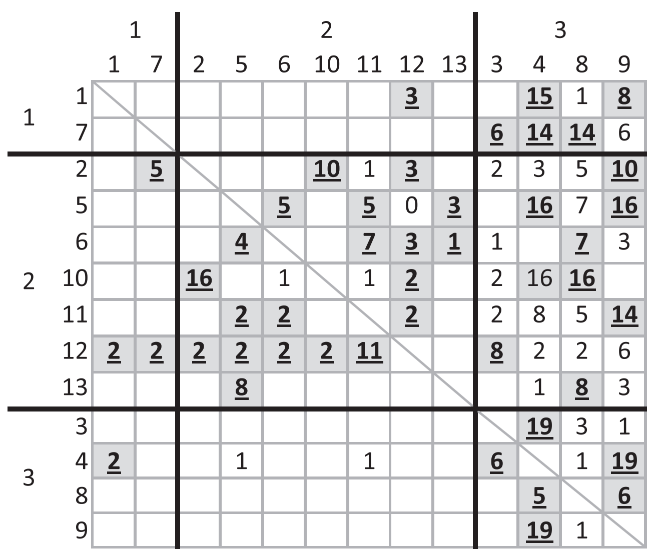

Fig. 2. Optimal regular blockmodel of note-sharing data using direct deviational approach (default (non-parametric) approach; pre-specified block image; shaded cells and bold/underline indicate ties in, respectively, *RB* and *CB*).

It is noteworthy that all of the penalties in the "free-search" solutions (Fig. 3a and b) stem from imperfect null blocks, possibly indicating that the generalized (heavy) penalty for regular blocks (*regg*) might be a bit too harsh in this context. Replacing[11] the penalty function for regular and null blocks with their lighter varieties (i.e. *regr* and *nulwr*; see Table 2), the optimal 3- and 4-positional solutions (with penalties at, respectively, 4 and 7; interpretational certainty values of 0.012) are more similar to each other as well as the solution found by Žiberna – see Fig. 3c–f. Whereas the almost identical 4-positional light-penalty solutions map with one of the 3-positional partitions (Fig. 3d), I deem[12] all four solutions (with corresponding block images; see Fig. 4) to be adequate solutions at, respective, the 3- and 4-positional partitions.

Despite the similarities between the optimal partitions and block images found when applying deviational blockmodeling (with light-penalty regular blocks) and the approach suggested by Žiberna on the note-sharing dataset, the two approaches capture different notions of what is meant by prominent ties in valued networks. There is thus no objective "best fit" blockmodel other than the one obtained using the approach that best represents what we intend to capture.

However, rather than providing a suitable function for the set of modified ideal block types and a threshold parameter, both reflecting particularities of the analyzed dataset, the deviational approach instead yields a measure of interpretational certainty of the resulting blockmodel. In addition, the provided categorization of ties into prominent and non-prominent, possibly also non-determined, allows for a direct (Boolean) comparison between the valued regular blocks of note-sharing patterns with the standard set of ideal blocks used in generalized blockmodeling.

6.2. Example dataset 2: Baker's social work journal citation

In his study of social work journals, Baker (1992) analyzes citation data between a set of 20 journals in 1985–1986. Whereas Baker determines journal prominence on the basis of a hierarchical clustering analysis, the valued directional data has also been used to exemplify the core–periphery methods of Borgatti and Everett (1999) as well as the 2-mode generalized blockmodeling approach of Doreian et al. (2005, p. 267ff). Whereas Borgatti and Everett (1999) symmetrized the original data and Doreian et al. excluded two of the journals, the analyses that follow use the original directional and valued 20-actor dataset as found in Baker (1992, p. 159).

Discarding all intra-journal citations, the 87 directional relations correspond to a topological density of 0.23. Relational values range from 2 to 124, with a median, mean and standard deviation at, respectively, 13, 20.9, and 21.8. The journal citation dataset is given in Table 4, sorted according to decreasing valued gross degree of journals.

Having calculated *RD* and *CD* for the citation data in Table 4, *RB* and *CB* are created using the default (non-parametric) formulas in (4a) and (4b). Of the 87 valued ties, 60 (69%) and 22 (25%) are

[11] As Ziberna only fits null and f-regular blocks to the note-sharing data in his example, i.e. as done in reqular equivalence blockmodeling, it makes sense to use their light-penalty functions (Doreian et al., 2005, p. 187) instead of those recommended for generalized blockmodeling.

[12] As in non-deviational blockmodeling approaches, direct as well as indirect, the choice of the number of positions/clusters (*k*) should not only depend on minimizing a criteria function but also rest on theoretical considerations and substantive information of the particular dataset (e.g. Doreian et al., 2005, p. 194, 232). Examining the partition lattice for various values of *k*, it can be noted that the second 3-positional solution (Fig. 3d) is connected to both the 2- and 4-positional solutions, as such indicating a higher degree of consistency for this particular 3-positional solution across different values of *k*, though both solutions are equally "correct" in a quantitative sense.

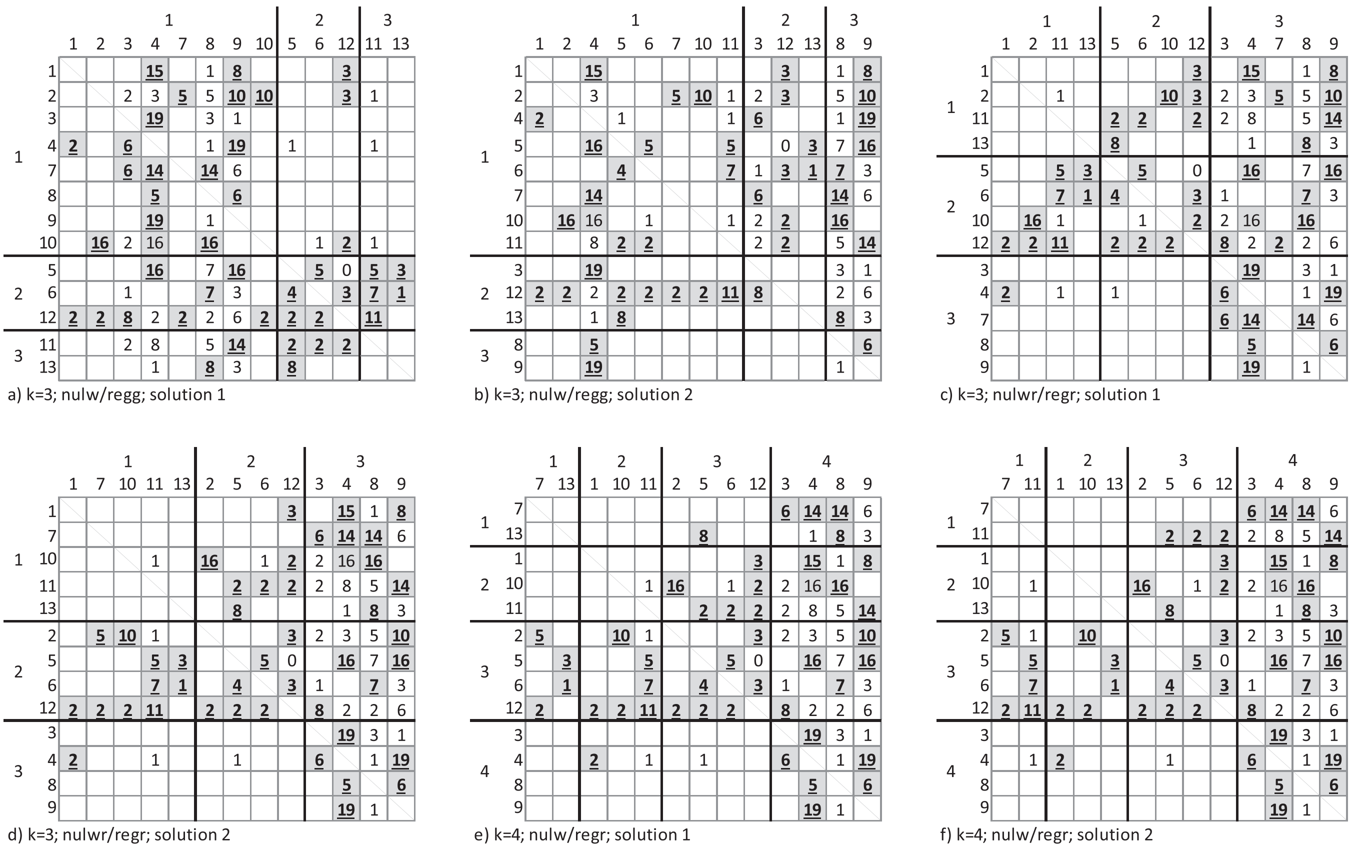

Fig. 3. (a–f) Optimal deviational blockmodels and partitions (k = 3, 4) for note-sharing data (light- vs heavy-penalty blocks).

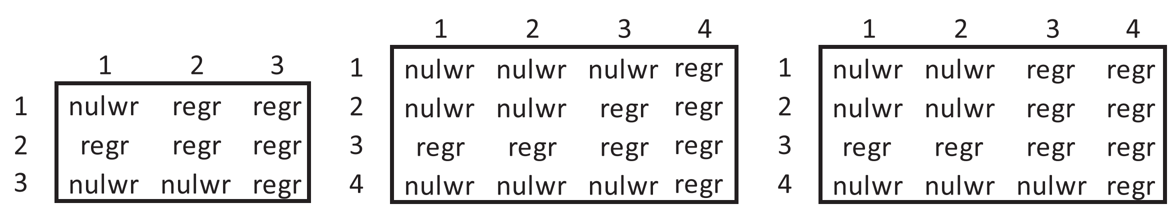

Fig. 4. Optimal block images (*nulwr* and *regr* ideal blocks; k = 3, 4) for direct deviational blockmodeling of note-sharing data (corresponding to blockmodels in, respectively, Fig. 3c–f).

deemed, respectively, prominent and non-prominent in both *RB* and *CB*. The remaining 5 ties are contradictory, i.e. where the corresponding values in *RD* and *CD* have different signs. The combined set of prominent ties in *RB* and *CB* ranges from 2 to 124 (mean at 22.0), and the ties classified as non-prominent in *RB* and *CB* range from 5 to 58 (mean at 17.8).

Both Baker (1992) and Borgatti and Everett (1999) find a distinct core–periphery structure in the journal citation data, although with different methods and ways of data handling. To control for journal size, Baker treats the data with repeated row and column marginal-normalization until convergence (1992, pp. 158, 166), subsequently applying agglomerative hierarchical clustering (aggregate Euclidean distances) to identify clusters of journals. The identified core cluster consists of 5 journals (SW, SCW, SSR, SWRA, JSWE), with five additional clusters (where three are deemed as roughly following journal themes). In Borgatti and Everett (1999), where the dataset[13] is symmetrized prior to analysis, both its valued and

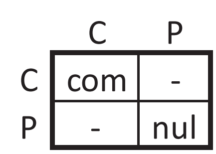

Fig. 5. Ideal core–periphery blockimage (ignoring off-diagonal blocks) as suggested by Borgatti and Everett (1999, p. 383).

dichotomized form is used to test different core–periphery models. Modeling intra-core and intra-peripheral ties as, respectively, a complete and a null block, they suggest a fitting function where non-diagonal ties are ignored (Borgatti and Everett, 1999, p. 383) – see Fig. 5. Using this model, the binary symmetrized[14] data reveals a 7-journal core (CW, CYSR, JSWE, SSR, SCW, SWRA and SW; fit:

[13] A couple of discrepancies can be noted between the symmetrized valued data in Borgatti and Everett (1999, p. 386) and the data found in Baker (1992, p. 159): the symmetrized SWRA-ASW relation should be 7 (not 20), the tie from ASW to JSWE should be 18 (not 8), and the SW to BJSW should be 19 (not 9), resulting in a miniscule difference in the optimal fitting value. The resulting partitions when using the symmetrized version of the original Baker data are however identical to those reported by Borgatti and Everett (1999).

[14] With non-symmetrized binary data, CYSR is placed in the periphery (Borgatti and Everett, 1999, p. 385).

Table 4
Baker citation data (valued; directional), excluding diagonal values, sorted by decreasing gross degree (indegree plus outdegree).

	SW	SCW	SSR	CW	JSWE	SWRA	ASW	CYSR	CSWJ	SWHC	SWG	PW	CAN	FR	JGSW	BJSW	CCQ	JSP	AMH	IJSW	Gross Degree	Full name
SW		58	53	52	33	8	15			43	15	19		9							967	Social Work
SCW	124		36	17	21	18	8		8	6	8		6	6	6						537	Social Casework
SSR	106	30		17	9	25	7														417	Social Service Review
CW	58	32	10		11			6				7	7								320	Child Welfare
JSWE	58	18	16			16	9			7											262	J of Social Work Education
SWRA	44	8	39	8	24																202	Social Work Research & Abstracts
ASW	73		21		18	7						13									171	Administration in Social Work
CYSR	28	8	14	70		5						6	12	4			5				158	Children & Youth Services Review
CSWJ	45	47	20																		120	Clinical Social Work J
SWHC	26	20																			111	Social Work in Health Care
SWG	40	9	7		9					9						2					99	Social Work with Groups
PW				4									7								56	Public Welfare
CAN	8	6		9																	55	Child Abuse & Neglect
FR	9	18																			46	Family Relations
JGSW	18	16																			40	J of Gerontological Social Work
BJSW	19				13																34	British J of Social Work
CCQ		3		12																	20	Child Care Quarterly
JSP			7																		7	J of Social Polocy
AMH	3																				3	Administration in Mental Health
IJSW	3																				3	Indian J of Social Work

Source: Baker (1992, p. 159).

Table 5
Core–periphery partitions of Baker citation data, various approaches.

	SW	SCW	SSR	CW	JSWE	SWRA	ASW	CYSR	...	IJSW	Penalties/optimal fit
Borgatti/Everett Core-Periphery	C	C	C	C	P	P	P	P	P	P	0.80 (corr), 0.93 (Hamming), 47.1 (density)
Ziberna ValBM [mean/max, m=13, 26]	C	C	C	C	C	P	P	P	P	P	92 (m=13), 185 (m=26)
Binary blockmodeling	C	C	C	C	C	C	P	P	P	P	10 penalties (2 in core, 8 in periphery)
Deviational [comw]	C	C	P	C	C	P	C	P	P	P	18 (7 in core, 11 in periphery)
Deviational [coms]: solution 1	C	C	P	C	C	P	C	P	P	P	21 (10 in core, 11 in periphery)
Deviational [coms]: solution 2	C	C	P	C	P	P	C	P	P	P	21 (5 in core, 16 in periphery)

0.86) whereas the valued symmetrized data results in a smaller core consisting of SSR, SW, and SCW (fit: 0.81–0.82; see footnote 14).

Fitting the Baker citation data in Table 4 to the ideal core–periphery model given in Fig. 5, different methods give different results[15] – see Table 5. The Borgatti/Everett core–periphery heuristic finds a 4-journal core consisting of SW, SCW, SSR and CW.[16] With Žiberna's valued blockmodeling approach, testing with both mean- and max-functions combined with median and double-median m parameter values, a 5-journal core consisting of the journals found with the Borgatti/Everett heuristic plus JSWE is identified. When applying binary blockmodeling, i.e. where tie values are ignored, SWRA joins the five actors identified as core by Žiberna's approach. It is noteworthy that the set of core journals identified by respective approach are consistently those with the largest gross degree and most in- and outbound ties.

As evident in Table 5, the results obtained from the deviational approach differ from those obtained using the more conventional approaches. Whereas the conventional approaches consistently categorize SSR as core, the deviational approach deems this journal to be peripheral. ASW, however, consistently categorized as peripheral in the conventional approaches, is deemed to be core in the deviational approach. It can be noted that the strong complete block criteria yields two equally optimal partitions, differing by the placement of JSWE. For the weak criteria, the singular optimal solution places JSWE in the core, as shown in the blockmodel in Fig. 6.

How certain and "stable" is the resulting blockmodel in Fig. 6? Applying a two-sided deviational threshold, increasing it from its default value of zero, a threshold value passing 0.028 results in the CW-to-JSWE going from prominent to non-determined ($cd_{CW,JSWE} = .028$; $rd_{CW,JSWE} = -.011$). Yielding an uncertainty penalty, the non-determined prominence of this tie results in two new optimal solutions: one with the same partition as before, but also a partition consisting of SW, SCW, SSR, CW and JSWE. Thus, the blockmodel in Fig. 6 is certain up to the two-sided deviational threshold of 0.028.

With non-diagonal blocks being ignored when fitting the blockmodel in Fig. 6, it can be noted that block B(C,P) is row-regular and that the B(P,C) block is column-regular. When fitting the network to the blockimage where these non-diagonal blocks are specified, we arrive at the same optimal solution as given in Fig. 5. The row-regular B(C,P) block thus implies that all core journals have prominent citations to at least one peripheral journal, although not all peripheral journals are prominently cited by core journals. All core journals are however prominently cited by peripheral journals, and almost all peripheral journals (with JSP being the exception) have prominent citations to core journals. Thus, apart from confirming[17] the existence of a core–periphery

[15] The Borgatti/Everett core–periphery results were obtained using the implementation in Ucinet (version 6.509), the Ziberna valued blockmodeling results were obtained using his R package "blockmodeling", and the binary blockmodeling results were obtained using Pajek64 (version 4.01a).

[16] This solution is obtained whether using correlation, Hamming or density as the fitting function. Observe that this partition was obtained using directional data, i.e. contrary to the analysis done in Borgatti and Everett (1999).

[17] In addition to optimizing the network according to these 2-positional block images, free-searching at $k = 3, 4, 5$ were also tested. All of these either yielded trivial results or resulted in blockmodels with higher penalty scores. Supported by previous studies, I conclude that Baker's journal citation data indeed is best seen as a core–periphery structure.

		SW	SCW	CW	JSWE	ASW	SSR	SWRA	CYSR	CSWJ	SWHC	SWG	PW	CAN	FR	JGSW	BJSW	CCQ	JSP	AMH	IJSW
		C					P														
C	SW		**58**	**52**	**33**	**15**	**53**	8			**43**	**15**	**19**		**9**						
	SCW	**124**		17	21	**8**	36	**18**		**8**	6	**8**		**6**	**6**	**6**					
	CW	**58**	**32**		**11**		10		**6**				**7**	**7**							
	JSWE	**58**	18			**9**	16	**16**			**7**										
	ASW	**73**			**18**		**21**	**7**					**13**								
P	SSR	**106**	30	17	9	**7**		**25**													
	SWRA	44	8	8	**24**		**39**														
	CYSR	28	8	**70**			14	5					**6**	**12**	**4**			**5**			
	CSWJ	45	**47**				**20**														
	SWHC	**26**	**20**																		
	SWG	**40**	9		**9**		7				**9**						**2**				
	PW			**4**										**7**							
	CAN	8	**6**	**9**																	
	FR	9	**18**																		
	JGSW	**18**	**16**																		
	BJSW	**19**			**13**																
	CCQ		**3**	**12**																	
	JSP						**7**														
	AMH	**3**																			
	IJSW	**3**																			

Model

	C	P
C	comw	-
P	-	nul

Penalties

	C	P
C	7	0
P	0	11

Fig. 6. Optimal core–periphery blockmodel of Baker citation data using direct deviational approach (default (non-parametric) approach; comw, nulw/nuls; pre-specified blockimage; shaded cells and bold/underline indicate ties in, respectively, *RB* and *CB*).

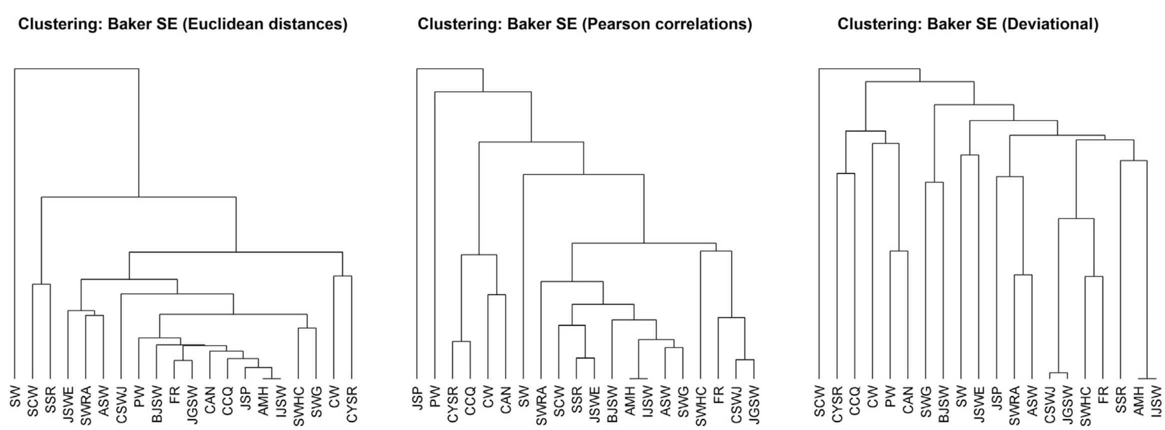

Fig. 7. Baker citation data: hierarchical clustering (weighted-average) of Euclidean, correlational, and deviational similarity measures.

structure as identified by both Baker (1992) and Borgatti and Everett (1999), the deviational approach also identifies patterns of prominent citations between core and peripheral journals, patterns that reflect ideal block types that, I argue, are reasonable to expect in the case of citation patterns of journals within a particular discipline.

6.3. *Baker citations: indirect measure of deviational structural equivalence*

Using the Baker citation data, this section compares the indirect measure of deviational structural equivalence with the more conventional indirect approaches to structural equivalence. For the latter, the indirect measures of structural equivalence are calculated using both the Euclidean distance measures and the Pearson cross-product correlation on the rows and columns in the raw citation data (Table 4). For the measures of deviational structural equivalence, formulas (7)–(9) are applied to the *RD* and *CD* matrices obtained from the Baker citation data. Subsets of structurally equivalent actors were identified through (weighted-average[18]) hierarchical clustering on the equivalence matrices for respective method – see Fig. 7. For comparative reasons, partitions were chosen at the 5- and 8-positional levels for respective approach, partitions given in Fig. 8.

Whereas the Euclidean measure of structural equivalence captures similarities based on tie strengths, the correlation measure is less sensitive to variations in mean and variance of tie values and is thus typically regarded as a better measure of pattern similarities (e.g. Wasserman and Faust, 1994, p. 374). This is evident in the dendrograms and partitions above: with the exception of CYSR at the 5-positional level, the gross degree span of journals in the Euclidean partitions does not overlap, whereas there is noticeable overlap in the gross degree spans of positional actors in the Correlation and Deviation partitions. For both the Euclidean and Correlation partitions, singleton positions dominate in both the 5- and 8-positional partition, where the journal with the most in- and outbound citations – Social Work – constitute a singleton position. This differs from the Deviational partition where SW is deemed equivalent to JSWE up until the 9-positional partition. In the Deviation partition, Social Casework (SCW) is identified as a singleton position at the top of the dendrogram, whereas this journal is part of the largest position in the Correlation partition.

In Baker's 1992 study, five journal positions were identified roughly following journal themes, and an additional residual position was found consisting of the three journals with the smallest

[18] Weighted-average clustering was chosen as it produced the least number of singleton positions in the three varieties. For consistency, this clustering approach was also used in the example on EU/EFTA trade.

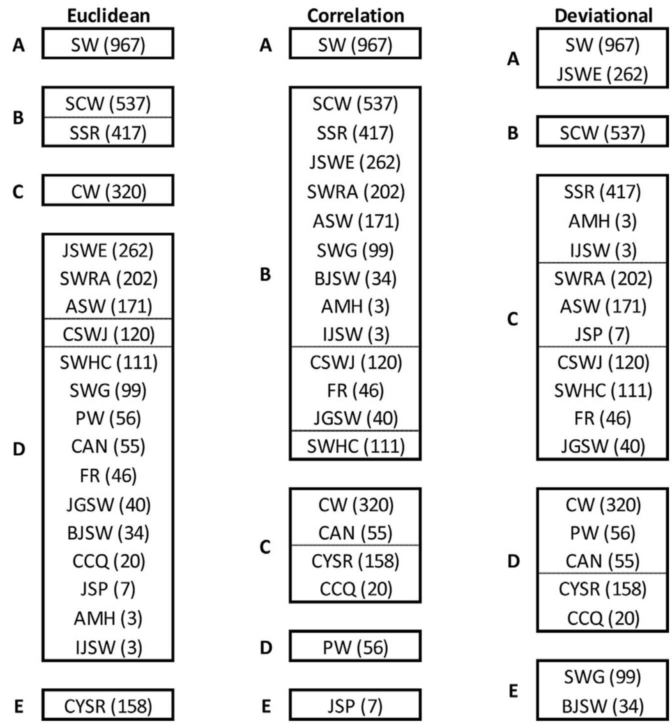

Fig. 8. Obtained partitions for Baker citation data using Euclidean, correlation and deviational approaches (journal gross degrees in brackets).

gross degrees: JSP, IJSW, and AMH. Apart from a core position consisting of 5 of the 6 journals with the largest gross degree, Baker finds 3 thematic positions consisting of, respectively, administrative (ASW, PW), child welfare (CYSR, CW, CAN), and method and practice (SWHC, SWG) journals, as well as a mixed position (CSWJ, JGSW, FR, CCQ, BJSW), covering gerontology, child care and family relations.

Whereas journal themes are poorly reflected in the Euclidean partition, the Correlation partition performs better at capturing such intuitive subsets. At the 5-positional level, we find 4 journals concerned with child welfare (CW, CAN, CYSR, CCQ), and at the 8-positional level we find two clinical-oriented journals – CSWJ and JGSW – forming a position together with FR. However, examining the Deviational partition at the 5- and 8-positional level, the herein suggested indirect metric seems to perform even better. Similar to the child welfare position in the Correlation partition, the same journals form a thematic cluster together with PW already at the 5-positional level. Similarly, at the 8-positional partition, the clinical-oriented position identified in the Correlation partition is equally found in the Deviational partition, this time however complemented by Social Work in Health Care (SWHC). Contrary to both Baker's findings and the Euclidean and Correlation partitions, the deviational approach puts two generic social work journals – SW and JSWE – in their own position, whereas the casework-oriented journal is found in its own singleton position. A somewhat surprising finding is that the two least connected (and perfectly structurally equivalent) journals AMH and IJSW form a position together with the well-connected and high-degree journal SSR.

To construct the deviational blockmodel, we use a two-sided deviational threshold to identify prominent (1) and non-prominent (0) ties. The deviational values in *RD* and *CD* are translated into four different states using a weak 1/0-cell criteria (see Table 3). Beginning at the default deviational threshold at zero where there are no non-determined ties, 5 cells are contradictory. Increasing the two-sided deviational threshold, the contradictory ties decrease as the number of non-determined ties increase, with some minor changes occurring regarding the number of determined ties. As the two-sided deviational threshold reaches 0.103, the final contradictory tie disappear, resulting in a total of 60 1-cells, 21 0-cells, and 6 non-determined ties.

Using this two-sided deviational threshold of 0.103, the blockmodel for the 5- and 8-positional Deviational partition is given in Fig. 9. In this figure, prominent and non-determined ties are shaded in, respectively, dark and light gray, and the 8-positional partition is indicated by dashed lines.

The density blockimage for a deviational structural blockmodel is determined in the conventional way, i.e. by dividing the number of 1-cells in a block with its total number of cells. For each block, the share of non-determined cells are also calculated, a block-wise uncertainty measure that can either be represented separately or by stating the density of a block as a value range.

Comparing the obtained density blockimages from the partitions obtained from the deviational and correlation-based indirect

		A		B	C									D					E	
		SW	JSWE	SCW	SSR	AMH	IJSW	SWRA	ASW	JSP	CSWJ	SWHC	FR	JGSW	CW	PW	CAN	CYSR	CCQ	SWG	BJSW
A	SW		33	58	53			8	15			43	9		52	19				15	
	JSWE	58		18	16			16	9			7									
B	SCW	124	21		36			18	8		8	6	6	6	17		6			8	
C	SSR	106	9	30				25	7						17						
	AMH	3																			
	IJSW	3																			
	SWRA	44	24	8	39										8						
	ASW	73	18		21			7								13					
	JSP				7																
	CSWJ	45		47	20																
	SWHC	26		20																	
	FR	9		18																	
	JGSW	18		16																	
D	CW	58	11	32	10											7	7	6			
	PW														4		7				
	CAN	8		6											9						
	CYSR	28		8	14			5					4		70	6	12		5		
	CCQ			3											12						
E	SWG	40	9	9	7							9									2
	BJSW	19	13																		

Fig. 9. Deviational structural blockmodel of Baker citation data, indirect approach.

Deviational, 5-positional

	A	B	C	D	E
A	100		35-40	20	25
B	0-50		50-60	20	50
C	40-45	40	8	2	
D	0-20	60	2	55	
E	100		5		50

B: Singleton

Correlational, 5-positional

	A	B	C	D	E
A		62	25	100	
B	85	17	6	8	
C	75	10	33		
D					
E					

A, D, E: Singleton

Deviational, 8-positional

	A	B	C1	C2	C3	D1	D2	E
A	100		17-34	50	38	33		25
B	0-50		0-33	67	75	33		50
C1	50			22				
C2	50		33	17		11		
C3	25-38	100	8					
D1	0-33	67				83	17	
D2		50			13	67	50	
E	100				13			50

B: Singleton

Correlational, 8-positional

	A	B1	B2	B3	C1	C2	D	E
A		67	33	100	50		100	
B1	78	28	4	11	17		11	
B2	100	15						
B3	100	11						
C1	100	17			50			
C2	50	11			75			
D								
E								

A, B3, D, E: Singleton

Fig. 10. Density blockmodels (percentages) for Baker data, correlational and deviational approaches, 5- and 8-positional partitions (empty blocks correspond to densities of zero).

measures of structural equivalence (Fig. 8), the density blockimages for respective approach at the 5- and 8-positional partitions are given in Fig. 10. Density ranges are stated for those blocks in the deviational density blockimages that contain non-determined cells. The correlation-based density blockimages were calculated from corresponding blockmodels where the raw valued data was dichotomized at a cutoff of 8, resulting in 64 prominent ties, this being closest[19] to the 60 prominent (and 6 non-determined) ties in the deviational structural blockmodel in Fig. 9.

The indirect deviational approach seemingly outperforms the conventional approaches with regards to capturing journal-thematic clusters. By applying an incremental (post-partition) two-sided deviation threshold, the threshold value that eliminates all contradictory ties from the blockmodel is used as a measure of its interpretational certainty. Translating the deviational values in *RD* and *CD* into prominent, non-prominent and non-determined ties, here using a weak 1/0-cell criterion, the resulting blockimage can be readily interpreted in terms of block-wise densities of prominent ties. In this case, the resulting density blockimage reflects a network structure that, similar to what is the case for the original dataset, is connected, something which only a cutoff at 7 or less would yield for the conventional correlational approach.

6.4. Example dataset 3: EU/EFTA commodity trade flow example

The final example dataset consists of total bilateral commodity trade value as of 2010 (UNCTAD, n.d.), measured in million USD, between 30 countries within EU and EFTA. Croatia, joining the EU in 2013, and Liechtenstein, whose trade data is part of that of Switzerland, are excluded. With total traded values at 3.4 trillion USD, the mean and median flows are 3.9 billion and 482 million, respectively. The largest tie value (103 bn USD) is the trade from Germany to France, and although Iceland and Cyprus do trade with each other, the low value of their two bilateral flows results in the only statistical zeros in the dataset.[20]

With the default (non-parametric) interpretation of prominence, 14 contradictory ties are found in RD and CD. Applying a two-sided deviation threshold, contradictory ties disappear at a deviational cutoff at 0.083, where the combined tally of tie categorizations in *both RB* and *CB* is 416 (24%) prominent, 1236 (71%) non-prominent, and 88 (5%) non-determined ties. With a strong 1/0-cell criteria (see Table 3), i.e. where a tie is deemed prominent (non-prominent) if corresponding values in *RD* and *CD* are *both* above (below) the (negative) deviation threshold, we arrive at 200 prominent ties. When determining the prominence of ties using an absolute cutoff in the original valued matrix, the same number of prominent ties – 200 – appears at a cutoff of 2848 million USD.

Building on the traditional gravity model in economic geography, we begin with the hypothesis that countries tend to form clusters of more intense trade on the basis of their spatial proximity. Roughly following De Blij and Muller (2004), we pre-partition the actors into five regions: West, Mediterranean, East, North, and the Baltics. Sorting actors in each position by decreasing gross trade, the blockmodel in Fig. 11a depicts the 200 prominent ties as determined by an absolute-value dichotomization at 2848 million USD. With 50 of these prominent ties are intra-West and an additional 108 ties connect Western and non-Western countries, the Baltic countries, Cyprus, Malta and Iceland are depicted as complete isolates due to lacking any bilateral trade flows above the stipulated 2848 million USD threshold.

Using the same partition as above but instead categorizing ties in terms of deviations from expected, Fig. 11b highlights the 200 prominent ties as determined by the two-sided deviational threshold of 0.083. Compared to the intra-Western cohesiveness in Fig. 11, the intra-West block resemblances more a regular block in the deviational approach. Other intra-positional densities increase down the diagonal, with the Baltic (1.0) and the North (.90), possibly also the East (.79), being better candidates for diagonal complete blocks.

Whereas most prominent ties in Fig. 11b seem intuitive from the theory of trade regionalism, the tie from Cyprus to Latvia is an interesting case study. Although the 22 million USD of trade from Cyprus to Latvia is less than 1% of the stipulated dichotomization threshold in Fig. 11a, its deviational values of +3.0 in both *RD* and *CD* clearly categorize this flow as prominent for both these countries. Equally significant is the very special economic relationship between these two countries. With the flight of capital from Cyprus in the last decade (Sprūds, 2012), further accelerated by the 2013 bailout of Cypriot banks (Eglitis, 2013), the relatively weak regulations of Latvian banks have made the latter a significant

[19] As pointed out by a reviewer, it is somewhat problematic to choose a cutoff that results in a given number of prominent ties. However, as the aim here is to compare and highlight the differences between how the deviational and conventional indirect correlation determine prominent ties, I argue that it makes sense to keep the densities of respective blockmodel as similar as possible in this case. The same approach is used in the subsequent example on EU/EFTA trade.

[20] All example datasets available in the Appendix and as part of the demonstrational software client.

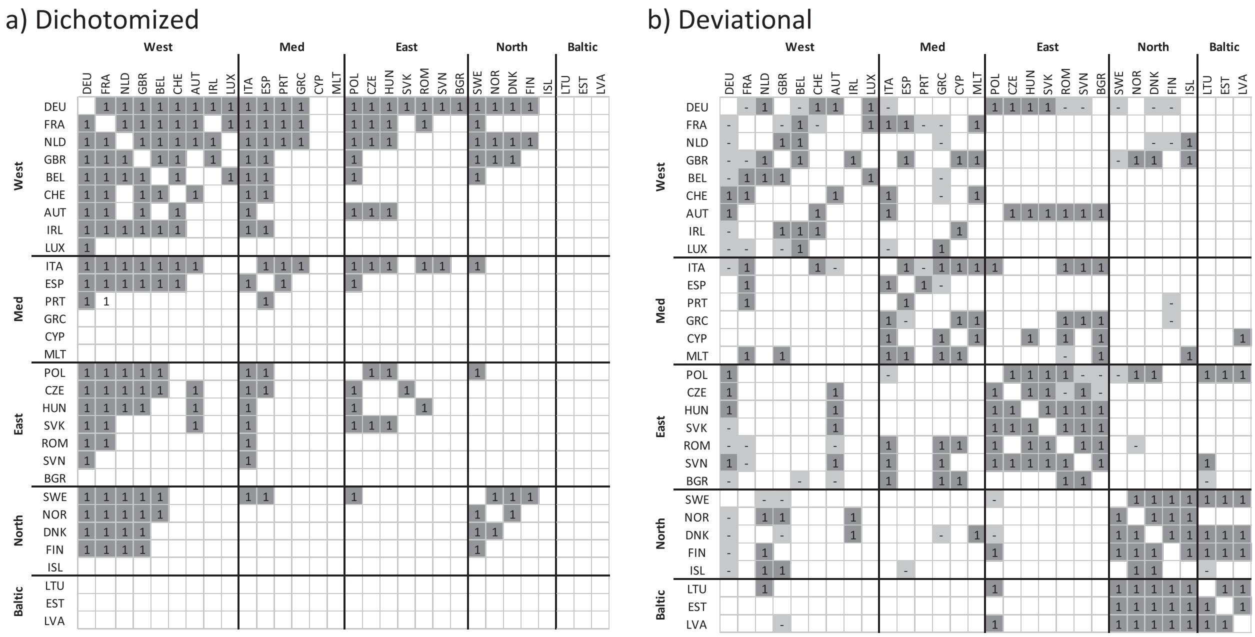

Fig. 11. (a and b) EU/EFTA trade: pre-specified regional partition; (a) dichotomized valued data (cutoff = 2848 million USD; (b) deviational approach (two-sided threshold = 0.083; strong 1/0-cell criteria; non-determined as dashes).

recipient of Russian money leaving Cyprus (Eglitis, 2013). Facilitating this, with ministerial negotiations beginning in 2006 (Consulco, 2011), a bilateral double-tax agreement between the two countries was approved by the Latvian government in 2014. It thus seems likely that the observed deviational significance of this particular bilateral trade flow in commodities indeed reflects this unique economic relationship between Cyprus and Latvia and an ongoing transfer of funds from the latter to the former (see also Brovkin, 2001).

Contrasting the identified prominence of the Cyprus–Latvia trade link, it is noteworthy that the largest absolute trade flow, from Germany to France, is not deemed as prominent in the deviational blockmodel above. Although prominent in an absolute sense, this trade flow evidently does not constitute an anomaly at the chosen deviation threshold with respect to total German exports ($rd_{DEU,FRA} = 0.005 < 0.083$), only with respect to total French imports ($cd_{DEU,FRA} = 0.129 > 0.083$).

Compared to approaches where tie prominence is determined using a network-wide dichotomizing cutoff value, the proposed deviational heuristic lends more support to regional theories of trade. In what follows, the indirect measure of deviational structural equivalence lends additional support for regionalism in intra-European commodity trade.

6.5. Indirect approach: deviational structural equivalence in European trade patterns

This section analyzes the EU/EFTA trade data using both the conventional and deviational indirect approaches to structural equivalence, comparing the partitions and blockmodels obtained from respective approach. Beginning with the conventional approach, this example only uses the Pearson cross-product correlation of pair-wise rows and columns in the original valued network, excluding all correlations involving the diagonal.

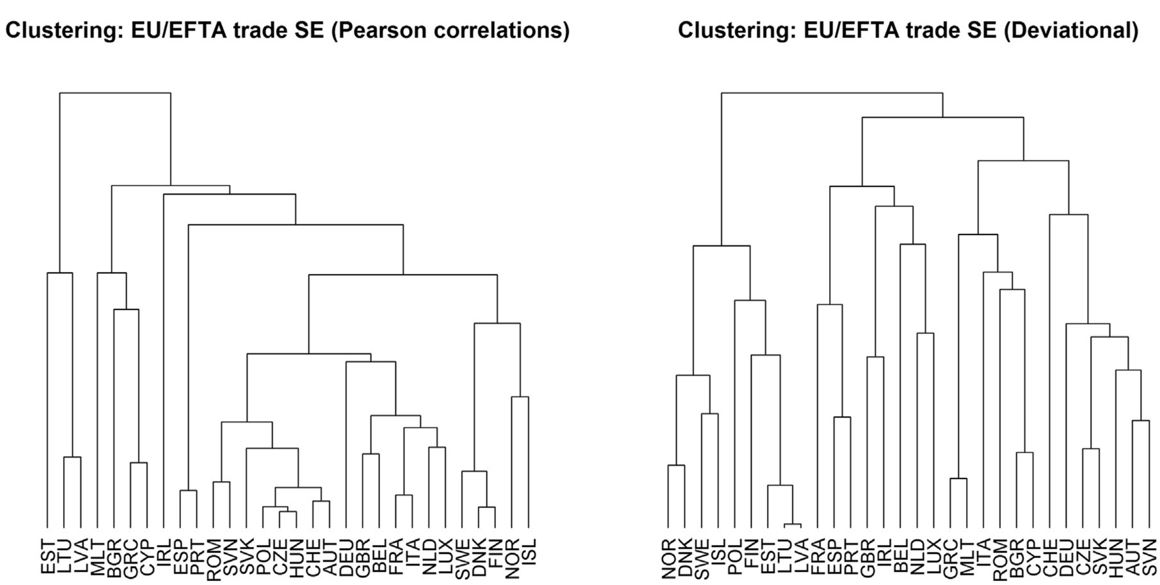

Fig. 12. Hierarchical clustering (weighted-average) dendrograms of conventional (Pearson correlations) and deviational structural equivalence of EU/EFTA 2010 trade data.

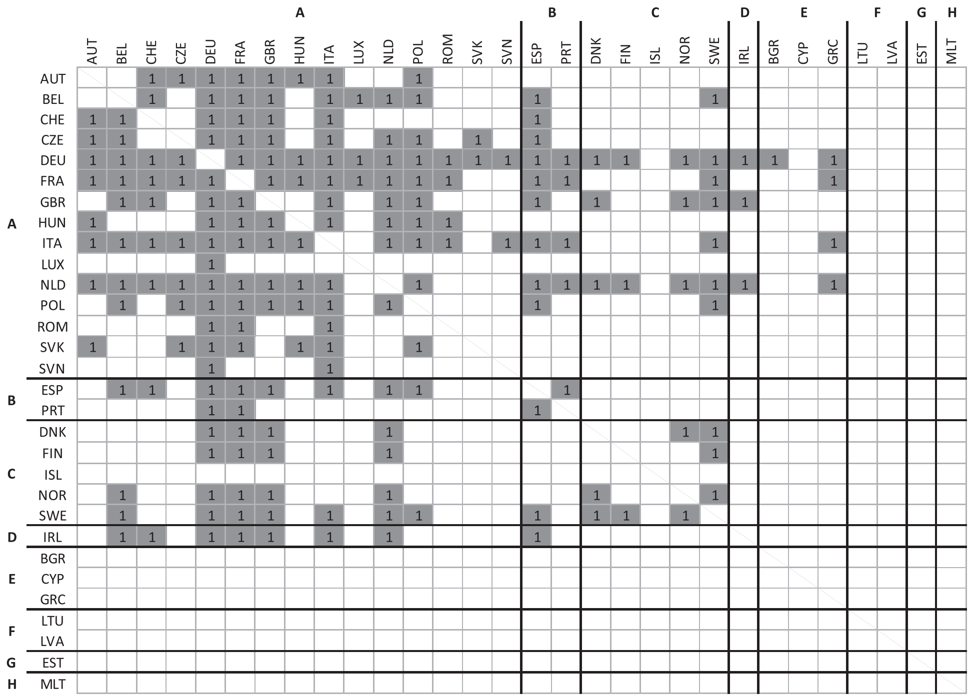

Fig. 13. Conventional (indirect) structural blockmodel of EU/EFTA dataset: 8-positional partition (tie values above 2848 million USD are deemed prominent).

Applying formulas (7)–(9) to calculate pair-wise correlations of rows in *RD* and columns in *CD*, a matrix of deviational similarities for the EU/EFTA trade data is created. Applying weighted-average hierarchical clustering on these two matrices of conventional and deviational measures of structural equivalence results in the two dendrograms given in Fig. 12.

Choosing cutoffs that yield 8 positions in respective dendrogram, the conventional cluster analysis results in three singleton positions characterized by their small trade volumes – Ireland, Malta, and Estonia – whereas the deviational cluster analysis results in one singleton position characterized by its unique role in the European economic landscape: Switzerland. The conventional approach yields one position containing half of the countries, whereas the corresponding partition of the deviational dendrogram results in positions that are more equal in size. Based on these two partitions, applying the same parameters that identified 200 ties as prominent for the pre-determined partition above, we arrive at the conventional and deviational structural blockmodels in, respectively, Figs. 13 and 14.

In the conventional correlation-based indirect approach to structural blockmodeling of EU/EFTA trade, a very "forgiving" blockmodel interpretation would be that of the typical

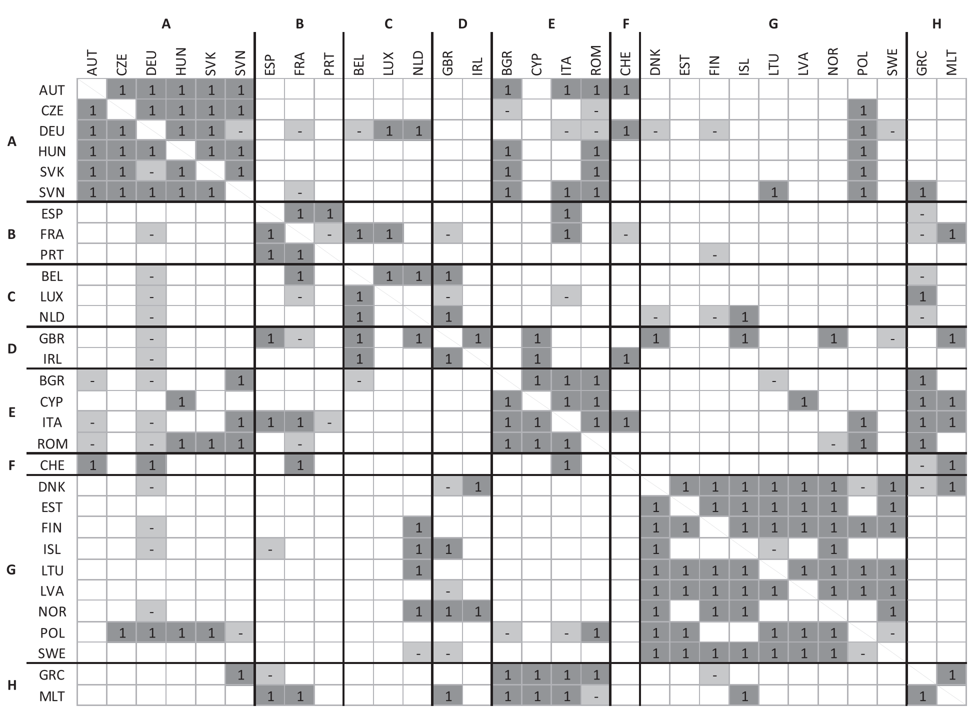

Fig. 14. Deviational (indirect) structural blockmodel of EU/EFTA dataset: 8-positional partition (two-sided deviational threshold = 0.083; strong 1/0-cell criteria).

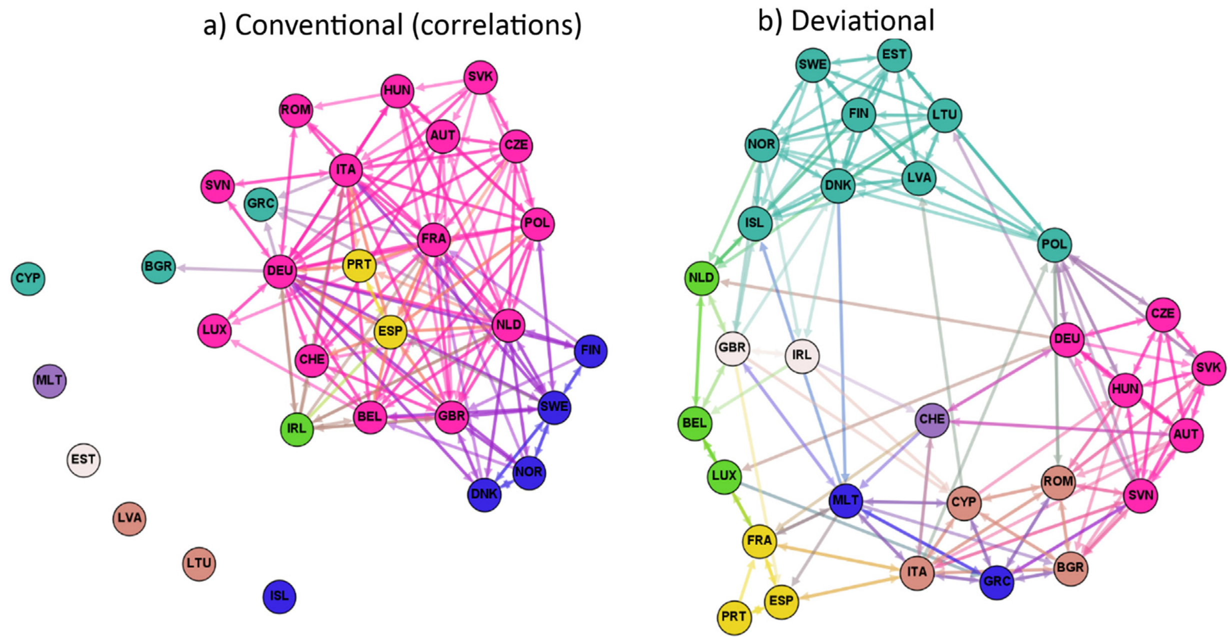

Fig. 15. (a and b) Force-directed layout of, respectively, conventional correlation-based (a) and deviational (b) blockmodels, node colors indicate positional membership.

'core–periphery' or 'centralized' image (Wasserman and Faust, 1994, p. 423). Contrasting this, the deviational structural blockmodel resemblances the typical 'cohesive subgroups' image/template (Wasserman and Faust, 1994), i.e. with complete blocks in the diagonal, here being in support of the existence of regional trade effects in the European economy. In addition to this, the inter-positional ties classified as prominent in the deviational approach are either between neighboring countries, or can be explained, similar to the Cyprus–Latvia link, by the particularities surrounding each such dyad.

Using a force-directed layout algorithm, indicating positional membership by different colors, the graphs in Fig. 15 corresponds to the conventional (a) and the deviational (b) blockmodels in Figs. 13 and 14. Noteworthy, the obtained layout of the deviational structural blockmodel broadly overlaps with the European geography, an aspect that is not as evident in the conventional structural blockmodel graph.

Although force-directed graphs are less than ideal for interpreting blockmodels, Fig. 15b does hint at possible applications of the suggested deviational approach to valued networks outside the scope of role-analysis and blockmodeling. The deviational approach per se – i.e. the calculation of *RD* and *CD*, their subsequent (parametric or non-parametric) transformation to *RB* and *CB*, and the (weak or strong) interpretation of the binary *RB* and *CB* matrices into a binary directional network, similar to the one in Fig. 15b – is indeed independent of would-be applications in blockmodeling and role-analysis. As discussed below, this could potentially provide a bridge between valued networks and existing binary metrics, as such providing a new family of deviational micro-, meso- and macro-level metrics for valued networks.

7. Conclusion

Addressing the inherent dilemmas of valued blockmodeling and the shortfalls of existing approaches for comparing ideal binary blocks with valued empirical data, this article has proposed a deviational approach to valued blockmodeling. Rather than determining empirical block patterns and the prominence of valued ties on the basis of absolute values, the proposed heuristic conceptualizes tie prominence in terms of deviations from expected values, the latter determined from the distributions of valued in- and outdegrees. By default non-parametric, the heuristic utilizes the full range and data resolution of valued relational data, and is applicable for networks irrespectively of their valued degree distributions. Operationalizing an alternative conceptualization of tie prominence in valued networks, the proposed heuristic allows for direct blockmodeling of valued networks, mapping the overlap between empirical blocks and the standard set of ideal blocks on a per-tie basis.

This article also proposes an alternative indirect way to establish deviational structural blockmodels. By correlating the row and column deviations from expected values for each pair of actors, the resulting coefficients of deviational structural equivalence can, similar to the conventional Pearson correlation measure, be analyzed to determine suitable partitions. The interpretation of the resulting deviational structural blockmodel is facilitated by the same heuristic used in the direct approach, allowing for a direct comparison between empirical valued blocks and ideal (binary) complete and null blocks.

Both the direct and indirect approach are tested on three datasets: Hlebec's note-sharing data, Baker's journal citation data, and commodity trade among European (EU/EFTA) countries as of 2010. For the note-sharing dataset, an almost identical partition is found when pre-specifying the "best-fit" block image found by Žiberna. However, rather than introducing novel types of ideal blocks and providing an arbitrarily set parameter value, the deviational approach not only provides a goodness-of-fit measure but also yields a measure of interpretational certainty for each hypothetical blockmodel. Without a pre-specified block image, slightly different optimal partitions and block images are found, also depending on the applied criteria function for regular blocks.

The Baker citation example confirms the core–periphery structure identified in previous studies. However, rather than consistently categorizing high gross-degree journals as core, the deviational approach diverges somewhat from previously identified core–periphery partitions for this dataset. In addition, although the direct blockmodeling fitting function ignore the patterns of ties *between* core and periphery, the deviational approach allows for identifying prominent ties across the blockmodel, finding citation patterns between core and periphery that follow the ideal row- and column-regular blocks. Comparing the indirect measure of deviational structural equivalence with the more conventional measures of, respectively, Euclidean distances and Pearson cross-product correlations, the deviational approach seems better at capturing thematic journal clusters. Applying a two-sided deviation threshold where contradictory ties disappear, the resulting prominent and non-prominent ties are used to construct the resulting density blockimage, where the densities of blocks containing non-determined ties are stated in terms of value ranges.

With the example data on intra-European commodity trade of 2010, the differences between the conventional and deviational indirect measures of structural equivalence, and the subsequent interpretations respective approach allows for, is emphasized further. Whereas the conventional indirect correlation-based measure hints at something resembling a European core–periphery trade structure, the deviational approach lends strong support to intra-European regionalism. In its categorization of European countries into structurally equivalent positions, it is evident that the deviational approach performs better at capturing prominent bilateral trade on the basis of total imports and exports of pairs of countries than what is the case for the conventional approaches.

Operationalizing a different approach to the notion of patterns of prominent ties in valued networks, yielding direct and indirect blockmodel results that evidently differ from those obtained by existing approaches, the deviational approach has to be tested and evaluated using other valued datasets, comparing how obtained results differ from more conventional approaches and, when applicable, to compare resulting partitions and tie classifications with would-be pre-understandings of roles and structures in the analyzed networks. However, related to this, it is crucial to understand that the herein suggested deviational approach is not a replacement to existing approaches to direct and indirect blockmodeling, but an alternative approach taken from a somewhat different viewpoint. With a fundamentally different conceptualization of what constitutes prominent and non-prominent ties in valued networks implies an equally different conceptualization of equivalence in networks. For instance, whereas both the Pearson correlation and the deviational correlation capture notions of structural equivalence, their respective operationalization of what equivalence translates to might indeed be meaningful for studying a particular valued network, but their respective conceptualization of equivalence do not rule each other out. Thus, depending on the research question and how we choose to conceptualize equivalence in the valued networks of commodity trade, the European economy might be perceived as *both* a core–periphery block image *and* a cohesive subgroup block image.

Whereas this article has focused exclusively on blockmodeling of valued (one-mode) networks, a possible future research agenda is to explore whether the deviational approach can be equally applied to other methods and algorithms primarily designed for binary networks. One such extension could be to modify the REGE algorithm to use *RD* and *CD* when measuring dyadic regular equivalence, i.e. similar to the alternative REGE implementation suggested by Žiberna (2007b, p. 35). The calculation of *RD* and *CD* and the subsequent transformation, non-parametrically or using a one- or two-sided deviation threshold, to the binary matrices *RB* and *CB* could possibly also be applied to existing binary-oriented metrics and heuristics, e.g. centrality indices, cluster coefficients, cohesive subgroup analysis, triadic analyses etc., to the study of patterns of prominent ties in valued networks.

Acknowledgements

This research was fully sponsored by Budapesti Közép-Európai Egyetem Alapítvány (CEU BPF). The views expressed in this article are those of the author and do not necessarily reflect the views of CEU BPF. Thanks to my colleagues and friends at the Center for Network Science and the Department of Political Science, John Harboard at the Center for Academic Writing, Central European University, and fellows and friends at RECENS, MTA, Hungarian Academy of Science. I am particularly grateful for the two anonymous reviewers who gave feedback on a previous version of this manuscript.

Appendix A.

Baker original noselfties.

	SW	SCW	SSR	CW	JSWE	SWRA	ASW	CYSR	CSWJ	SWHC	SWG	PW	CAN	FR	JGSW	BJSW	CCQ	JSP	AMH	IJSW
SW	0	58	53	52	33	8	15	0	0	43	15	19	0	9	0	0	0	0	0	0
SCW	124	0	36	17	21	18	8	0	8	6	8	0	6	6	6	0	0	0	0	0
SSR	106	30	0	17	9	25	7	0	0	0	0	0	0	0	0	0	0	0	0	0
CW	58	32	10	0	11	0	0	6	0	0	0	7	7	0	0	0	0	0	0	0
JSWE	58	18	16	0	0	16	9	0	0	7	0	0	0	0	0	0	0	0	0	0
SWRA	44	8	39	8	24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ASW	73	0	21	0	18	7	0	0	0	0	0	13	0	0	0	0	0	0	0	0
CYSR	28	8	14	70	0	5	0	0	0	0	0	6	12	4	0	0	5	0	0	0
CSWJ	45	47	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SWHC	26	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SWG	40	9	7	0	9	0	0	0	0	9	0	0	0	0	0	2	0	0	0	0
PW	0	0	0	4	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0
CAN	8	6	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FR	9	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
JGSW	18	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BJSW	19	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CCQ	0	3	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
JSP	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AMH	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IJSW	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

EFTA.

	DEU	FRA	NLD	GBR	ITA	BEL	ESP	CHE	AUT
DEU	0	103434	78104	71401	78067	62839	37030	56286	59429
FRA	81804	0	19137	36177	42501	43135	33872	14993	4286
NLD	91099	25188	0	40634	26426	72764	14179	7988	4281
GBR	51108	25974	29337	0	13231	21871	14298	6843	2290
ITA	57851	44962	9496	22037	0	11812	22127	17969	10190
BEL	44632	46993	42243	26347	17683	0	7933	4927	2427
ESP	29486	37185	9330	15494	22153	8151	0	4724	2348
CHE	43552	14600	2809	9185	13540	3855	3576	0	8138
AUT	45234	5876	2334	4056	11197	2442	2460	7625	0
POL	37648	9124	6130	9388	9570	3765	4034	1239	2512
SWE	17526	7474	7002	10082	4590	7325	3064	1366	1653
CZE	39246	6640	6024	6141	5938	3110	2929	2049	5547
NOR	22663	6491	11177	29730	2073	4623	2412	276	769
IRL	18591	7745	5295	19745	4153	19810	4341	5629	805
HUN	22126	4233	2855	5016	4763	1459	2388	891	4150
DNK	14700	3524	3612	6312	2808	1484	2222	937	601
PRT	5504	5579	1909	2667	1837	1439	11319	484	529
SVK	12341	3841	1652	2492	3383	955	2105	491	3485
FIN	7979	2923	4211	3339	1921	2206	1410	812	588
ROM	8860	3973	1249	1907	6185	720	1382	295	1248
GRC	2592	811	523	1040	2490	336	713	214	182
SVN	5035	1829	382	547	2868	208	291	247	1664
LUX	3945	1983	859	1441	1399	2761	594	248	246
BGR	2302	857	259	353	2144	1002	474	102	427
LTU	1955	790	1102	851	372	323	254	51	83
EST	583	191	325	237	114	149	52	85	48
LVA	802	173	228	588	125	84	107	31	31
CYP	348	46	146	160	222	86	16	5	47
MLT	404	481	116	257	332	104	139	11	13
ISL	940	150	834	546	53	89	169	64	21

	POL	SWE	CZE	NOR	IRL	HUN	DNK	PRT	SVK	FIN
DEU	37784	26758	32057	9507	4627	20990	17116	10460	10352	9040
FRA	7504	6814	4151	2763	2442	3238	2685	5466	2337	2671
NLD	6400	9196	4048	2888	2938	3950	5856	3866	662	3682
GBR	4676	8336	2544	4580	19488	1641	4999	2847	1134	2114
ITA	9726	4366	4897	2096	1032	3739	2816	4292	2174	1917
BEL	4104	5709	2276	1418	1436	1927	2736	2150	679	1620
ESP	3490	1818	2253	1015	882	1097	1183	23557	647	897
CHE	1499	1247	1353	802	1123	663	852	480	392	774
AUT	2914	1603	4231	614	219	5414	775	383	1630	625
POL	0	4342	8041	1941	425	4590	2463	466	2632	1200
SWE	3263	0	1168	10807	452	794	11009	768	344	6907
CZE	6442	1789	0	767	255	2832	951	470	6643	802
NOR	2671	12971	1037	0	1712	29	3306	700	62	1573
IRL	1092	1943	1052	798	0	457	974	725	184	509
HUN	3057	991	2734	348	268	0	588	353	2758	335
DNK	2115	12309	772	4801	1032	649	0	416	272	1635
PRT	443	510	401	257	131	148	367	0	118	402
SVK	3599	1122	6505	393	69	3624	339	135	0	278
FIN	1765	7814	450	1973	252	434	1370	163	162	0
ROM	1225	307	708	477	141	2281	125	159	462	134
GRC	304	197	175	43	39	107	184	142	122	182
SVN	683	199	531	97	17	856	235	46	247	97
LUX	242	355	215	84	36	105	139	66	59	78
BGR	333	90	176	46	14	202	69	35	99	46
LTU	1022	787	119	529	56	67	527	35	24	300
EST	166	1682	59	423	4	17	238	12	18	1664
LVA	220	578	40	257	33	20	310	5	21	284
CYP	34	27	39	6	4	77	31	2	9	7
MLT	45	14	15	9	15	21	13	30	4	21
ISL	93	34	58	208	13	9	101	18	14	14

	ROM	GRC	SVN	LUX	BGR	LTU	EST	LVA	CYP	MLT	ISL
DEU	10361	6709	4273	5220	2958	2553	1398	1282	774	391	294
FRA	3674	3128	1432	3035	833	602	273	270	435	448	72
NLD	2181	3368	517	842	711	1028	311	441	385	228	333
GBR	1420	1910	389	407	400	371	289	177	705	474	200
ITA	7171	6159	4156	552	1876	764	372	371	802	1411	116
BEL	1313	2167	385	5069	482	756	187	193	177	91	53
ESP	1303	1924	705	239	479	281	123	132	238	144	45
CHE	584	1043	515	152	235	61	143	162	104	143	54
AUT	2539	706	2091	127	882	172	94	129	39	24	15
POL	2318	482	574	177	533	2065	718	878	28	26	48

EFTA. (*Continued*)

	ROM	GRC	SVN	LUX	BGR	LTU	EST	LVA	CYP	MLT	ISL
SWE	352	472	178	107	133	763	1121	396	60	20	204
CZE	1475	311	633	139	482	332	141	163	30	12	22
NOR	117	263	39	35	11	79	238	117	9	2	354
IRL	323	551	95	84	71	42	57	27	234	39	32
HUN	5383	301	757	115	791	187	92	148	31	19	13
DNK	270	644	78	53	127	398	181	255	36	111	276
PRT	251	146	46	75	74	18	16	12	45	9	8
SVK	959	169	371	77	276	139	40	80	19	4	6
FIN	233	345	90	25	126	414	1521	547	30	27	207
ROM	0	607	252	17	1769	42	24	8	91	18	5
GRC	827	0	105	8	1511	17	9	10	1616	77	3
SVN	413	185	0	22	200	75	25	35	9	4	3
LUX	63	208	72	0	17	13	7	16	6	1	1
BGR	1900	1375	159	4	0	51	7	19	42	4	2
LTU	46	33	9	1	38	0	790	1900	2	1	18
EST	14	6	6	2	5	668	0	797	3	1	33
LVA	31	9	3	1	15	1461	705	0	2	1	16
CYP	83	719	1	5	37	4	2	22	0	8	0
MLT	29	38	3	2	34	1	1	1	134	0	7
ISL	4	10	12	2	1	13	3	4	0	1	0

Notesharing.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0	0	0	15	0	0	0	1	8	0	0	3	0
2	0	0	2	3	0	0	5	5	10	10	1	3	0
3	0	0	0	19	0	0	0	3	1	0	0	0	0
4	2	0	6	0	1	0	0	1	19	0	1	0	0
5	0	0	0	16	0	5	0	7	16	0	5	0	3
6	0	0	1	0	4	0	0	7	3	0	7	3	1
7	0	0	6	14	0	0	0	14	6	0	0	0	0
8	0	0	0	5	0	0	0	0	6	0	0	0	0
9	0	0	0	19	0	0	0	1	0	0	0	0	0
10	0	16	2	16	0	1	0	16	0	0	1	2	0
11	0	0	2	8	2	2	0	5	14	0	0	2	0
12	2	2	8	2	2	2	2	2	6	2	11	0	0
13	0	0	0	1	8	0	0	8	3	0	0	0	0

References

Baker, D.R., 1992. A structural analysis of the social work journal network: 1985–1986. J. Soc. Serv. Res. 15 (3-4), 153–168.

Borgatti, S.P., Everett, M.G., 1992. Regular blockmodels of multiway, multimode matrices. Soc. Netw. 14, 91–120.

Borgatti, S.P., Everett, M.G., 1999. Models of core/periphery structures. Soc. Netw. 21 (4), 375–395.

Breiger, R.L., Boorman, S., Arabie, P., 1975. An algorithm for clustering relational data with applications to social network analysis. J. Math. Psychol. 12, 329–383.

Brovkin, V., 2001. Moving money, making money, and parking money overseas: front companies in offshore jurisdictions. Demokratizatsiya 9 (2), 150–166.

Consulco, 2011. Cyprus Tax Treaty List and Treaty Withholding Rates. Consulco House, Nicosia, Cyprus.

De Blij, H.J., Muller, P.O., 2004. Geography Realms, Regions, and Concepts. Wiley.

Doreian, P., Batagelj, V., Ferligoj, A., 2005. Generalized Blockmodeling. Cambridge University Press, Cambridge, UK.

Doreian, P., 2006. Some Open Problems for Generalized Blockmodeling. In: Batagelj, V., Bock, H., Ferligoj, A., Žiberna, A. (Eds.), Data Science and Classification. Springer verlag, Berlin, pp. 119–130.

Eglitis, A., 2013. Latvia bank deposits to grow as cash leaves Cyprus, Moody's says. Bloomberg, Downloaded http://www.bloomberg.com/news/2013-05-02/latvia-bank-deposits-to-grow-as-cash-leaves-cyprus-moody-s-says.html (03.05.13).

Ferligoj, A., Doreian, P., Batagelj, V., 2011. Positions and roles. In: Carrington, P., Scott, J. (Eds.), Handbook of Social Network Analysis. SAGE Publications, Newbury Park, pp. 434–446.

Galtung, J., 1966. East–west interaction patterns. J. Peace Res. 3 (2), 146–177.

Hartigan, J.A., 1972. Direct clustering of a data matrix. J. Am. Stat. Assoc. 67 (337), 123–129.

Hlebec, V., 1996. Metodoloske znacilnosti anketnega zbiranja podatkov v analizi omrezji: Magistersko delo. FDV, Ljubljana.

Lorrain, F., White, H.C., 1971. Structural equivalence of individuals in social networks. J. Math. Sociol. 1, 49–80.

Nordlund, C., 2007. Identifying regular blocks in valued networks: a heuristic applied to the St. Marks carbon flow data, and international trade in cereal products. Soc. Netw. 29 (1), 59–69.

Sprūds, A. (Ed.), 2012. The Economic Presence of Russia and Belarus in the Baltic States: Risks and Opportunities. Latvian Institute of International Affairs, Centre for East European Policy Studies, Soros Foundation-Latvia, Riga.

UNCTAD, n.d. United Nations Comtrade Trade Database. http://comtrade.un.org (accessed 01.02.14).

Wasserman, S., Faust, K., 1994. Social Network Analysis: Methods and Applications. Cambridge University Press.

White, H.C., 1974a. Models for interrelated roles from multiple networks in small populations. In: Knopp, P.J., Meyer, G.H. (Eds.), Proceedings of the Conference on the Application of Undergraduate Mathematics in the Engineering, Life, Managerial and Social Sciences. Georgia Institute of Technology, Atlanta, pp. 52–102.

White, H.C., 1974b. Null Probabilities for Blockmodels of Social Networks. Department of Sociology, Harvard University, Unpublished paper.

White, H.C., Boorman, S.A., Breiger, R.L., 1976. Social structure from multiple networks. I. Blockmodels of roles and positions. Am. J. Sociol. 81 (4), 730–780.

Žiberna, A., 2007a. Generalized blockmodeling of valued networks. Soc. Netw. 29, 105–126.

Žiberna, A., PhD thesis 2007b. Generalized Blockmodeling of Valued Networks (POSPLOŠENO BLOČNO MODELIRANJE OMREŽIJ Z VREDNOSTMI NA POVEZAVAH). Ljubljana University, Faculty of Social Science, Ljubljana.