Discussion Paper, Series A, No. 2010-221

A Network Structure of ROSCAs

(Rotating Savings and Credit Associations):

ERGMs (Exponential Random Graph Models)
Applied to a Leaders' Network in Rural Uzbekistan

Masato Hiwatari

February, 2010

Graduate School of Economics & Business Administration
Hokkaido University
Kita 9 Nishi 7, Kita-Ku, Sapporo 060-0809, JAPAN

A Network Structure of ROSCAs (Rotating Savings and Credit Associations): ERGMs (Exponential Random Graph Models)Applied to a Leaders' Network in Rural Uzbekistan*

Masato Hiwatari†

February 19th, 2010

Abstract

This paper empirically analyzes a network structure created by the ROSCAs (Rotating Savings and Credit Associations) related to a leaders' network in rural Uzbekistan. The estimation methodology is based on the recent development of ERGMs (Exponential Random Graph Models) whose approximate maximum likelihood estimators are produced by MCMC (Markov Chain Monte Carlo) algorithms. The paper reveals the tendencies of the transitive triad structure of the network that can facilitate the tracking of defecting members.
JEL: O1, I3
Keywords: ROSCAs; Networks; Risk-Sharing; ERGMs; MCMC

I. Introduction

This paper empirically analyzes the network structure created by ROSCAs (Rotating Savings and Credit Associations), by focusing on the particular case of a

*I thank participants of the fifth joint seminar of Yeungnam University and Hokkaido University and the joint workshop of EUIJ Kansai and the institute of economic research, Kyoto University.
†Graduate School of Economics and Business Administration, Hokkaido University; *E-mail address*: mhiwatari@econ.hokudai.ac.jp

leaders' network in rural Uzbekistan. This study has two aims: the first concerns a methodological issue involving an empirical analysis of network formation in rural villages in developing countries and the other is about the application of the network perspective to ROSCAs.

In recent years, development economists have increasingly started taking cognizance of the role of sub-groups within the local communities, or more specifically, the ”structure” or the ”architecture” of the networks among local villagers. Previously, development economists had largely taken the community or the network to be an exogenous entity; i.e., they had either interpreted village as exogenous units of study, or while examining their effects on some outcomes, had considered some exogenous characteristics – such as ethnicity, caste, or other indices of village level averages – as proxies of the community or the social network (e.g., Townsend [1994]; Deaton [1997], chap 6). More recently, however, they have begun stressing the importance of the state and structure of intra-village network-based components. This is particularly true of the literature on the risk-sharing networks of rural households in developing countries (e.g., Genicot and Ray [2003]; De Weerdt and Dercon [2006]; Fafchamps and Lund [2003]; Fafchamps and Gubert [2007]). It is now recognized that the risks faced by rural households – such as unpredictable income fluctuation or expenditure needs – are, in reality, not shared by exogenous units such as villages, but by small, tightly-knit sub-groups or the networks that are endogenously formed in the process of coping with information and enforcement issues. Some attempts have already been made to empirically explore the formation of these sub-groups or networks. For instance, through their dyadic regression model, Fafchamps and Gubert [2007] examined the determinants of the formation of the intra-village mutual insurance links in rural Philippines and confirmed the effects of social and geographical proximity.

The past decade was also marked by the embarking game-theoretic literature on endogenous network formation. The main focus of the literature has been on characterizing the architecture of the equilibrium network, which has lead to predictions about the conditions under which specific networks are stable (e.g., Jackson and Wolinsky [1996]; Bala and Goyal [2000]; Jackson [2008]; Bloch et al. [2008]). Most recent empirical studies on developing countries are on the verge of examining the empirical implication of the endogenous network theory (Comola [2008]; Krishnan and Sciubba [2009]). For instance, by using the empirical data of rural Ethiopia, Krishnan and Sciubba [2009] confirmed the role of the architecture of the network of labor-sharing arrangements by examining the network properties that are necessarily shared by the possible equilibriums of their model. However, this perspective has been subjected to little empirical challenge. As far as the methodological technique for carrying out empirical analyses is concerned, there might still be plenty of scope for effecting further improvements in it or finding substitutable approaches.

Thus, the first aim of this paper is to make a methodological contribution to the empirical studies on informal network formation in rural villages. This paper, rather than testing a specific equilibrium prediction derived from the endogenous network formation theory, attempts to identify the determinants of network formation by modeling a probability distribution of the entire network of a local community, where the probability of observing a network is dependent on the presence of various *configuration*s (subgraphs) expressed by the models. In particular, I introduce the Markov random graph models – a particular sub-class of exponential random graph models (ERGMs) – that enable us to deal with the effects of a specific dependence structure between networks links. The advantage of this modeling framework over the other approaches such as the dyadic regression approach lies in its capacity to represent the effects of complex local network structures involving more than three actors as well as nodal level, dyadic level, or general degree-based effects. They are parameterized in terms of the prevalence of small *configuration*s in the network and can be interpreted as describing the combinations of local social process from which a given network emerges. This approach is preferable to other approaches, such as testing the summary statistics of specific network properties, in that we can avoid a loss of information as regards the dependent structure of links and examine the effects of more local and specific architectures within the network.

The estimation methodology of this paper is based on the recent development of ERGMs whose approximate maximum likelihood estimators are produced by Markov Chain Monte Carlo (MCMC) algorithms. Statistical approaches to social networks span a long history, being prevalent since the 1930s. The class of exponential random graphs was first proposed in the statistical literature in the 1980s (Frank and Strauss [1986]) on the basis of existing works in spatial statistic literature (e.g., Besag [1974]). The models were derived from an explicit hypothesis about dependencies among network links. More sophisticated extensions of the models were proposed in the 1990s (Strauss and Ikeda [1990]; Wasserman and Pattison [1996]) and used widely. The exponential random graph models (ERGMs), also known as $p*$ models, are now considered to be the most promising class of statistical models for expressing a social network observed at a particular moment in the network literature. In recent years, the models have experienced remarkable improvement in terms of both model parameterizations and computer algorithms (Snijders et al. [2006]; Robins et al. [2007]), which are the factors on which the estimation methodology of this paper is mainly based. The algorithms will be explained in details in the Appendixes.

The second aim of this paper is to make the pioneering attempt to explore the network structure of ROSCAs. ROSCAs are informal financial associations widely reported in the developing world (e.g., Geertz [1962]; Ardener [1964]).

They are formed by members who agree to make regular contributions to a fund and the total sum thus raised is then given to each contributor in rotation. In economics literature, the ROSCA has been understood as a system that, as compared to an autarky, can facilitate earlier gain of utility from the use of indivisible goods (Besley et al. [1993]; Kovsted and Lyk-Jensen [1999]). This function of the ROSCA has been tested in certain empirical studies of development economics (e.g., Besley and Levenson [1996]; Handa and Kirton [1999]). Obviously, the ROSCA system can be also utilized for any relatively large, one time expenditure such as small business investments (e.g., Kimuyu [1999]; Johnson [2004]). Moreover, the ROSCA fund can be utilized for risk-sharing or insurance purposes among the association's members, thus facilitating urgent access to money, provided the members allow for some flexibility in the sequence of fund's rotation among themselves or in the amount of the contribution at each meeting (Calomiris and Rajaraman [1998]). As regards the ROSCAs as a savings system, a recent theoretical study has featured certain psychological effects emanating from participation in a ROSCA, such as increased self-control and restraint as regards minor expenses (Ambec and Treich [2007]; Gugerty [2007]).

While a growing amount of economics literature views ROSCAs in the manner described above, the network structure of ROSCAs has not attracted much research' interest thus far. One reason for this is that ROSCAs have been simply classified as groups or affiliation networks; they are considered to have no variations in their architecture, and the relationship between their members is interpreted as exclusively symmetric (e.g., Krishnan and Sciubba [2009], 918). However, I would like to emphasize that this viewpoint is not necessarily justified. In actuality, it is reported that participation in multiple ROSCAs can be regarded as a rather common phenomena in transition countries such as Uzbekistan (Hiwatari [2008b]), and that the numerous layers of interaction resulting from this type of participation have given rise to complex network structures that extend over the entire village. Thus, this paper – which deals with empirical data sourced from rural Uzbekistan – examines whether actors choose their ROSCA partners solely on the basis of personal status, attributes or affiliation or, whether the potential network structure resulting from their additional membership is also an underlying consideration. Specifically, I focus on a transitive triad structure of the network that provides actors with dyadic constraints and thus facilitates monitoring and enforcement. Owing to the lack of relevant data, this paper presents a rather preliminary analysis of a certain village's leaders' network, considering only a specific network comprising 45 nodes. I intend to conduct a more comprehensive analysis comprising the relevant data of all the households in a village in the immediate future.

The rest of the paper is organized as follows. Section II describes the data as well as provides background information. Section III and IV discuss the model specification and the estimation strategy. Section V reveals the results. Some technical issues are mentioned in the Appendixes.

II. Data and Setting

Specifically speaking, this paper deals with the network data of certain traditional associations – called *gap*s – in rural Uzbekistan. *Gap*s and similar associations, in which the members take turns to invite other members of the group to their houses for feasts or entertainment, have been deeply pervasive in Uzbekistan since the Soviet period and have been observed even earlier. Such associations have been mentioned in the local literature as one of the essential components of traditional communal life in Uzbekistan. Their function as ROSCAs, in fact, has attracted attention only recently (Kandiyoti [1998], Hiwatari [2008b]). Hiwatari [2008b] examined the quantitative magnitude of the associations' prevalence and their function as ROSCAs on the basis of field research conducted in the Andijan Region of Uzbekistan. In the researched *mahalla*[1](neighborhood community), henceforth referred as *Oqmahall*, 81.6% of the sample households had at least one family member who belonged to a *gap* and had participated in 2.47 *gap*s on an average. These numbers clearly confirmed that most of the households in *Oqmahalla* were intertwined in a dense web of *gap* networks. Households whose members belonged to *gap*s were expected to receive 70% of their monthly household incomes 2.5 times a year from their *gap* funds. The positive correlation between *gap* participation and the prevalence of durable goods was also recognized. These results pointed to the role of recent *gap*s, such as ROSCAs, as well as their impacts on residents' economic life.

Furthermore, Hiwatari [2008b] discussed some peculiar characteristic of *gap*s by considering the issues of sustainability constraints and default problems, and cited two distinctive features: the high propensity to multiple belongings and the scarcity of default problems. First, residents have often stated that there is no restriction to belong to multiple *gap*s, and a considerable number of people did participate in multiple *gap*s. This propensity was particularly high for local magnates or leaders as will be discussed below. Second, we obtained a notable result that any experience of a serious default problem was not reported throughout the survey. This fact is quite uncommon with respect to the situation in other coun-

[1]*Mahalla*s are neighborhood communities that have been reformed into formal institutions or lowest administrative units of local municipalities (*khokimiyats*) during the 1990s and till present. They have existed since ancient times as traditional institutions in areas such as Bukhara, Samarkand, Tashkent, and the Ferghana Valley. In some urban areas of Uzbekistan, *mahalla*s do not have deep historical roots and have been recently created as lowest administrative units by the Uzbek government (See, for example, Sievers [2002]; Hiwatari [2008a]).

tries where defaults in ROSCAs often cause major problems such as suicides or demands for governmental regulations (e.g. Ardener [1964], 216; Besley and Levenson [1996], 43). In *Oqmahalla*, most of the respondents agree with the view that the inability of a member to pay for his or her obligation in a gathering is unlikely to lead to a major problem. *Gap*s are never disbanded and the defaulter is not banished from the group on the grounds of his or her default.

These two features of *gap*s imply some discrepancy in the realities with respect to the theories; that is, in general, increasing the number of groups to which the members belong as well as increasing the group size must raise the risks of defaulting of the former group not only because of the enlarging obligations of the members, but also as a result of the increasing monitoring and enforcement costs and the ambiguity of the boundaries of membership. However, in *gap*s, few problems of defaults or other destructive conflicts were observed, though members have a high propensity to belong to multiple *gap*s. To understand this issue, Hiwatari [2008b] revealed a necessity to not consider theoretical frameworks that deal with each interest group such as a ROSCA as a separate unit, and to apply a network-related perspective, but only visually alluded the importance of an overlapping complex structure of the *gap* network. In light of this, this paper introduces statistical modeling for the analysis of the network structure and attempt to clarify the structural propensities of the *gap* network.

In order to examine a network structure of *gap*s, I use a list of members who hold any committee post in the *Oqmahalla mahalla* committee. The *mahalla* committee is a representative organ of the *mahalla* formed by its residents. In *Oqmahalla*, it comprises a council (*kengash*), advisors (*maslahatchi*), and ten subordinate commissions that work on specific tasks. These members are, in principle, elected by the residents every 2.5 years. Four of the members at the main posts receive remuneration for their work and their elections need to be approved by the local municipality (*khokimiyat*). In reality, however, most of the other members are informally elected by the residents and the committee posts are only circulated among the members of the same lineup in successive terms. Since the *mahalla* faces a perennial shortage of qualified people, many members are forced to hold two or three posts concurrently as the *Oqmahalla* committee comprises as many as 63 posts (repeated names are omitted in the table). In other words, it can be said that the committee's members includes the names of almost all the influential persons or leaders in *Oqmahalla*. Thus, the analysis in this section aims to clarify the structural tendencies of the leaders' network in the *mahalla*.

Table 1 presents information regarding the *gap* participation of 45 *mahalla* committee members. It classifies the *gap* participation of these committee members by classmate *gap*s, *mahalla gap*s, relative *gap*s, and other *gap*s. The symbol ● indicates that the concerned person is a *gap* leader (*djurabashi*), ○ indicates that he or she is a normal *gap* member, and – indicates that he or she is not a mem-

Table 1: *Gap* participation of the *mahalla* committee members in *Oqmahalla*

No.	Age	Sex	Position	Classmate	Mahalla	Relative	Other
Council (*kengash*)							
1	47	M	chairman (*rais*)	○1	●	●1	−
2	57	F	deputy	○	●	●	−
3	46	M	secretary	○2	●	●2	−
4	57	M	member	●	○	−	●
5	47	M	member	●1	●	●	○
6	58	F	member (*dasturhonchi*)	○	●3	●3	−
7	53	F	member (*maslahatchi*)	○	●4	○	−
8	78	M	member（*oqsoqol*）	−	●	●	−
9	63	M	member (*oqsoqol, ex-imom*)	−	−	●	−
Advisors (*maslahatchi*)							
10	60	M	member	○	●1、●2	−	−
11	32	M	member (*shirkat rais*)	○3	○	−	●
12	63	F	member	−	●	●	−
13	54	M	member (*imom*)	●	○1	○1	○
14	67	M	member	−	○	−	●
15	64	F	member	−	●	●	○
16	51	M	member	●4	○1, ○2	○2	−
Members of subordinate commissions							
17	46	M	Member（*ex-rais*）	●2	○1, ●	●4	−
18	52	F	member	○	○3	●	−
19	46	M	commission head	○2	●	−	●
20	36	M	member	○5	−	−	●
21	48	M	member	○1	●	●5	●
22	48	F	member	●	○4	−	−
23	35	M	member	●6	−	−	●
24	32	M	member (*posbon*)	○3	○	○4	●
25	36	M	commission head	●5	○	−	−
26	40	M	member	○	−	−	○
27	54	M	member		○1, ○2	−	−
28	42	M	member	●	●3	−	○
29	47	M	commission head	○1	○	○6	○
30	38	F	member	○	●	○	−
31	49	M	member	○4	○	●6	−
32	44	M	member	●5	○	○2	●
33	47	F	member	○	−	●	−
34	25	M	member	●	○	−	−
35	37	M	member	−	○1, ○2	○1	−
36	51	F	member	−	○	○	−
37	35	M	commission head	○6	−	−	●
38	41	M	member	○7	−	○2	−
39	46	M	member	○2	●	○5	−
40	35	M	commission head	○6	●	○	−
41	51	M	member	○4	●	○	−
42	41	M	member	●7	●	−	○
43	49	F	commission head	−	○3, ○4	○3	−
44	45	F	member	●	●	○	−
45	48	F	member	○	−	●	○

Note: ● indicates belonging to a *gap* as a leader *(djurabashi*), ○, belonging as a normal member, and − indicates not belonging to a *gap*. The figures are noted on the right-hand side of the marks of the *gap* s to which more than one member in the list belong, in order to identify the common *gap* s.
Sources: Hiwatari (2008a, 188-189, Table6.1.3)

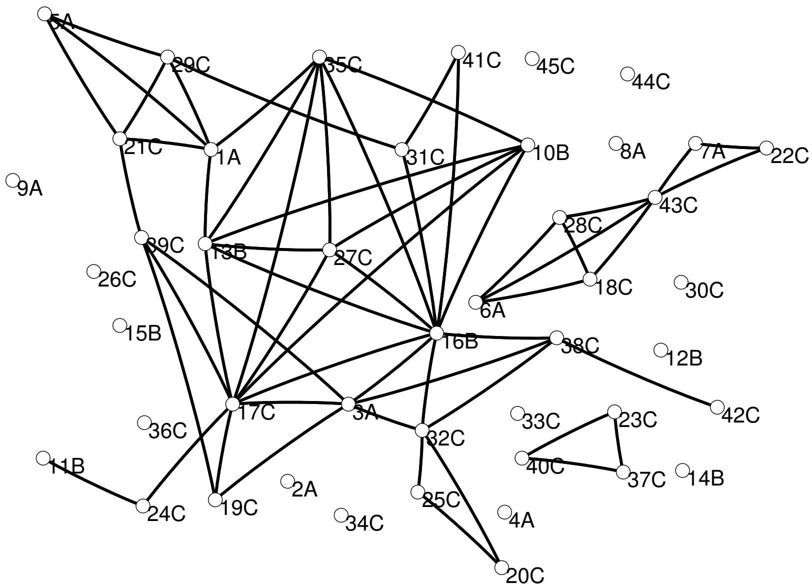
ber of any *gap*. In order to indicate the *gap*s whose membership the committee members share in common, the figures as substitutes of the unique names of each *gap* in each category are noted on the right-hand side of the marks of the *gap*s to which more than one member in the list belong. It is clear from the table that the committee members participate in more *gap*s than ordinary residents. All of them participate in at least one *gap*, while 29 out of the total 45 participate in more than three *gap*s. Moreover, these individuals tend to hold responsible positions in their *gap*s. In fact, 38 of the total 45 committee members hold a leadership position in at least one *gap*.

It should be noted that while several *gap*s contain multiple committee members, with the exception of one *gap*, the number is confined to only between two and four committee members in each *gap*. With regard to the members within the council (*kengash*) that forms the core of the committee, one pair merely belongs to the same *gap* of classmates. However, although each *gap* can unite only a few members within the committee, many committee members are expected to be indirectly united through their participation as members' of other *gap*s, and therefore, here, it is particularly meaningful to apply the perspective of a *network*, instead of a *group*. Figure 1 visually shows the relational structure of committee members resulting through their *gap* memberships. Here, the network data was converted into a one-mode network; the circles indicate persons and the lines, *gap*s. The numbers below the circle represent the member numbers and the alphabets, the committee posts: "A" denotes that the person is a council member (*kengash*); "B", that he or she is an advisor (*maslahatchi*); and "C", that the individual is a member of a subordinate commission. It is clear from Figure 1 that a large number of members are actually united through multiple *gap*s, and thereby, a large network is formed, as expected above. In light of this revelation, what kind of propensities can we derive from this network structure? This is the question that we attempt to answer in the following sections.

III. Model Specification

The network is represented by a symmetric $n \times n$ matrix $\mathbf{Y}$ and an $n \times q$ matrix $\mathbf{X}$ of nodal covariates, where n is the number of nodes (actors). The entries of the $\mathbf{Y}$ matrix, termed the *adjacency* matrix, are all zeros and ones, with $Y_{ij} = 1$, indicating the presence of a link (*tie*) between i and j. We specify y_{ij} as the observed value of Y_{ij} with $\mathbf{y}$ the matrix of observed links of the network. In a case of the *gap* network that is a non-directed network as shown in Figure 1, the *adjacency* matrix becomes symmetric matrix ($y_{ij} = y_{ji}$). The nodal covariate matrix $\mathbf{X}$ includes various attributes of the individuals in the network. As mentioned below, we use such covariates as positions in the committee, ages, gender, wage levels,

Figure 1: The Leaders' *Gap* Network in *Oqmahalla* (as a One-Mode Network)



and education levels, and consider these attributes to be fixed and exogenous.

The exponential random graph model of the adjacency matrix **Y**, conditional on the covariate matrix **X** takes the form of a probability distribution of graphs:

$$P_\theta\{\mathbf{Y} = \mathbf{y}|\mathbf{X}\} = exp(\boldsymbol{\theta}'\mathbf{u}(\mathbf{y}, \mathbf{X}) - \psi(\boldsymbol{\theta})), \quad (1)$$

where $\mathbf{u}(\mathbf{y}, \mathbf{X})$ indicates the vector of network statistics corresponding to any *configuration* or/and nodal covariates, $\boldsymbol{\theta}$ indicates the vector of parameters corresponding to $\mathbf{u}$, and ψ depicts a normalizing constant ensuring that the probabilities are proper (sum to 1) (e.g. Snijders [2002], 3; Wasserman and Robins [2005], 152–153; Snijders et al. [2006], 104; Robins et al. [2007], 194). Here, a *configuration* means a set of nodes and a subset of links among them. Different sets of configuration types represent different models. The selection of configuration types is crucial for the analysis and depends on our theoretical assumption about dependencies among possible network links [2] (e.g. Bernoulli random graph, Markov random graph and *partial dependence* graph). The concrete examples of $\mathbf{u}$ will be shown below. In this manner, the probability of observing a graph is assumed to be dependent on the presence of various configurations expressed by the model. This family is called an exponential random graph model (ERGM)

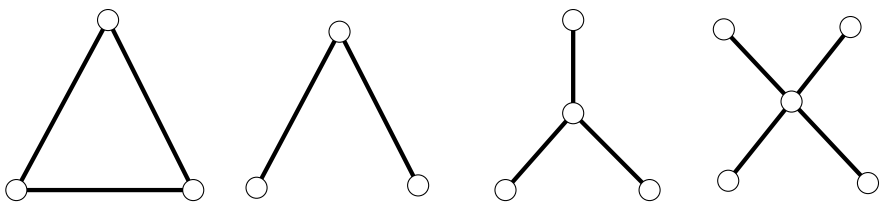
[2]For instance, if the model does not include such configurations formed by more than three nodes as triangles, each link will occur independently with other links, as per the Bernoulli random graph assumption.

because this is an example of what a statistician calls an exponential family distribution with $\mathbf{u}(\mathbf{Y})$ as the sufficient statistic. The model is also known as the $p*$ model (Wasserman and Pattison [1996]).

To begin with, we employ familiar parameters based on the Markov random graph assumption; transitive triad, k-star parameters. The Markov random graphs are a particular sub-class of exponential random graph models in which a possible link from i to j, conditional on the rest of the graph, is assumed to be conditionally dependent only on other possible links involving i and/or j; thus, two links are independent from each other unless they share an actor. This type of dependency resembles a Markov spacial process, and hence, these dependencies were defined as a Markov graph by Frank and Strauss [1986]. A *transitive triad* or a *triangle* is a subset of three mutually connected nodes, and a k-star is a subset of k+1 nodes in which one node is connected by a link to each of the other k nodes (Figure 2). As it can be proved from Hammersley-Clifford theorem that a probability model of a Markov graph depends only on the cliques of the dependence graph [3], that is, the form represented by k-stars or triangles in the original graph, these parameterizations have been broadly employed in network literature.

According to a convention, we also impose the homogeneity constraint by

Figure 2: transitive triad (triangle), 2-, 3- and 4-star



which parameters for isomorphic configurations of nodes are equated (that is, subgraphs are indistinguishable once the node labels are removed). Accordingly, we simply define $u_T(\mathbf{y})$ as the number of transitive triads and $u_k(\mathbf{y})$ as the number of k-stars. Thus, the parameters corresponding to $u_T(\mathbf{y})$ and $u_k(\mathbf{y})$ can be interpreted as the structural effects of transitive triads and k-stars, respectively.

We primarily focus on the structural effect of transitive triad because its concept is closely related to a kind of restriction through indirect relationships, known as the *dyadic constraint* in sociological literature. The simple picture is as follows: A and B belong to a ROSCA, but B also belongs to another ROSCA. From the general perspective regarding a ROSCA, the multiple belonging of B puts the former ROSCA at risk by making it difficult to keep track of B's defecting. However,

[3]The dependence graph is a graph whose nodes correspond to the links of the original graph and represent dependencies among those links.

in the case of *gap*s that have a high density of prevalence, the following situation can frequently occur: A also belongs to the other *gap* that has a member, say C, who belongs to B's *gap* to which A does not belong. Here, C can function as a monitor, and B's behavior against A's interests will be restricted even in the *gap* without A. In this manner, we can hypothesize that the participants of *gap*s that need to be sustainable despite the members' tendencies for multiple belongings are expected to exhibit a higher propensity to engage in the transitive triad structure of the network created by multiple *gap*s.

Further, the parameters of k-stars are important with respect to controlling for degree distribution. For instance, a certain actor may be very popular, and hence, attract links, including those from other popular actors. This process may result in a core-periphery network structure with popular actors at the core. Transitive triads are likely to occur in the core as an outcome of link formation based on popularity. In this case, the number of triads should be considered on the basis of the distribution of the actors' degrees without referring to the transitive triad. Since the number of stars is a function of the degrees, including the parameters of k-stars is equivalent to modeling the degree distribution.

Moreover, it is important to control for node-level effects such as actor attributes, particularly in our analysis. It must be noted that each *gap* is originally organized as an affiliation network (*group*), and thereby, each organizing process follows specific principles. In this sense, links do not occur randomly here, and effects generated by these principles should be controlled. For instance, it is apparent from the observation that the *gap*s follow a principle that they are basically organized by members of the same gender. A woman tends to be connected through a *gap* with another woman, who also tends to be linked to yet another woman. In this manner, triangulation may occur as a result of other principles of the network, without any consideration of dyadic constraints.

As for the nodal attributes, $\mathbf{X}$, we consider the so-called x-related single covariate effects, x-related similarity effects, and x-related identity. For example, considering the individual wage level, it is plausible from previous knowledge about the *gap*s that the actors with higher wages can afford to attend more *gap*s or actors among the same wage level tend to be linked to each other through common *gap*s. To control and examine these effects, we include the wage-level related covariate effects and the wage-level related similarity effects respectively. As for similarity effects, we also examine those of other attributes such as the actors' ages or education levels. Gender or position-related identities are also included in order to control for the tendencies that the actors of the same gender or of the same position in the committee might tend to be linked with each other through *gap*s. These variables are included in the model not only as control variables, but also, evidently, with particular interests for their effects themselves in addition to the transitive triad. The mathematical definitions of all the effects mentioned above are depicted in Appendix A.

In the next step, we attempt to use a new specification proposed by Snijders et al. [2006] that includes the statistics of higher order measures such as alternating k-stars or alternating k-triangles. The motivation of these parameters first came from the failure of Markov random graph models to fit observed networks (See Snijders et al. [2006]; Robins et al. [2007], 195–197). It is often reported that the Maximum Likelihood estimates of the Markov graph cannot be found (MCMC estimation procedure does not converge) and that only the phenomenon called *degeneracy* can be observed (e.g. Handcock et al. [2003])[4]. These new statistics, proposed as successful solutions to avoid degeneracy problems, have important meanings that reflect specific structural properties in the network.

The alternating k-star is proposed as a new higher order star structure. In the conventional Markov model, the star parameters are often limited to a maximum of 3-stars, assuming that the values of the parameters for higher order stars are zero, in order to limit the number of parameters and achieve an identifiable model for which parameters can be estimated. The alternating k-star assumption proposes that, instead of setting higher order star parameters to equal zero, all star parameters should be retained in the model, but with a linear constraint among parameters values, such that, for all $k >= 2$, $\theta_k = -\theta_{k-1}/\lambda$ for some λ greater than 1; the parameters have alternating signs so that positive parameters of some k-star counts are balanced by negative parameters of other k-star counts, and the impact of higher order stars is reduced for higher k.

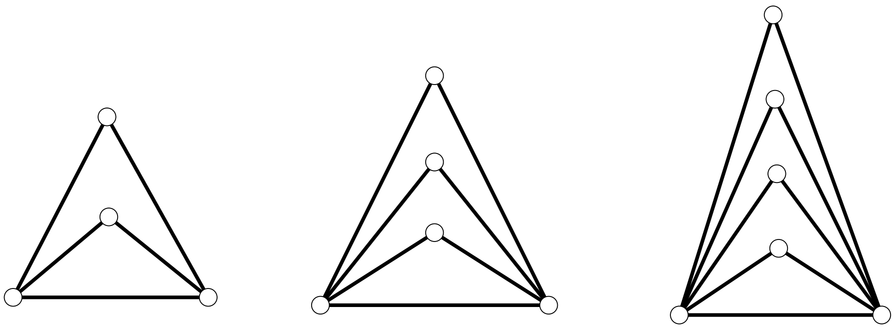
The interpretation of alternating k-star parameters is not straightforward and dependent on other structural parameters, but is broadly interpreted that a positive alternating k-star parameter usually implies a graph that exhibits preference for connections between a large number of low degree nodes and a smaller number of higher degree nodes, akin to a core-periphery structure that is attained by popularity process (Snijders et al. [2006]). Moreover, it is considered to be more preferable to control for degree distribution by alternating k-star parameters, instead of 2 or 3-stars parameters, assuming that the parameters for high orders stars are zero.

The alternating k-triangle is proposed for a new higher order transitive triad structure (Snijders et al. [2006]). A k-triangle is a combination of k individual triangles that all share one link(*tie*). 2-, 3-, and 4-triangles are depicted in Figure 3. The k-triangles can be combined into one statistic in the same way as that for the alternating k-stars. The interpretation of a positive k-triangle parameter is also dependent on other effects in the model, but can be broadly interpreted

[4] It is termed as *degeneracy* or *near-degeneracy* when a model implies that very few distinct graphs are probable, often only empty or complete graphs. For more details, see Snijders et al. [2006] or Robins et al. [2007], 195–197.

as evidence of transitive triad effects in the network. More specifically, a positive k-triangle parameter suggests elements of the core-periphery structure owing to triangulation effects in which the core is built from overlapping triangulation instead of popularity effects. In terms of the hypothesized dependence structure of the network, it must be noted that the alternating k-triangle parameter extends dependence assumption beyond the class of Markov graph models. The Markov graph assumption should be less restrictive and *partial conditional independence* defined by Pattison and Robins [2002] should be assumed, wherein whether two links are independent conditionally on the rest of the graph depends not only on whether they share nodes, but also on the patterns of links in the rest of the graph (For more details of this concept, see Pattison and Robins [2002]; Robins and Pattison [2005]; 202–205; Snijders et al. [2006], 116–117).

Figure 3: 2-, 3- and 4-triangles.



IV. Estimation

The following likelihood function of (1) is simple:

$$L(\boldsymbol{\theta}) = exp(\boldsymbol{\theta}'\mathbf{u}(\mathbf{y}, \mathbf{X}) - \psi(\boldsymbol{\theta})). \tag{2}$$

However, the function is not easy to use because $\psi(\boldsymbol{\theta})$ is dependent on the unknown parameter $\boldsymbol{\theta}$, and hence, is difficult to calculate. Hitherto, two approaches were mainly proposed in literature: (1) Pseudo-likelihood estimation using logistic regression techniques and (2) ML estimation based on MCMC algorithms. To date, pseudo-likelihood estimation proposed by Strauss and Ikeda [1990] has been the most common method for the estimation of ERGMs. It is considered to be intuitively appealing and easily implemented by means of standard logistic regression techniques. However, the properties of the pseudo-likelihood estimators are not well understood, and it is not clear as to when these estimates may be acceptable. In recent years, as a substitute, MCMC algorithms for maximum likelihood (ML) estimation of the parameters in ERGMs were proposed by Snijders [2002]. The key idea of the MCMC method is refining the approximate parameter estimates by comparing the observed graphs against a distribution of random graphs generated by stochastic simulation using the approximate parameter values. The process is repeated until the parameter estimates are stabilized (converged). If the parameter estimates never converge, the model is likely to degenerate. When convergent estimates are obtained, simulation from the estimates will produce the distribution of graphs in which the observed graph is typical for all of the effects in the model. One of the advantages of this new method is that one can also obtain reliable standard errors for the estimate. This paper conducted MCMC estimation using the SIENA program (Snijders et al. [2007]), version 3.2, by implementing the Metropolis Hastings algorithm for generating draws from the exponential random distribution based on the stochastic approximate algorithm described in Snijders [2002]. More details of the MCMC algorithms for ERGMs will be provided in Appendix B.

V. Results

The estimation results are presented in Table 2. In addition to estimates and standard errors, the output includes a convergence t-ratio for each estimate. This convergence t-ratio indicates how well the estimate has converged, with a small value close to zero that suggests good convergence. After obtaining the estimate, the program performs a large number of Metropolis-Hastings steps with these parameter values and calculates deviations between observed values of statistics ($\mathbf{u}(\mathbf{y})$) and averages of simulated values of the statistics, which are the estimated expected values ($E_{\hat{\theta}}\mathbf{u}(\mathbf{Y})$). The convergence t-ratios are averages of these deviations divided by their standard deviations. It is considered to be excellent when the ratio are less than 0.1 and good when they are less than 0.2 in absolute value (Snijders et al. [2007], 29). Meanwhile, the parameter estimates were tested by using the t-statistics defined as parameters divided by standard error, and by referring these to an approximating standard normal distribution as the null distribution. If these values are larger than 2 in absolute values, those estimates are regarded as significant at the 0.05 significance level.

Table 2 represents only the reduced models that could sufficiently converge[5]. As for structural effects, the transitive triad parameter in the Markov model

[5] Here, the conditional simulation/estimation method that keeps the total numbers of links(*ties*) fixed was employed. The converged estimates were obtained after the 2-4 repeated runs of the programs using the previously obtained estimates as the initial values for the new estimation and tuning constants such as probabilities of inversion steps or multiplication factors, according to the guidance by Snijders et al. [2007]. See also Appendix B. The convergences were much easier to

Table 2: MCMC Parameter Estimates for the Leaders' *Gap* Network in *Oqmahalla*

Variables	Leaders' gap network (N=45)					
	Model I		Model II		Model III	
	Convergence t-ratio	Estimated Coefficient (S.E.)	Convergence t-ratio	Estimated Coefficient (S.E.)	Convergence t-ratio	Estimated Coefficient (S.E.)
Transitive triads	-0.115	1.1320 (0.1559)*	-0.133	1.1323 (0.1741)*	0.103	1.1495 (0.1651)*
2-stars	-0.125	0.1629 (0.1818)	-0.133	0.1570 (0.1737)	0.084	0.1380 (0.1860)
3-stars	-0.125	-0.1030 (0.0585)	-0.112	-0.1004 (0.0557)	0.090	-0.1001 (0.0573)
Same gender	-0.094	1.6788 (0.5631)*	-0.125	1.6495 (0.5420)*	-0.097	1.6122 (0.5687)*
Same position	-0.066	-0.5675 (0.3490)	-0.085	-0.5806 (0.3505)	0.008	-0.5972 (0.3531)
Age similarity	0.014	5.3698 (1.4327)*	-0.060	5.3555 (1.4292)*	0.162	5.3772 (1.3746)*
Wage size	-	-	-	-	0.082	0.0259 (0.0924)
Wage similarity	-	-	-0.272	0.3273 (0.5874)	-0.006	-0.0182 (0.6662)
Education similarity	-	-	-	-	-0.112	0.7505 (0.5973)
	Model IV		Model V		Model VI	
	Convergence t-ratio	Estimated Coefficient (S.E.)	Convergence t-ratio	Estimated Coefficient (S.E.)	Convergence t-ratio	Estimated Coefficient (S.E.)
Alternating k-stars, λ=3	0.014	-0.6481 (0.1350)*	-0.001	-0.6609 (0.1303)*	-0.007	-0.6720 (0.1531)*
Alternating k-triangles, λ=3	0.050	1.3098 (0.1760)*	0.054	1.3121 (0.1650)*	-0.055	1.3062 (0.1773)*
Same gender	-0.021	1.2290 (0.4967)*	0.011	1.2357 (0.5014)*	-0.034	1.2951 (0.5002)*
Same position	0.032	-0.5285 (0.3364)	0.088	-0.4905 (0.3400)	-0.080	-0.6044 (0.3500)
Age similarity	0.109	4.5478 (1.2729)*	0.059	4.7177 (1.1881)*	-0.103	4.6674 (1.2075)*
Wage size	-	-	0.078	0.0535 (0.0806)	0.019	-0.0098 (0.0836)
Wage similarity	-	-	-0.248	-0.1637 (0.4911)	-0.020	-0.6237 (0.5921)
Education similarity	-	-	-	-	0.098	0.9286 (0.4663)

* $p < 0.05$

(Model I, II and III) and the alternating k-triangle parameter in the higher order model (Model IV, V and VI) have t-statistics greater than 2, and they suggest tendencies for triangulation in this network. Furthermore, we can observe that the negative k-star parameter is significant in the higher order model; the interpretation of these complex structural effects will be discussed in the next paragraph. With regard to nodal covariates, the age similarity and gender identity parameters are significant as expected. however, we cannot find evidence that there are any important effects distinguishing partners in their level of wage or education because the education similarity or wage similarity parameters are not significant.

A notable feature in respect of the network structure is that both of the negative k-star and positive k-triangle estimates are significant in the result of the higher order model. The combination of negative k-star and positive k-triangle estimate is, in actuality, not uncommon and is discussed in the literature (e.g. Robins et al. [2007], 205–207). It is considered that there are two countervailing tendencies: one toward a triangulated core-periphery structure and one against a degree-based core-periphery structure. Particularly, *single-core periphery structure implied by the positive k-triangle effect has been segmented by the inclusion of negative k-star parameter into a chain of smaller dense regions of the network*

obtain in the higher order model than the Markov random graph model, as was expected.

(Robins et al. [2007], 206). From the simulation, it can be shown that the shape of the network with a fixed positive k-triangle parameter moves from centralization to segmentation when k-star parameters range from zero to increasingly negative values.

What implication can we derive from these findings regarding the structural characteristic of the leaders' network through the *gap*s in *Oqmahalla*? The results implied that the *gap*s appear to contribute to the formations of the transitive triad structure within the leader's network, but work against the degree-based core-periphery structure. From the perspective of sustainability issues of ROSCAs, the strong propensity of transitive triad structure can be related to a feature of *gap*s: the scarcity of default problems. As stated previously, the triangulated triad structure of the network can provide actors with dyadic constraints and thereby help them to keep track of defecting members. If the tendencies of members to belong to multiple *gap*s can encourage more formation of transitive triad relationships, the generous attitude toward multiple belongings to *gap*s in *Oqmahalla* will not necessarily be irrational even with regard to their sustainability issues. On the other hand, it is not straightforward to detect meaningful linkages between propensities against the core-periphery structure and *gap*s' features or sustainability issues because of the lack of existing works that deal with a network structural aspect of ROSCAs. A certain recent study that deals with the stability of the self-enforcing insurance mechanism from a network-related perspective implies that a network of intermediate density tends to be unstable, as compared to those of high or low densities (Bloch et al. [2008]). If the propensity against core-periphery structure leads the network to that of intermediate density without the core of high density or the periphery of low density, the *gap*'s network in a mass might be unexpectedly unstable. These questions remain to be addressed. To answer them, more detailed studies concerning the theoretical aspects of network formation of ROSCAs as well as empirical studies covering their longitudinal process will be required.

VI. Concluding Remarks

This paper attempted to clarify the structural propensities of a network created by ROSCAs or *gap*s in rural Uzbekistan, by applying ERGMs to a network of *mahalla* committee members and estimating the parameters corresponding to structural effects on the basis of MCMC estimation methods. The results implied that the *gap*s contributed to the formation of the transitive triad structure within the leaders' network, but work against the degree-based core-periphery structure. The propensity of the network toward forming transitive triads appears to contribute to the network's stability. However, to understand the meaning of the results further, we should further conduct theoretical studies on ROSCAs that form part of the network framework as well as empirical studies covering their longitudinal process.

Some comments on the methodology applied to the paper are provided as follows. To the best of our knowledge, this is the first paper applying ERGMs to the network of the ROSCAs, or the subject of development economics. Though the linkage between the paper's empirical results and theories is still weak, the methodology can be extended easily and applied to many research works in various directions. The targeted network in this paper was, in fact, not necessarily suitable for the application of an ERGM. This model will further enable the conduction of analyses on directed networks such as private income transfers, gift-giving, information flows, or aid delivery. The results from these directed networks may be applicable for practical purposes such as the projection of various development programs, considering that network simulations based on estimated or user-defined parameters can be easily implemented. In addition, the size of sample nodes in this paper was rather small. Recent methodological advances in statistical network analysis have greatly improved our ability to conduct inferences and allowed for analyses that can cover more than 1,600 nodes (e.g., Goodreau [2007]. The analysis of the network structure of an entire community based on complete household surveys will contribute to our understanding of a community structure. Furthermore, if provided with panel data, we can adopt the longitudinal models to which MCMC methods can be more directly applied, and thus, obtain more suggestive results regarding the formation process of a community.

Appendix A: Mathematical Definition of Effects

The notation in this part is the same as that of the other part of the paper. Replacing an index by a "+" indicates summation over this index.

–The number of two stars
$u_2 = \frac{1}{2}\sum_i \sum_{h \neq k} y_{hi} y_{ik}$

–The number of three stars
$u_3 = \sum_i \binom{y_{i+}}{3}$

–The number of transitive triads
$u_T = \frac{1}{6}\sum_{i,j,h} y_{ij} y_{ih} y_{jh}$

–The x_i-related single covariate effect
$u_{co} = \sum_i y_{i+} x_i$

–The x_i-related similarity effect
$u_{sim} = \sum_{i,j} y_{ij}(sim^x_{ij} - \overline{sim^x})$,
where $\overline{sim^x}$ is the mean of all similarity scores, which are defined as $sim^x_{ij} = \frac{\Delta - |x_i - x_j|}{\Delta}$ with $\Delta = max_{ij}|x_i - x_j|$ being the observed range of the covariate x

–The x_i-related identity
$u_{id} = \sum_{i,j} y_{ij} I\{x_i = x_j\}$,
where $I\{x_i = x_j\}$ is 1 if the condition $x_i = x_j$ is satisfied, and 0 if it is not.

–The alternating k-stars
$u_{kS} = S_2 - \frac{S_3}{\lambda} + \frac{S_4}{\lambda^2} - ... + (-1)^{n-2}\frac{S_{n-1}}{\lambda^{n-3}}$
$= \sum_{k=2}^{n-1}(-1)^k \frac{S_k}{\lambda^{k-2}}$
$= \lambda^2 \sum_{i=1}^{n}\{(1 - \frac{1}{\lambda})^{y_{i+}} + \frac{y_{i+}}{\lambda} - 1\}$,
where $S_k = \sum_i \binom{y_{i+}}{k}$ $(k \geq 2)$ (number of k-stars)

–The alternating k-triangles
$u_{kT} = 3T_1 - \frac{T_2}{\lambda} + \frac{T_3}{\lambda^2} - ... + (-1)^{n-3}\frac{T_{n-2}}{\lambda^{n-3}}$
$= \sum_{i<j} y_{ij} \sum_{k=1}^{n-2}(\frac{-1}{\lambda})^{k-1}\binom{L_{2ij}}{k}$
$= \lambda \sum_{i<j} y_{ij}\{1 - (1 - \frac{1}{\lambda})^{L_{2ij}}\}$,
where $T_k = \sum_{i<j} y_{ij}\binom{L_{2ij}}{k}$ (for $k \geq 2$) (number of k-triangles),
$T_1 = \frac{1}{3}\sum_{i<j} y_{ij} L_{2ij}$ and $L_{2ij} = \sum_h y_{ih} y_{hj}$.

Appendix B: MCMC Estimation for ERGMs

The MCMC algorithms for maximum likelihood (ML) estimation of the parameters in ERGMs were proposed by Snijders [2002]. This is now implemented in computer programs SIENA (Snijders et al. [2007]) and statnet (Handcock et al. [2008]). This appendix describes the major aspects of the algorithms with an emphasis on the manner followed in the case of SIENA. For further information, please see Snijders [2002]; Snijders et al. [2006]; Snijders et al. [2007]. The notation is the same as that of the other part of the paper. Random variables are indicated by capital letters, vectors, and matrices by bold letters.

MCMC estimation algorithms in the exponential random graph model are attained by repeatedly using MCMC simulation algorithms. Here, MCMC estimation algorithms refer to the estimation of the parameter for a given data set, whereas MCMC simulation algorithms mean simulating a random draw from the exponential random graph distribution with a fixed parameter value.

–MCMC simulation

The new digraph Y_{ij} is assumed to be generated by a Monte Carlo method according to the exponential random graphs model with parameter $\boldsymbol{\theta}$, dependent on the current step. A straightforward way to generate random samples from exponential random distributions is to use Gibbs sampler. Y_{ij} is assumed to be generated according to conditional distribution as follows:

$$P_{\boldsymbol{\theta}}\{Y_{ij}^{(t+1)} = a|\mathbf{Y}^{(t)} = \mathbf{y}^{(t)}\} = P_{\boldsymbol{\theta}}\{Y_{ij} = a|\mathbf{Y}_{hk} = \mathbf{y}_{hk}^{(t)} \quad for\, all\, (h,k) \neq (i,j)\} \quad (a = 0,1). \tag{3}$$

The left-hand side is the transition probability. The matrix $\mathbf{Y}^{(t)}$ and $\mathbf{Y}^{(t+1)}$ differ in only one element. Now, let us cycle through the set of all random variables $Y_{ij}(i \neq j)$ and simulate each in turn according to the above conditional distribution. The right-hand side is the same conditional distribution as that used in the pseudolikelihood procedure. Thus, to obtain conditional probabilities (1) from exponential random distribution, the well-known formulation to derive the pseudolikelihood function can be directly applied. In this formulation, the conditional distribution is defined by a logistic regression model where sufficient statistics are given by the difference between the values for $\mathbf{u}(\mathbf{y})$ obtained when letting $y_{ij} = 1$ or $y_{ij} = 0$, and leaving all other elements of $\mathbf{y}$ unchanged. For details, see Strauss and Ikeda [1990] or Wasserman and Pattison [1996].

If the selection of pair (i, j) cycles systematically through all elements of the adjacency matrix, it is termed as *cycling*. In contrast, if the selection of pair (i, j) occurs randomly under the condition $i \neq j$, it is called *mixing*.

Another alternative is the Metropolis-Hastings Algorithm. In this algorithm, after the selection of the pair (i, j), the probability for changing the cell value Y_{ij}

is given by

$$P_\theta\{Y_{ij}^{(t+1)} = 1 - y_{ij}^{(t)} | \mathbf{Y}^{(t)} = \mathbf{y}^{(t)}\} = min\{1,\ exp(\boldsymbol{\theta}'(\mathbf{u}(\mathbf{y}^{(ijc)}) - \mathbf{u}(\mathbf{y})))\}, \tag{4}$$

where $\mathbf{y}^{(ijc)}$ is the adjacency matrix obtained from $\mathbf{y}$ by changing element y_{ij} into $1 - y_{ij}$ and leaving all elements unchanged.

SIENA uses Gibbs sampler or Metropolis-Hastings algorithms for updating the steps. Moreover, the steps can be updated not only by dyad, but also by single arcs or triplets.

SIENA performs an inversion step with a rather small probability (default is 0.01) at each step, instead of that of the basic algorithm. Here, inversion means the biggest step that changes the graph to its components. This is because the basic algorithms occasionally converge very slowly owing to the very small steps taken. The Gibbs sampler and Metropolis-Hastings algorithms define specific updating probabilities for the inversion steps respectively (Snijders [2002], 20).

The present paper employed the Metropolis-Hastings algorithm for updating each dyad because the Metropolis-Hastings algorithm changes $\mathbf{Y}^{(t)}$ more frequently than the Gibbs sampling, and therefore, is often considered to be more efficient. The paper also employed the inversion step of the default value for the first 2–3 repeated runs. Later, it decreased the value to 0.0001 for the last run.

–MCMC estimation

The maximum likelihood (ML) estimate $\hat{\boldsymbol{\theta}}(\mathbf{y})$ for an exponential family is also the solution of the moment equation. That is (Snijders [2002], 22),

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = \mathbf{u}(\mathbf{y}), \tag{5}$$

where $\mathbf{u}(\mathbf{y})$ is the observed value and $\boldsymbol{\mu}(\boldsymbol{\theta})$ is defined as the expected value

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}}\{\mathbf{u}(\mathbf{Y})\} \tag{6}$$

that $\boldsymbol{\mu}(\boldsymbol{\theta})$ is also the gradient of $\psi(\boldsymbol{\theta})$

$$\mu_k(\boldsymbol{\theta}) = \partial\psi(\boldsymbol{\theta})/\partial\theta_k, \tag{7}$$

that the covariance matrix $\sum(\boldsymbol{\theta})$ =cov($\mathbf{u}(\mathbf{Y})$) of $\mathbf{u}(\mathbf{Y})$ with elements $\sigma_{hk}(\boldsymbol{\theta})$ is the matrix of derivatives of $\boldsymbol{\mu}(\boldsymbol{\theta})$,

$$\sigma_{hk}(\theta) = \partial\mu_k/\partial\theta_h = \partial^2\psi(\theta)/\partial\theta_h\partial\theta_k, \tag{8}$$

and that the asymptotic covariance matrix of the ML estimator $\hat{\boldsymbol{\theta}}$ is given by

$$cov_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}) = (\sum(\boldsymbol{\theta}))^{-1}. \tag{9}$$

However, it is not straightforward to apply this method to derive ML estimators for exponential random graph model because the calculations of the functions $\psi(\boldsymbol{\theta})$ and $\boldsymbol{\mu}(\boldsymbol{\theta})$ become extremely difficult except for the simplest model. To derive the ML estimator, SINENA employs the Robbins-Monro algorithm based on Monte Carlo simulation (Snijders [2002], 25–28), which intends to solve equations of the form

$$E\{\mathbf{Z}_\theta\} = 0, \tag{10}$$

with $\mathbf{Z}_\theta$ given by

$$\mathbf{Z}_\theta = \mathbf{u}(\mathbf{Y}) - \mathbf{u}_0 \tag{11}$$

where $\mathbf{u}_0 = \mathbf{u}(\mathbf{y})$ is the observed value of the sufficient statistic and $\mathbf{Y}$ has probability distribution (1) with arbitrary parameter $\boldsymbol{\theta}$.

The iteration step in the Robbins-Monro procedure for solving (10), with step-size $\mathbf{a}_n$, is

$$\hat{\boldsymbol{\theta}}^{(n+1)} = \hat{\boldsymbol{\theta}}^{(n)} - \mathbf{a}_n' \mathbf{D}_n^{-1} \mathbf{Z}(n), \tag{12}$$

where $\mathbf{Z}(n)$ for $n = 1, 2, ...$ are random variables such that the conditional distribution of $\mathbf{Z}(n)$ given $\mathbf{Z}(1), ..., \mathbf{Z}(n-1)$ is the distribution of $\mathbf{Z}_{\boldsymbol{\theta}}$ obtained for $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}^{(n)}}$. $\mathbf{D}_n$ is the covariance matrix of $\mathbf{u}(\mathbf{Y})$ at the step n (given by (8)). The step sizes $\mathbf{a}_n$ are a sequence of positive numbers converging to zero (e.g. $a_n = 1/n$).

References

Ambec, Stefan and Nicolas Treich, “Roscas as Financial Agreements to Cope with Self-Control Problems,” *Journal of Development Economics*, January 2007, *82* (1), 120–137.

Ardener, Shirley, “The Comparative Study of Rotating Credit Associations,” *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 1964, *94(2)*, 201–229.

Bala, Venkatesh and Sanjeev Goyal, “A Noncooperative Model of Network Formation,” *Econometrica*, September 2000, *68* (5), 1181–1230.

Besag, Julian, “Spatial Interaction and the Statisitical Analysis of Lattice Systems,” *Journal of the Royal Statistical Society*, 1974, *series B 36*, 192–225.

Besley, Timothy and Alec R Levenson, “The Role of Informal Finance in Household Capital Accumulation: Evidence from Taiwan,” *Economic Journal*, January 1996, *106* (434), 39–59.

— **, Stephen Coate, and Glenn Loury**, "The Economics of Rotating Savings and Credit Associations," *American Economic Review*, September 1993, *83* (4), 792–810.

Bloch, Francis, Garance Genicot, and Debraj Ray, "Informal insurance in social networks," *Journal of Economic Theory*, November 2008, *143* (1), 36–58.

Calomiris, Charles W. and Indira Rajaraman, "The role of ROSCAs: lumpy durables or event insurance?," *Journal of Development Economics*, June 1998, *56* (1), 207–216.

Comola, Margherita, "The network structure of informal arrangements: Evidence from rural Tanzania," PSE Working Papers 2008-74, PSE (Ecole normale sup?rieure) August 2008.

Deaton, Angus, *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy (World Bank)*, World Bank, 8 1997.

Fafchamps, Marcel and Flore Gubert, "The formation of risk sharing networks," *Journal of Development Economics*, July 2007, *83* (2), 326–350.

— **and Susan Lund**, "Risk-sharing networks in rural Philippines," *Journal of Development Economics*, August 2003, *71* (2), 261–287.

Frank, Ove and David Strauss, "Markov Graphs," *Journal of the American Statistical Association*, 1986, *81*, 832–42.

Geertz, C., "The Rotating Credit Association: A "Middle Rung" in Development," *Economic Development and Cultural Change*, 1962, *10(3)*, 241–263.

Genicot, Garance and Debraj Ray, "Group Formation in Risk–Sharing Arrangements," *Review of Economic Studies*, January 2003, *70* (1), 87–113.

Goodreau, Steven M., "Advances in exponential random graph (p*) models applied to a large social network," *Social Networks*, 2007, *29(2)*, 231–248.

Gugerty, Mary Kay, "You Can't Save Alone: Commitment in Rotating Savings and Credit Associations in Kenya," *Economic Development and Cultural Change*, January 2007, *55* (2), 251–82.

Handa, Sudhanshu and Claremont Kirton, "The economics of rotating savings and credit associations: evidence from the Jamaican 'Partner'," *Journal of Development Economics*, October 1999, *60* (1), 173–194.

Handcock, Mark S., David R. Hunter, and Carter T. Butts et al., “statnet: Software Tools for the Representation, Visualization, Analysis and Simulation of Network Data,” *Journal of Statistical Software*, 2008, *24.*

_ **, Garry Robins, Tom Snijders, Jim Moody, and Julian Besag**, “Assessing Degeneracy in Statistical Models of Social Networks,” *Journal of the American Statistical Association*, 2003, *76*, 33–50.

Hiwatari, Masato, *The Customary Economy and Economic Development: The Community-based Structure of a Mahalla in Uzbekistan (in Japanese)*, University of Tokyo Press, 2008.

_ , “Traditions and the Informal Economy in Uzbekistan: A Case Study of Gaps in the Andijan Region,” *Acta Slavica Iaponica*, 2008, (25), 43–66.

Jackson, Matthew O., *Social and Economic Networks*, Princeton Univ Pr, 2008.

_ **and Asher Wolinsky**, “A Strategic Model of Social and Economic Networks,” *Journal of Economic Theory*, October 1996, *71* (1), 44–74.

Johnson, Susan, “Gender Norms in Financial Markets: Evidence from Kenya,” *World Development*, August 2004, *32* (8), 1355–1374.

Kandiyoti, D., “Rural livelihoods and social networks in Uzbekistan: Perspectives from Andijan,” *Central Asian Survey*, 1998, *17*, 561–578.

Kimuyu, Peter Kiko, “Rotating Saving and Credit Associations in Rural East Africa,” *World Development*, July 1999, *27* (7), 1299–1308.

Kovsted, Jens and Peter Lyk-Jensen, “Rotating savings and credit associations: the choice between random and bidding allocation of funds,” *Journal of Development Economics*, October 1999, *60* (1), 143–172.

Krishnan, Pramila and Emanuela Sciubba, “Links and Architecture in Village Networks,” *Economic Journal*, 04 2009, *119* (537), 917–949.

Pattison, Philippa and Garry Robins, “Neighborhood-Based Models For Social Networks,” *Sociological Methodology*, 2002, *32*, 301–337.

Robins, Garry and Philippa Pattison, “Interdependencies and Social Processes: Dependence Graphs and Generalized Dependence Structure,” In Carrington, Scott & Wasserman (Eds), *Models and Methods in Social Network Analysis*, 2005, Cambridge University Press, 192–214.

___ **, Tom Snijders, Peng Wang, Mark Handcock, and Philippa Pattison**, "Recent developments in exponential random graph (p*) models for social networks," *Social Networks*, May 2007, *29* (2), 192–215.

Sievers, Eric W., "Uzbekistan's Mahalla: From Soviet to Absolutist Residential Community Associations," *The Journal of International and Comparative Law at Chicago-Kent*, 2002, *2*, 91–158.

Snijders, Tom A. B., "Markov Chain Monte Carlo Estimation of Exponential Random Graph Models," *Journal of Social Structure*, 2002, *3*.

Snijders, Tom. A. B., Christian Steglichand, and Michael Schweinberger et al., *Manual for SIENA version 3* 2007.

Snijders, Tom A. B., Philippa E. Pattison, Garry L. Robins, and Mark S. Handcock, "New Specifications for Exponential Random Graph Models," *Sociological Methodology*, December 2006, *36* (1), 99–153.

Strauss, David and Michael Ikeda, "Pseudolikelihood Estimation for Social Networks," *Journal of the American Statistical Association*, 1990, *85*, 204–212.

Townsend, Robert M, "Risk and Insurance in Village India," *Econometrica*, May 1994, *62* (3), 539–91.

Wasserman, Stanley and Garry Robins, "An Introduction to Random Graphs, Dependence Graphs, and p*," In Carrington, Scott & Wasserman (Eds), *Models and Methods in Social Network Analysis*, 2005, Cambridge University Press, 148–161.

___ **and Philippa Pattison**, "Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and p*," *Psychometrica*, 1996, *61(3)*, 401–425.

Weerdt, Joachim De and Stefan Dercon, "Risk-sharing networks and insurance against illness," *Journal of Development Economics*, December 2006, *81* (2), 337–356.