2010

The emergence of local elite networks: Structure or preference? - An econometric approach -

Christian Assmann
University of Bamberg, christian.assmann@uni-bamberg.de

Christian H.C.A. Henning
University of Kiel, chenning@ae.uni-kiel.de

The emergence of local elite networks: Structure or preference?
- An econometric approach -

Christian Aßmann**and Christian Henning**
University of Bamberg* and University of Kiel**

May 17, 2010

Abstract

This paper analyzes the determinants and causes of communication in local elite networks. The database comprises four rural county elite networks from Poland and Slovakia. Socio spatial processes allowing a flexible incorporation of individual specific information are embedded within a logit framework. Empirical analysis focuses on the assessment of the hypotheses, whether preferences measured by socio demographic factors and political ideology or institutional settings (structure) influence individual communication in local elite networks. The results suggest that while in high performing communities institutional settings, i.e. a common membership in local organizations, are the most important factors determining communication, in low performing communities communication ties are stronger determined by actors' preferences, i.e. ideological distances and socio demographic factors. Moreover, communication is more centralized for the latter when compared to the former socio spatial process.

*Corresponding address: Tel.: +49-951-863 3423. *E-mail:* christian.assmann@uni-bamberg.de (C. Aßmann)

1 Introduction

Early studies of policy networks focused on social network structures among governmental and non-governmental organizations to explain political decision-making (Parsons, 1963) and (Coleman, 1963). In particular, Laumann and Knoke (1987) and Knoke et al. (1996) have developed social influence models to explain opinion formation within a political communication process, where governmental actors partly adopt their policy positions to the positions communicated by other non-governmental organizations. However, while early policy network studies relate network structures nicely to the political influence of individual actors, these studies do neither provide a rational model of political influence nor do they relate policy network structures to political performance at the macro level. In contrast, Henning (2009) suggest a theoretical model that provides a rationality behind social influence of non-governmental actors on the political position of governmental agents and allows the relation of policy network structures with the efficiency of political decision-making at the macro level. In particular, the rationality of political influence follows from the fact that from the viewpoint of political agents political decision making is characterized by a fundamental uncertainty regarding the impact of policies on the state of the world. Thus, while most politicians have a clear preference regrading the desirable state of the world, they have only limited and incomplete information on the political technology, i.e. how different policy instruments actually translate into a specific state of the world. Accordingly, agents have to choose among policy alternatives although they are uncertain regarding the evaluation of different alternatives. But in a world of uncertainty it turns out that maximization of individual utility can only be achieved by some supplement strategies. For example, to be able to make a rational choice in these situations agents form beliefs regarding the uncertain impact of various policy alternatives on the state of the world and thus on their utility. Thus, in a social context characterized by uncertainty information available on the benefits and harms that might results from a decision is a crucial factor that influences the decision of agents. In this context Henning and Saggau (2010) analyze in a agent based-model framework how political communication network structures among a local elite influences information aggregation via communication and hence overall efficiency of local government. In particular, Henning and Saggau (2010) demonstrate that political communication among governmental and non-governmental actors implies both a more efficient learning of the true political technology and a policy bias towards particular interest of local community. Given a policy bias of local elite the overall impact of communication on the efficiency of local government decision-making depends on the network structures, i.e. random networks are c.p. more efficient when compared to clustered or centralized communication network. Therefore, communication networks correspond to information aggregation mechanisms and hence can be interpreted as social capital in the sense of Coleman or Burt, where concrete individual and collective values of a communication network structures depend on the specific framework conditions.

Please note that also economists have taken up the idea of social influence models to explain agents opinion formation, e.g. models of herding behavior (Krause, 2004), where Battiston et al. (2004) explicitly analyzed the role of social networks in agents' opinion formation and decision. Moreover, Bala and Goyal (1998) analyze belief formation in a social network. However, they do not analyze how specific network structures influence agents' belief formation. Later Gale and Kariv (2003) as well as Choi et al. (2004) or Celen et al. (2004) explore the interaction between network structures and beliefs, but they focus their analysis on small networks (3 nodes) and have not considered large and more complex networks. More recently Currarini et al. (2009) analyzes in a very interesting theoretical paper the impact of communication network structures among a set of actors on their opinion formation.

However, Currarini et al. (2009) have not yet analyzed how communication network structures influence the overall efficiency of collective decision making.

Given the importance of social network structures on economic and political behavior at the micro level and induced outcomes at the macro level, it is interesting to understand how social network structures emerge in the first place or can be changed.

In this regard Moody (2004) pointed out that the formation of network structures are basically determined by two different processes: actors' preferences and the structure of the meeting process among the set of relevant actors. On the one hand actors choose their network contacts according to their preferences, while on the other hand the meeting process determines the probability that two actors actually have the opportunity to form a network tie.

In this context Currarini et al. (2009) have recently analyzed homophily and self-selection in schools focusing on the relative importance of preferences versus matching structures in an extremely interesting paper. In particular, Currarini et al. (2009) provide a theoretical model to show that observed friendship structures, e.g. homophily, imbreeding homophily and the tendency that larger groups from more friendships per capita, could be nicely derived from a simple theoretical model of friendship formation assuming both a biased meeting (matching) process and biased preferences for a friendship with the same-type. Furthermore, using aggregate friendship data taken form the Add Health data set they could also empirically specify the parameter of their simple model implying that both biased preferences and a biased matching process generates observed friendship patterns in American schools. However, Currarini et al. (2009) did not provide a microeconometric estimation of the underlying network generating process taking into account individual characteristics to estimate network ties among pairs of individual actors.

In this paper we suggest an empirical framework allowing adequate analysis of communication structures between individual member of a local political elite. Technically, we apply a socio spatial Gaussian kernel process suggested by Linkletter et al. (2006) to perform aggregation of individual specific information.

Empirical analysis of political communication networks among local elites in two Polish and four Slovakian rural communities suggest that structure is more important than preference, i.e. the number of common organizational memberships is among the most important determinants of social communication in local elite networks. Furthermore, with respect to individual specific variables, political party membership is identified as an factors shaping the distance in the socio-spatial space. However, these findings are not confirmed for all six communities, i.e. especially in low performing communities ideological distances between actors are the most determining factor of social communication. Interestingly, local elite network structures are not only more biased towards special interest in low performing when compared to high performing communities, but elite network structures are also stronger determined by actors' ideological preferences and less by structural meeting opportunities like overlapping organizational memberships in the former.

The paper proceeds as follows. The database is described in Section 2. Section 3 reviews the two different model frameworks, the estimation methodology and the data generating process. The empirical results are discussed within Section 4. Section 5 concludes.

2 Data Description

Via personal interviews local elite communication networks have been surveyed in political counties in Poland and Slovakia. Within the interviews, several questions have been devoted to enquiry of political attitudes in different political fields, e.g. question referring to political priorities concerning public services, infrastructure, etc. This personal attitudes are summarized within a political conflict index providing dyad specific information on the probability to observe communication between local elite members.[1] A further dyad specific information variable is provided via the also interviewed organizational memberships of local elite members, i.e. it is counted how often two local elite members are member of the same organizations, e.g. social, political, or religious ones.[2] As individual specific information serve the variables age (in years), highest educational degree[3], a dummy variable indicating political party membership, personal reputation assessed via a number of nominations within a asked reputation network, and a social prestige index referring to job descriptions. The social prestige index is constructed based on the asked job occupations of local elite network members, see van der Gaag (2005) for details on the construction approach. Summary statistics for the variables under consideration are provided in Table (1). The summary statistics show differences with respect to educational level and age between the counties, with Chotza having youngest local elite on average and a lower educational level in comparison with the other counties. This structural differences possibly point at differences in the communication process within the counties, which are assessed within the empirical subsequent empirical analysis.

3 Model Framework and Estimation

Social network models often provide metric or binary measurements on links between n network constituents, which can be summarized via

$$\begin{pmatrix} -- & y_{12} & \cdots & y_{1n-1} & y_{1n} \\ y_{21} & -- & & & \\ \vdots & & -- & & \vdots \\ y_{n-11} & & & \ddots & y_{n-1n} \\ y_{n1} & y_{n2} & \cdots & y_{nn-1} & -- \end{pmatrix}.$$

If the network is assumed to be symmetric implying $y_{ij} = y_{ji}$ the network provides a total of $n(n-1)/2$ observations summarized within the vector Y. Social network models provide then a link between observation network relations and explaining factors. This is often done using a regression function of the type

$$g(y_{ij}) = X_{ij}\beta + e_{ij},$$

where $g(\cdot)$ is given as

$$g(y_{ij}) = \begin{cases} y_{ij}, & \text{if } y_{ij} \text{ has metric scale;} \\ \eta_{ij}, & \text{if } y_{ij} \text{ is binary,} \end{cases}$$

[1]The index is calculated as follows

$$C_{ij} = \sqrt{\sum_{k=1}^{K} |p_{ik} - p_{jk}| d_{ik} d_{jk}},$$

where d_{ik} denotes the interest of individual i in policy field k and p_{ik} the position of individual i in policy field k.

[2]All organizations are classified according to their main purpose as follows: (1) policy and administration, (2) agricultural economics, (3) industrial economics, (4) handcraft economics, (5) trade economics, (6) consumer economics, (7) cultural/educational/media, (8) associations and clubs, (9) religious/church, (10) other organization purposes.

[3]The following ordinal scheme is used: (1)- unfinished primary school, (2)-completed primary school, (3)-vocational training, (4)-unfinished high school, (5)-completed high school without matura, (6)-completed high school with matura, (7)-general gymnasium certificate, (8)-completed professional training, (9)-Bachelor diploma at university, (10)-completed university study, Magister.

where the link between η_{ij} and y_{ij} is

$$\eta_{ij} > 0, \quad \text{if } y_{ij} = 1, \qquad \eta_{ij} \leqslant 0, \quad \text{if } y_{ij} = 0.$$

X_{ij} thereby reflects dyad specific characteristics, e.g. physical distance in business networks (for all network constituents summarized in X). Next to dyad specific characteristics available, information is often gathered with respect to network constituents not with respect to social relations. Hence, a natural approach to augment the regression equation via suitable chosen sufficient network statistics typically given as the functions of the total number of ties in the network or the number of reciprocal ties, encounters difficulties in estimation with respect to the ad hoc choice of suitable variable transformations and network indicators, see Linkletter (2007). Furthermore, neglecting individual specific structures causes a considerable lack of model fitness, see Hoff (2005), Snijders (2002), and Handcock (2003) for discussion.

To overcome the matter of lacking fit when using network statistics for explaining social relations, Hoff et al. (2002) discuss several network models, where latent variables are used for characterization of the relative position of the actors in an *unobserved* social space. Furthermore, Hoff (2007) provides a characterization of the latent social space via a reduced rank matrix obtained from a singular value decomposition. Since these characterizations of the latent social space do not allow observed characteristics of the network constituents to influence the relative positions network individuals take to each other, Linkletter (2007) and Handcock and Tantrum (2007) suggest representations of the latent social space to dealing with these issues. The different characterizations of the latent space are summarized in the upcoming sections.

The model considered in Linkletter et al. (2006) provides an extension dealing with several drawbacks of the latent space approach suggested by Hoff et al. (2002). The socio economic position of an individual is assumed to depend on socio economic attributes, thereby providing a parsimonious parametrization of the the latent space. This parsimonious parametrization possibly immunizes the model against over fitting the data as in the approach followed by Hoff et al. (2002).

The network is modeled as follows. The likelihood of a observed connection between two actors i and j is assumed to be logistic, hence a latent regression model is used to link dyad specific regressors to observed network relations, i.e.

$$\eta_{ij} = X_{ij}\beta + h_{ij} + e_{ij}.$$

Assuming logistic errors, the likelihood is then

$$\mathcal{L}(Y|X, \beta) = \prod_{i>j} \frac{\exp\{\eta_{ij}\}^{y_{ij}}}{1 + \exp\{\eta_{ij}\}}. \tag{1}$$

For the relative position h_{ij} Linkletter (2007) assumes

$$h_{ij} = -|z(q_i) - z(q_j)|,$$

where $z(\cdot)$ is a function mapping the observed characteristics of social agents $q_i \in \mathfrak{Q}$ to the real line thereby subsuming the observed characteristics within a position in the latent space. It is flexibly parameterized, where parameters are left unspecified a priori, but are simultaneously estimated to specify the transformation $z(\cdot)$. This approach allows to deal most accurately with transformation reflecting the homophily of attributes in the considered network. The consideration of a mapping function $z(\cdot)$ allows an interesting interpretation in the sense of clustering. In regions of the observed space $\mathfrak{Q}$, where $z(\cdot)$ is flat, differences have only little impact on the connection probability. These regions form groups of actors with similar characteristics showing high potential to connect with each other. Conversely, large changes in the functional $z(\cdot)$ indicate regions, where only small changes of large influence on the cooperation probability. These features possibly create a set of boundaries separating different groups of actors. It should be noted that the modeling of the relative position in the latent space as given above can be interpreted as a special case of the approach suggested by Hoff et al. (2002) as it provides a restriction on the latent space, which is absent in Hoff et al. (2002).

Note that for the number of variables p defining the observed space $\mathfrak{Q}$ are assumed to be known. Extensions as given also in Linkletter et al. (2006) are also concerned about selection of adequate variables forming the observed space. Furthermore the variables forming the latent space $\mathfrak{Q}$ are standardized to $[0,1]^p$ via the transformation[4]

$$\tilde{q}_{ih} = \frac{q_{ih} - \min \mathfrak{Q}_h}{\max(\mathfrak{Q}_h) - \min(\mathfrak{Q}_h)}, \quad h = 1, \ldots, p,$$

or alternatively with $q^*_{ih} = \frac{q_{ih} - \frac{1}{n}\sum_{i=1}^n q_{ih}}{\sqrt{\frac{1}{n}\sum_{i=1}^n (q_{ih} - \frac{1}{n}\sum_{i=1}^n q_{ih})^2}}$

$$\tilde{q}_{ih} = \frac{\arctan q^*_{ih} + \pi/2}{\pi}, \quad h = 1, \ldots, p.$$

A Gaussian process is chosen as a prior for the functional $z(\cdot)$. However, instead of defining a Gaussian process for $Z = z(Q) = (z(q_1), \ldots, z(q_n))\prime$ by a mean vector and a covariance matrix causing the number of parameters to increase with order $O(n^2)$, Linkletter et al. (2006) propose to construct a Gaussian process by convolution of a Gaussian white noise process $\alpha(q)$ with a smoothing kernel $k(q)$. Following Higdon (2002) the white noise process $\alpha(q)$ is discretized, which can then be controlled by fewer parameters. The discretized convolution is given by

$$z(q) = \sum_{r=1}^{m} \alpha_r k_\rho(q - w_r),$$

where m denotes the dimensionality of the considered process, weights $\alpha_r = \alpha(w_r)$, and w_r gives the Gaussian white noise process. A specific choice for the kernel $k(\cdot)$, which may influence the results, is to chose an independent p-dimensional Gaussian kernel given by

$$k_\rho(q_i - w_r) = \prod_{d=1}^{p} \rho_d^{(w_{rd} - q_{id})^2}, \tag{2}$$

where $\rho = (\rho_1, \ldots, \rho_p)\prime$ model correlations, and w_{rd} and q_{id} denote the d^{th} element of w_r and q_i respectively. A standard notation of ρ_d is given as

$$\rho_d = e^{-\frac{1}{2\sigma_d}},$$

where σ_d is the standard deviation of the kernel in the d^{th} direction. However, Equation (2) is convenient in terms of MCMC exploration since $0 \leqslant \rho_d \leqslant 1,\ d = 1, \ldots, p$. Interpretation of the parameters ρ_d can be based on the notion of spatial correlation. A large value of ρ_d implies a larger scaling of distance in dimension of variable x_{id}. A large value of ρ_d corresponds therefore to a space surface were a evenly spread distances are more likely to occur, since not only a small regions of space has considerable large value for the kernel k_ρ. In combination with weights $\alpha_r,\ r = 1, \ldots, m$ this representation of latent

[4]This standardization is required for identifications reasons, similarly to the standardization of latent positions in the model-based clustering approach.

space allows very flexible kinds of third order dependencies. Figure (5) illustrates the effect of different values of ρ on the scaling of distances within the latent socio-spatial process. Figure(5) indicates that small changes in the relative position with respect to variables have large effect, with the magnitude of the effect depending on the chosen parameter constellation with parameters nearer to 1 scaling down difference within the socio-spatial space.

Computation of the kernel $k_\rho(\cdot)$ requires realizations of the Gaussian white noise process $W = (w_1, \ldots, w_m)\prime$. These realizations are chosen as a grid with high coverage over the hypercube $[0,1]^p$. The number of points m forming the grid is chosen via a rule of thumb known as $m = 10p$ roles suggesting to use 10 points per dimension in order to provide an accurate approximation. High coverage is achieved using a Latin hypercube design with space filling optimization, see McKay et al. (1979) and Jones et al. (1998) for details.

The random field $z(Q)$ is hence governed via the parameters α and ρ for which the following prior assumptions are used. The prior for alpha is given as

$$\alpha \sim \mathcal{N}(0, I_m)$$

and the prior for ρ is

$$\rho_d \sim \mathcal{U}(0,1), \quad d = 1, \ldots, p.$$

The above outlined socio-spatial process can be used for network implementation and has several advantages over the latent factor model approach. Since the latent position h_{ij} is conceptualized via the functional $z(\cdot)$ only the knowledge of the corresponding variables q_i and q_j is necessary to provide a forecast of the relative latent position. This allows to implement network relations subject to nominated non-response and non-nominated non-response in case where subjective and dyadic attributes are known.

The model frameworks adapted above will be estimated within the Bayesian framework. Therefore the following paragraphs will provide the sampling techniques and MCMC sampling blocks used to perform Bayesian inference in the above given model frameworks.

Markov Chain Monte Carlo (MCMC) sampling is a device to produce a sample from distributions of interest. In a Bayesian analysis, where properties of the joint posterior analysis are of interest, MCMC sampling is used to obtain a sample from this joint posterior distribution. This sample and moments derived thereof serve then as Bayesian estimates. The use of MCMC sampling is necessary, since analytical derivation of moments of the posterior distribution serving as Bayesian estimators via analytical integration is most often hindered via the high complexity of integrals. MCMC sampling is performed via iterative sampling from the full conditional densities of an adequately chosen partition of the parameter vector into blocks. This iterative sampling constitutes a Markov chain, which ensures under general regularity conditions provided in Chib (2001) and not fulfilled only in pathological cases convergence towards the joint posterior distribution.[5] Before describing the employed sampling schemes, some general notes

[5]Following Chib (2001), the transition from $\theta_k^{(r)}$ to $\theta_k^{(r+1)}$ is accomplished via sampling from $p(\theta_k|\theta_1^{(r)}, \ldots, \theta_{k-1}^r, \theta_{k+1}^{(r-1)}, \ldots, \theta_K^{(r-1)}, S)$. The transition of the Markov chain constituting out of K blocks is then described for continuous full conditional distributions as

$$K(\theta^r, \theta^{r+1}) = \prod_{k=1}^{K} p(\theta_k|\theta_1^{(r)}, \ldots, \theta_{k-1}^r, \theta_{k+1}^{(r-1)}, \ldots, \theta_K^{(r-1)}, S).$$

Sufficient conditions for convergence can then be stated as follows. Let $K(\theta, \theta')$ denote the transition density of the Gibbs sampler and let $K^R(\theta_o, \theta')$ be the density of θ' after R iterations of the Gibbs sampler given the initialization θ_0. The

$$\|K^R(\theta_o, \theta') - p(\theta|S)\| \to 0 \text{ as } R \to \infty,$$

where $\|\cdot\|$ denotes the total variance distance. As it is shown by Roberts and Smith (1994), convergence is ensured under the following conditions on the applied sampling algorithms and the data augmentation device will be given.

Metropolis-Hastings algorithm and Data augmentation Chib and Greenberg (1995) provide an extensive review of the different Metropolis-Hasings algorithms, which serve as building blocks in the applied sampling schemes. The Metropolis-Hastings algorithm is a device to produce a sample from a target density f, where the proposal density q is allowed to depend on the current state of the Markov-Chain. A draw from the proposal or jumping distribution y is accepted with probability

$$\alpha = \min \left\{ \frac{f(x)q(x,y)}{f(y),q(y,x)}, 1 \right\},$$

where x denotes the current state of the Markov Chain. The choice of the proposal density q is left to the applied researcher. However, there are several basic types discussed in Chib and Greenberg (1995). Choosing the candidate value $y = x + u$, where u is called increment value, is referred to as random walk Metropolis-Hastings algorithm, since the proposal is given via the current value plus noise. Since this proposal density can be implemented in many situation and provides a fast sampling device, it is a common choice for sampling from full conditional distributions, which are not accessible to direct sampling. However, the use of a Metropolis-Hastings Random Walk algorithm can be problematic, since mixing over the parameter space may be slow and the acceptance rate may be low in some applications. A possibly slow mixing over the parameter space under consideration may be indicated via high persistence within the Gibbs sweeps, i.e. high autocorrelation.[6] Alternatively, the proposal density q can be chosen to be independent of the current state y. Such a chain is labeled as *independent chain* by Tierney (1994).[7]

For each type of candidate density scale and for the independent chain approach also the mean parameter have to be specified. Note that numerical accuracy depends crucially on the choice of these parameters. With respect to scale parameters, their choice affects accuracy in terms of acceptance ratio and coverage of sample space. While choosing a small scale of the proposal density ensures to obtain candidates from a high density region and therefore a high acceptance rate, it causes possibly a poor coverage of the sample space and vice versa. Roberts et al. (1994) discusses some guidelines for choosing the scale parameter in the context of the random walk proposal density. In context of normal proposals and target densities, they argue to use a scaling yielding an acceptance rate ranging from .25 for higher dimensional problems to .50 in one dimensional problems. Müller (1991) also recommends in context of the random walk chain algorithm a scale parameter providing an acceptance ratio around .50. While these choice have shown reasonable performance in applications, they lack consideration of dependencies between elements of parameter vectors. As noted by Geweke (1989) for independent chain samplers importance sampling can be used to construct a proposal density. However, to ensure convergence the importance density must dominate the target density in the tails. Furthermore, all these recommendations are subject to the general caveat that a proposal density with nearly optimal acceptance probability may exhibit excessive autocorrelation within draws. To ensure the validity of estimates in such circumstances alternative families of proposal densities have then to be analyzed. As will illustrated below, (optimized) importance sampling can provide a generic tool to produce proposal densities providing high acceptance ratios and lower autocorrelation within sequences of draws.

1. $p(\theta|S) > 0$ implies there exists an open neighborhood N_θ containing θ and $\xi > 0$ such that, for all $\theta' \in N_\theta$, $p(\theta\prime) \geqslant \xi > 0$;
2. $\int f(\theta) d\theta_k$ is locally bounded for all k, where θ_k is the kth block of parameters;
3. the support of θ is arc connected.

[6] Shephard and Pitt (1997) discuss an alternative measure for numerical accuracy of sampling procedures, see also Liesenfeld and Richard (2006).

[7] Chib and Greenberg (1995) mention further classes of proposal densities such as use of pseudodominating densities, autoregressive chains, and kernel chains.

Data augmentation as proposed by Tanner and Wong (1987) includes latent variables of the model into the parameter vector. In the present context of social network models the latent variables η_{ij} for all i and j are hence included. The joint posterior of the augmented parameter vector is then subject to analysis via MCMC schemes. While this augmentation on the hand complicates the matter of sampling directly from the joint posterior distribution, it is applied when simplifying the matter of sampling from the full conditional distributions. Here the inclusion of the latent variable η_{ij} alters the problem of sampling β in into sampling the parameters from a linear regression model.

Parameters to be estimated are summarized in θ referring to the parameters governing the conditional mean and random field $z(Q)$. To obtain Bayesian estimates of parameters of interest, the posterior distribution has to be calculated. Since the considered model framework does not allow an analytical treatment, estimation is based on MCMC techniques to obtain draws from the posterior distribution and used the empirical moments as estimators. To obtain draws from the posterior distribution its form must be known up to an unknown constant, which is provided via the assumed likelihood and prior functions

$$p(\theta|Y,X) \propto \mathcal{L}(Y|X,\theta)\pi(\theta).$$

The prior $\pi(\cdot)$ has therefore to be specified. For parameters concerning the conditional mean a multivariate normal prior is a straightforward choice. Concerning the random field $z(Q)$ prior assumptions are chosen with respect to $\{\rho_d\}_{d=1}^{p}$ as uniform over interval $[0,1]$ and with respect to α as multivariate normal with zero mean and diagonal unit variance.[8]

Given this setup the posterior distribution for $\theta = \{\mu, \alpha, \rho\}$ can be summarized as

$$\begin{aligned} p(\beta,\alpha,\rho|Y,X) \quad \propto \quad & \prod_{i>j} \frac{\exp\{\eta_{ij}\}^{y_{ij}}}{1+\exp\{\eta_{ij}\}} \prod_{d=1}^{p} I_{(0,1)}(\rho_d) \\ & \exp\{-\frac{1}{2}(\beta-\mu_\beta)'\Omega_\beta^{-1}(\beta-\mu_\beta)\}\exp\{-\frac{1}{2}\sum_{r=1}^{m}\alpha_r^2\}. \end{aligned}$$

Since direct sampling from the posterior is not possible and also the full conditional distribution of the parameter blocks are non standard, draws $\beta^{(s)}, \alpha^{(s)}, \rho^{(s)}$ are obtained via a Metropolis-Hastings algorithm. It has the following structure.

- Given an initialization $\beta^{(0)}, \alpha^{(0)}, \rho^{(0)}$ with $p(\beta^{(0)}, \alpha^{(0)}, \rho^{(0)}|Y,X) > 0$ generate new draws for the parameters as follows. Repeat for $s = 1, \ldots, S$.

  1. Draw a candidate value β^* for a known symmetric distribution $f(\beta^*|\beta^{(s-1)})$ given the previous value $\beta^{(s-1)}$. This distribution is called *jumping distribution*, see Gelman et al. (1995) for a complete discussion of the Metropolis-Hastings algorithm including regularity conditions. Proposal draws are obtained from a multivariate normal with mean $\beta^{(s-1)}$ and a diagonal covariance. Note that Hoff et al. (2002) use in the scalar case a uniform jumping distribution centered around $\beta^{(s-1)}$. Alternatively, the mean of the proposal distribution is identified as the solution of the following maximization problem, i.e.

$$\beta^* = \arg\max_\beta \prod_{i>j} \frac{\exp\{\eta_{ij}\}^{y_{ij}}}{1+\exp\{\eta_{ij}\}} \exp\{-\frac{1}{2}(\beta-\mu_\beta)'\Omega_\beta^{-1}(\beta-\mu_\beta)\}.$$

[8] Alternatively, concerning prior beliefs on ρ George and McCulloch (1993) and Clyde (1999) discuss a mixture prior of the form

$$\pi(\rho_d) = \gamma I_{[0,1]}(\rho_d) + (1-\gamma) I_{\{1\}},$$

where γ allows to incorporate a prior belief on the fraction on inactive factors.

The covariance is then provided as the corresponding inverse Hessian. This proposal density provides higher acceptance ratios within the Metropolis-Hastings algorithm and hence enhances numerical accuracy of the algorithm.

2. Calculate the ratio

$$\begin{aligned} r &= \frac{p(\beta^*, \alpha^{(s-1)}, \rho^{(s-1)}|Y, X)/f(\beta^*|\beta^{(s-1)})}{p(\beta^{(s-1)}, \alpha^{(s-1)}, \rho^{(s-1)}|Y, X)/f(\beta^{(s-1)}|\beta^*)} \\ &= \frac{p(\beta^*, \alpha^{(s-1)}, \rho^{(s-1)}|Y, X)f(\beta^{(s-1)}|\beta^*)}{p(\beta^{(s-1)}, \alpha^{(s-1)}, \rho^{(s-1)}|Y, X)f(\beta^*|\beta^{(s-1)})} \end{aligned}$$

3. Set

$$\beta^{(s)} \begin{cases} \beta^*, & \text{with probability } \min(r, 1), \\ \beta^{(s-1)}, & \text{otherwise .} \end{cases}$$

- Apply the Metropolis-Hastings scheme outlined above for α. Use as a *jumping distribution* for each element of α a uniform distribution centered around the previous draw. Also the use of a normal distribution with mean given by the previous draw and diagonal covariance matrix is possible. Thereby the variance is varied between values between 1 and 0.1.
- Apply the Metropolis-Hastings scheme outlined above for ρ. Again a natural choice for the *jumping distribution* for each element of ρ a uniform distribution centered around the previous draw, where the range is $\pm\lambda_\rho$ varying from 0.15 to 0.05.

In order to ensure mixing across the parameter space, the number of draws S has to be set sufficiently large. To enhance numerical efficiency, the MH-steps are iterated until a proposal is accepted, which improves also mixing across the parameter space. Furthermore, since the sequences of draws for the parameters are autocorrelated, only a fraction of draws is used to obtain estimates, e.g. each 5^{th} or 10^{th} draw.

The parameter estimates are hence provided by the sampler averages of the drawn sequences, i.e.

$$\hat{\beta} = \frac{1}{S'}\sum_{s=1}^{S'} \beta^{(s)}, \qquad \hat{\alpha} = \frac{1}{S'}\sum_{s=1}^{S'} \alpha^{(s)} \quad \text{and} \quad \hat{\rho} = \frac{1}{S'}\sum_{s=1}^{S'} \rho^{(s)}.$$

Assessing the estimated latent position of actors is possible via

$$\hat{z}(q) = E[z(q)|Y, X] = \frac{1}{S'}\sum_{s=1}^{S'}\sum_{r=1}^{m} \alpha_r^{(s)} k_{\rho^{(s)}}(q - w_r),$$

what allows also to give the posterior probability of an observed network connection

$$\hat{\pi}_{ij} = E\left[\frac{\exp\{\eta_{ij}\}}{1 + \exp\{\eta_{ij}\}}|Y, X\right] = \frac{1}{S'}\sum_{s=1}^{S'} \frac{\exp\{\eta_{ij}^{(s)}\}}{1 + \exp\{\eta_{ij}^{(s)}\}},$$

where

$$\eta_{ij}^{(s)} = X_{ij}\beta^{(s)} - |z(q_i)^{(s)} - z(q_j)^{(s)}|.$$

The adequacy of the whole model setup can be assessed via cross validation techniques. Splitting the observed network into a training and a validation sample allows to judge, whether the estimated parameter and latent space parameter provide a valid representation of the network under investigation. The next section elaborates techniques on this specific issue of model selection.

In the following approaches suggested within the literature for issues of model specification and selection are considered and adapted to the features of the suggested model frameworks, which are not subject to posterior inference. This features are the number of mixing components, the dimension of the latent space, and the matter of what variables to choose for valid characterization of the latent socio-spatial space process.

Selection of Variables in the Socio-Spatial Process Model In the above outlined model, the variables shaping the latent socio-spatial process have been assumed as given and known. In empirical analysis however, although one might have good theoretical knowledge about variables influence the individual position in the latent space, the set of variables should be also subject to analysis and testing.

The problem of variable selection corresponds to choosing the best model for the observed network. A fully Bayesian analysis would investigate all 2^p possible subsets of variables according to the marginal likelihood criteria, which is potentially computationally costly, see Chipman et al. (2001) for a discussion of related issues. Instead, Linkletter et al. (2006) propose to judge the significance of variables based on the posterior estimates of the parameters ρ_k. The more different ρ_k is from 1, the less likely it is not a determining factor of the latent space. One important caveat is that in the outlined approach of Linkletter et al. (2006) the network response variable is considered to be metric and not binary and needs hence to be adapted to binary network data.[9]

The importance of single variables is judged on the basis of a comparison of the influence of an artificial reference variable and the set of variables under consideration. Comparing the influence of an artificial variable with the considered set of variables is advantageous over direct gauging the posterior distributions, since it provides a reference scenario for the behavior of an insignificant variable. A drawback of this reference distribution approach is an increase in the computational burden for construction of the reference distribution. The procedure can be summarized by the following algorithm.

1. Augment the set of variables Q by creating a reference variable Q_{p+1} with no significant influence. The creation is based on random sampling from the covariate space of the original variables.

2. Compute the posterior mean of $\tilde{\rho}_{p+1}$ of the added reference variable.

3. Repeat steps 1 and 2 M times. Use the obtained median estimates $\tilde{\rho}_{p+1}^{(m)},\ m = 1, \ldots, M$ to construct a reference distribution representing the situation of no influence.

4. Compare the estimates medians $\tilde{\rho}_d,\ d = 1, \ldots, p$ to the reference distribution to assess their significance.

The adequateness of the outlined approach is assessed in Linkletter et al. (2006) in various simulation experiments. The results show that the variable selection via reference distribution is suited to identify active variables influencing the shape of the latent socio-spatial process, see also the simulation study conducted in Chapter 4.

Model Comparison via Cross Validation Model fitness is assessed via cross-validation techniques which allow to compare also non-nested model setups. Cross-validation techniques split the observed network into a training sample employed for estimation and a forecasting sample. Based on estimates obtained from the training sample the prediction probabilities are calculated for the forecasting sample. Based in the forecasting probabilities from different models several fit criteria can be inspected and judgement on the best model specification made. Alternatively to approximating the marginal likelihood

[9] It will therefore be of special interest, whether the outline approach can be applied on dichotomous network data as well.

via the BIC, Hoff (2005) suggests for identification of the correct dimension of the latent space to assess the predictive performance of a model. In order to allow comparison across different models and different dimension of the latent space, he recommends the use of L-fold cross validation. The procedure can be summarized as follows:

1. Provide a partition of pairs of actors $\{(i,j) : i \neq j\}$ into L sets $A_1, \ldots, A_L$.
2. For different numbers of dimensions or different models
   (a) For each $l = 1, \ldots, L$
      i. perform estimation via the MCMC algorithm using only pairs $\{(i,j) \notin A_l\}$, but sample values y^*_{ij} for all pairs.
      ii. based on sampled value for y^*_{ij} compute posterior means $y^*_{ij} \in A_l$ and the log predictive probability $\text{lpp}(A_l) = \sum_{i,j \in A_l} \log p(y_{ij}|y^*_{ij})$.

   (b) Assess the predictive performance for a specific dimension of model setup via $LPP = \sum_{l=1}^{L} \text{lpp}(A_l)$.

This approach can also be extended for selection of the number of components in the model-based clustering approach via running the additional numbers runs for the considered selection dimension.

4 Empirical Results

This section presents the empirical estimation results. As a baseline model a standard probit model is estimated including only the dyad specific regressors, i.e. number of common organizational memberships and political conflict potential between individuals. Estimation of the standard probit model is performed via maximum likelihood. Estimation results for the standard probit and socio-spatial process model are presented in Table (2) for the counties of Budkovce (upper part) and Chotza (lower part), while Table (3) provides estimation results for the counties of Kamienienc and Porchovany.

Starting with the Slovakian county of Budkovce, standard probit regression analysis and socio-spatial model estimates underline the importance of the number of common organizational memberships as a dyad specific regressor for individual communication. A relevant factor shaping the relative distance in the socio-spatial space between individuals is political party membership, where the importance of this factor is highlighted by the reference distribution approach shown in the upper panel of Figure (5). Also personal reputation is documented to be among the variables showing significant impact on the socio-spatial distance.

For the Polish municipality of Chotza difference occur compared to Budkovce. On the one hand, political conflict potential is the significant dyad specific regressors, and not number of common organizational memberships. Furthermore, the estimation results suggest a positive link between political conflict potential and communication, i.e. with a higher political conflict potential making communication more likely. Furthermore, the reference distribution approach, see second panel of Figure (5), indicates the irrelevance of of individual specific regressors therefore of distances in the socio-spatial space for individual communication. Note that, descriptive statistics shown in Table (1) document a younger local elite on average compared with the other three municipalities and a lower educational level. Possibly, this structural difference are reflected in different structures of individual communication highlighted via the empirical analysis.

Within the municipality of Kamienienc, see for estimation results upper part of Table (3), also the number of common organization is the most influential dyad specific regressor. However, with respect to individual specific factors influencing the distance between local elite members within the latent social space, three variables namely educational level, political party membership, and personal reputation are documented to exhibit substantial influence, see Figure (5), third panel. Hence, personal communication within the county of Kamienienc is compared to the three other counties in Poland and Slovakia at most determined by personal characteristics of local elite members.

Concerning the fourth municipality of Porchovany, estimation results are provided in Table (3). For this county, the number of common organizations is the influential determinant of communication ties, next to the political party membership shaping the distance within the socio-spatial process. With respect to individual specific factors showing influence on the distance within the latent social space, no variables are found to have substantial influence, compare Figure (5), lowest panel.

Next to analysis of individual specific communication, a short note shall be made on the overall level of communication ties, individuals have within the local elite network. Therefore, a count data analysis is performed using a negative binomial setup.[10] Estimation is performed via maximum likelihood and estimation results are shown in Table (4). The empirical analysis reveals the importance of personal reputation for the overall number of ties within the social communication network. Furthermore, for the municipality of Chotza, with the youngest average local elite, a higher age significantly increases the overall number of communication ties.

5 Conclusion

This article analyzed social elite communication networks in four municipalities in Slovakia and Poland. The network structure and hence the dyadic nature of the dependent variable make empirical model structures necessary, which allow a flexible yet parsimoniously parameterized aggregation of individual specific information on the dyadic level of the binary dependent variable. This paper adapts a Gaussian field process for this purpose and shortly discusses alternative modeling approaches suggested and implemented within the literature. Estimation is performed using a Bayesian approach, which is implemented using MCMC techniques.

Empirical analysis reveals the importance of common organizational membership and the political party membership in explaining the occurrence of communication relationship between local elite members. Some difference between the municipalities were revealed for the county having the youngest local elite, where the political conflict potential between enhances the probability of communication between individuals. With respect to the overall activity in communication networks the poisson regression results document the importance of individual reputation.

Thus, this paper applied a flexible framework for aggregation of individual specific information on dyadic levels and provided an analysis of determinants of individual communication in local elite networks. Future research could aim at a formal comparison of alternative approach incorporating latent individual heterogeneity.

[10] Poisson regression results point at similar conclusion, however overdispersion is significantly estimated.

Table 1: Network and Variable Summary Statistics

	I	*II*	*II*	*IV*
	network statistics			
network density	0.2463	0.3126	0.3016	0.3597
# network constituents	29	37	36	28
	variable statistics (mean/std)			
# joint organizations	0.4754 / 0.7782	0.0736 / 0.2725	0.3571 / 0.5869	0.4229 / 0.6836
political conflict index	1.4884 / 0.8271	1.5362 / 0.9128	1.7376 / 1.0546	1.5459 / 1.0912
age (in years)	54.8929 / 11.8082	45.53571 / 11.10383	55.5926 / 9.5041	50.72727 / 10.20822
educational level	7.4286 / 2.4859	4.142857 / 1.693436	4.1481 / 1.7030	8.318182 / 1.783158
political party membership	0.2857 / 0.4600	.1428571 / .3563483	0.2593 / 0.4466	.2727273 / .4558423
personal reputation	0.2735 / 0.1955	.4942085 / .2547981	0.4115 / 0.2228	.3557312 / .2195079
social prestige index	56.1034 / 16.4758	50.28571 / 13.52599	49.7407 / 11.8924	60.95455 / 12.72222

Table 2: Estimation Results – (1)

	I-probit model			*II*-socio-spatial process model			
	mean	sd	90% HDI	mean	sd	90% HDI	*CD*
Bukovce							
dyadic specific regressors							
constant	-.7098	.1600	[−.9729; −.4466]	-.6497	.3464	[−1.1891; −.0617]	1.1692
# of joint organizations	.3051	.0840	[.1670; .4433]	.5151	.1520	[0.2709; .7681]	0.2652
political conflict index	-.0958	.0919	[−.2470; .0554]	-.1691	.1511	[−0.4191; .0783]	-2.5937
individual specific regressors							
age	–	–	–	.5928	.2871	[0.1207; .9829]	1.5515
educational level	–	–	–	.6230	.2794	[0.1442; .9654]	0.2971
political party membership	–	–	–	.2808	.1949	[0.0195; .6589]	1.7521
personal reputation	–	–	–	.2976	.2776	[0.0075; .7881]	1.1317
social prestige index	–	–	–	.8591	.1006	[0.6766; .9869]	1.4001
	mean	sd	90% HDI	mean	sd	90% HDI	*CD*
Chotza							
dyadic specific regressors							
constant	-.7148	.1026	[−.9159; −.5138]	-.5297	.3870	[−1.1145; .1623]	-1.7306
# of joint organizations	.2461	.1843	[−.1151; .6073]	.3469	.3076	[−.1650; .8525]	-0.1057
political conflict index	.1430	.0558	[.0337; .2523]	.2681	.1020	[.1051; .4388]	-1.1959
individual specific regressors							
age	–	–	–	.7201	.1142	[.5408; .8854]	-0.3049
educational level	–	–	–	.9382	.1089	[.6790; .9984]	-0.8316
political party membership	–	–	–	.7979	.1596	[.4838; .9846]	0.2325
personal reputation	–	–	–	.7033	.1948	[.3712; .9669]	-1.8311
social prestige index	–	–	–	.7991	.2485	[.3171; .9887]	2.6536

Notes: Estimation is based on 20000 MCMC draws, where initial 10000 were discarded for burn-in.

Table 3: Estimation Results – (2)

	I-probit model			*II*-socio-spatial process model			
	mean	sd	90% HDI	mean	sd	90% HDI	*CD*
Kamienienc							
dyadic specific regressors							
constant	-.7118	.1102	[−.9277; −.4958]	-.6620	.2731	[−1.1052; −.2089]	0.5861
# of joint organizations	.4921	.0906	[.3146; .6696]	.8487	.1624	[.5830; 1.1187]	-0.9062
political conflict index	-.0007	.0514	[−.1015; .1001]	.0614	.0928	[−.0916; .2123]	1.8179
individual specific regressors							
age	–	–	–	.6517	.2089	[.2686; .9452]	4.7667
educational level	–	–	–	.3916	.1385	[.1711; .6336]	0.5180
political party membership	–	–	–	.1638	.1630	[.0012; .5351]	-1.9512
personal reputation	–	–	–	.3384	.2959	[.0335; .9259]	-0.5492
social prestige index	–	–	–	.8732	.1071	[.6626; .9876]	0.0825
Porchovany							
dyadic specific regressors							
constant	-.5789	.1631	[−.8986; −.2592]	-.6026	.3571	[−1.1360; .0125]	-1.1585
# of joint organizations	.6614	.1259	[.4147; .9082]	1.0563	.2203	[.7022; 1.4186]	1.2682
political conflict index	-.0497	.0809	[−.2083; .1089]	-.1088	.1339	[−.3313; .1074]	2.6365
individual specific regressors							
age	–	–	–	.7564	.1850	[.3843; .9758]	1.0785
educational level	–	–	–	.4124	.2523	[.0603; .8237]	4.5914
political party membership	–	–	–	.3423	.2964	[.0098; .8848]	-1.1251
personal reputation	–	–	–	.5193	.3113	[.0402; .9516]	-1.4956
social prestige index	–	–	–	.7810	.2243	[.2904; .9846]	-1.5541

Notes: Estimation is based on 20000 MCMC draws, where initial 10000 were discarded for burn-in.

Table 4: Analysis of Communication Activity - Negative Binomial Regression

	mean	sd	95% CI
	Budkovce – (SVK)		
constant	1.541979	.8000722	[−.0261341; 3.110091]
age	-.0105996	.0114773	[−.0330947; .0118954]
educational level	.1121329	.0622844	[−.0099422; .2342081]
political party membership	.2945179	.257805	[−.2107707; .7998064]
personal reputation	1.886588	.4564618	[.9919391; 2.781236]
social prestige index	-.0104323	.0090419	[−.0281541; .0072894]
	Chotcza – (PL)		
constant	-.1571744	.7074138	[−1.54368; 1.229331]
age	.0199043	.0100786	[.0001506; .039658]
educational level	.0080713	.0607763	[−.1110481; .1271907]
political party membership	.1162344	.3237661	[−.5183355; .7508043]
personal reputation	2.641798	.471151	[1.718359; 3.565237]
social prestige index	.0017585	.0072925	[−.0125344; .0160514]
	Kamieniec – (PL)		
constant	2.772421	.7966183	[1.211078; 4.333764]
age	-.0117793	.0094905	[−.0303803; .0068217]
educational level	-.0174434	.0681389	[−.1509932; .1161065]
political party membership	-.2415409	.2176573	[−.6681414; .1850597]
personal reputation	1.678827	.4521789	[.7925729; 2.565082]
social prestige index	-.0091044	.007929	[−.024645; .0064362]
	Parchovany – (SVK)		
constant	.6724824	1.458618	[−2.186357; 3.531321]
age	.0031036	.0194798	[−.0350761; .0412834]
educational level	-.1094293	.0819881	[−.2701229; .0512644]
political party membership	.2205007	.3014767	[−.3703828; .8113841]
personal reputation	1.281929	.68248	[−.0557072; 2.619565]
social prestige index	.0251362	.0160243	[−.0062709; .0565433]

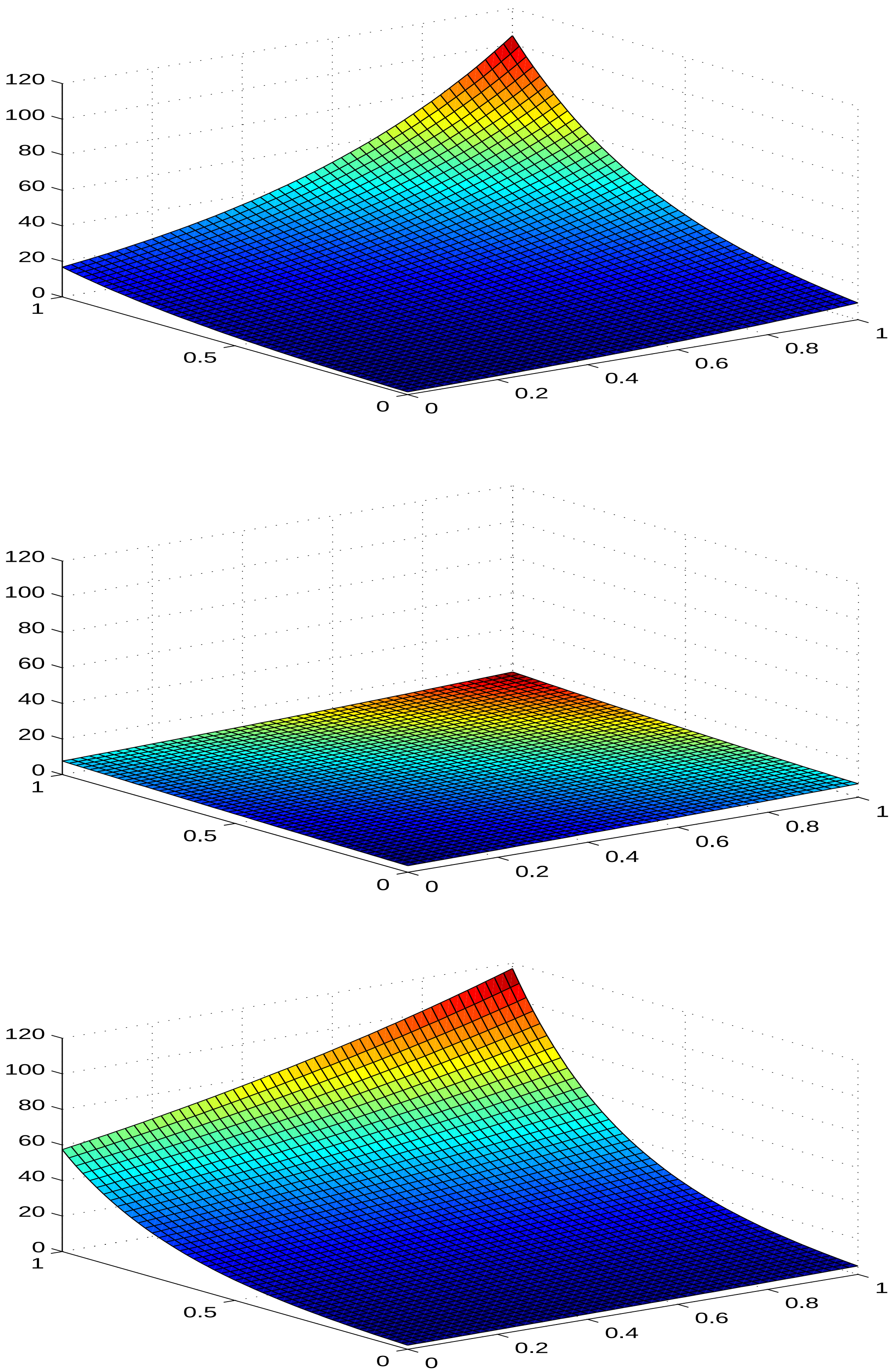

Figure 1: Socio Spatial Process with $\rho = (.3 \quad .4)$ (right panel) / $\rho = (.7 \quad .7)$ (left panel) / $\rho = (.2 \quad .7)$ (lower panel)

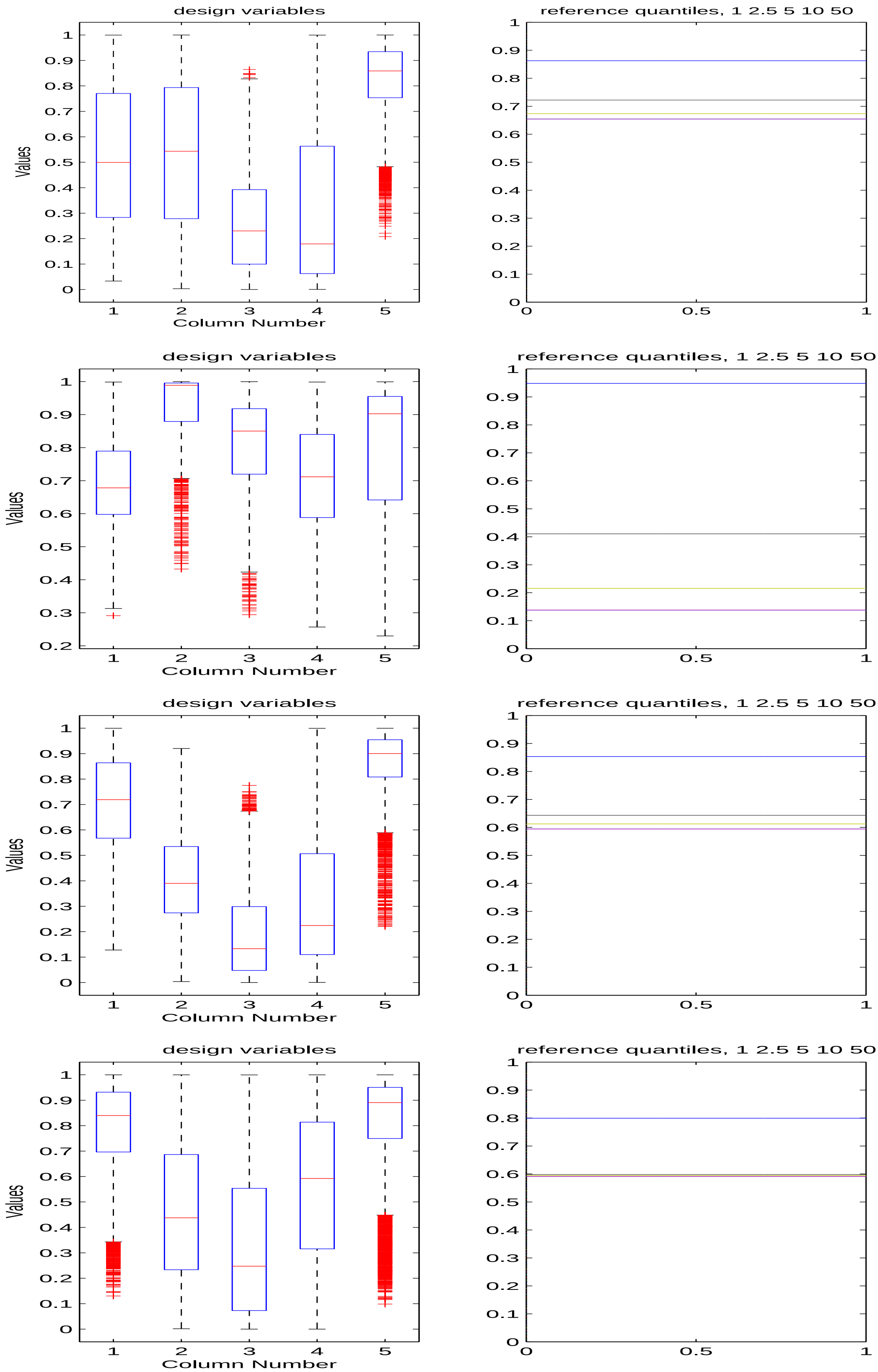

Figure 2: Variable Selection via Reference Distribution for four analyzed counties – Budkovce, Chotza, Kamienienc, Porchovany

References

Bala, Venkatesh and Sanjeev Goyal, "Learning from Neighbours," *Review of Economic Studies*, July 1998, *65* (3), 595–621. available at http://ideas.repec.org/a/bla/restud/v65y1998i3p595-621.html.

Battiston, S., E. Bonabeau, and G. Weisbuch, "Impact of Corporate Boards Interlock on the Decision Making Dynamics," in M. Gallegati, A. P. Kirman, and M. Marsili, eds., *The Complex Dynamics of Economic Interaction. Essays in Economics and Econophysics*, Berlin [u.a.]: Springer, 2004, pp. 355–378.

Celen, B., S. Kariv, and A. Schotter, "Learning in Networks: An Experimental Study," *mimeo*, 2004.

Chib, S. and E. Greenberg, "Understanding the Metropolis-Hastings Algorithm," *The American Statistician*, November 1995, *49* (4), 327–335.

Chib, Siddhartha, *Markov Chain Monte Carlo Methods: Computation and Inference*, Vol. 5 of *Handbook of Econometrics*, Elsevier Science B.V,

Chipman, H.A., E.I. George, and R.E. McCulloch, "The practical implementation of Bayesian model selection," *Institute for Mathematical Statistics Lecture Notes-Monograph Series*, 2001, *38*, 65–116.

Choi, S., D. Gale, and S. Kariv, "Learning in Networks: An Experimental Study," *mimeo*, 2004.

Clyde, M., *Bayesian Statistics*, Oxford University Press,

Coleman, James, "Comment on 'The concept of influence'," *Public Opinion Quarterly*, 1963, *27*, 63–82.

Currarini, S., M.O. JAckson, and P. Pin, "An Economic Model Of Friendship: Homophily, Minorities, and Segregation," *Econometrica*, July 2009, *77* (4), 1003–1045.

Gale, D. and S. Kariv, "Bayesian Learning in Social Networks," *Games and Economic Behavior*, 2003, *45* (2), 329–346.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin, *Bayesian Data Analysis*, London: Chapman & Hall, 1995.

George, E.I. and R.E. McCulloch, "Variable Selection via Gibbs Sampling," *Journal of the American Statistical Association*, September 1993, *88* (423), 881–889.

Geweke, J., "Bayesian inference in econometric models using Monte Carlo integration," *Econometrica*, 1989, *57*, 1317–1340.

Handcock, M.S., "Assessing degeneracy in statistical models of social networks," *Working Paper, Center for Statistics and the Social Science*, 2003, (39).

Henning, Christian H. C. A., "Networks of Power in the CAP System of the EU-15 and EU-27," *Journal of Public Policy*, 2009, *29* (Special Issue 02), 153–177.

_ **and Volker Saggau**, "Information Networks and Knowledge Spillovers: Simulations in an Agent-based Model Framework„" in Neri Salvadori, ed., *Institutional and social dynamics of growth and distribution*, Edward Elgar, 2010.

Higdon, D.M., *Quantitative methods for current environmental issues*, London: Springer Verlag,

Hoff, P.D., "Bilinear Mixed-Effects Models for Dyadic Data," *Journal of the American Statistical Society*, March 2005, *100* (469), 286–295.

_ **, A.E. Raftery, and M.S. Handcock**, "Latent Space Approaches to Social Network Analysis," *Journal of the American Statistical Society*, December 2002, *97* (460), 1090–1098.

Hoff, Peter D., "Model Averaging and Dimension Selection for the Singular Value Decomposition," *Journal of the American Statistical Association*, June 2007, *102* (478), 674–685.

Jones, D.R., M. Schonlau, and W.J. Welch, "Efficient global optimization of expensive black-box functions," *Journal of Global Optimization*, 1998, *13*, 455–492.

Knoke, David, Franz Urban Pappi, Jeffrey Broadbent, and Yutaka Tsujinaka, *Comparing Policy Networks. Labor Politics in the U.S., Germany, and Japan*, Cambridge Univ. Press, Cambridge, 1996.

Krause, Andreas, *Herding Behavior of Financial Analysis: A Model of Self-Organized Criticality*, Springer, Berlin,

Laumann, Edward O. and David Knoke, *The Organizational State* 1987.

Liesenfeld, R. and J.-F. Richard, "Classical and Bayesian Analysis of Univariate and Multivariate Stocastic Volatility Models," *Econometric Reviews*, 2006, *25* (2-3), 335–360.

Linkletter, C., D. Bingham, N. Hengartner, and D. Hidgon, "Variable Selection for Gaussian Process Models in Computer Experiments," *Technometrics*, November 2006, *48* (4), 478–490.

Linkletter, Crystal D., "Spatial Process Models for Social Network Analysis." PhD dissertation, Simon Fraser University 2007.

M., Raftery A.E. Handcock and J.M. Tantrum, "Model-based clustering for soical networks," *Journal of the Royal Statistical Society A*, 2007, *170* (2), 301–354.

McKay, M.D., R.J. Beckman, and W.J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, 1979, *21*, 239–245.

Müller, Peter, "A generic approach to posterior integration and Gibbs sampling," Technical Report 91-09, Purdue University 1991.

Moody, James, "The Structure of a Social Science Collaboration Network: Disciplinary Cohesion from 1963 to 1999," *American Sociological Review*, 2004, *69* (2), 213–238.

Parsons, Talcott, "On the concept of influence," *Public Opinion Quarterly*, 1963, *27* (1), 37–62.

Roberts, G.O. and A.F.M. Smith, "Some simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms," *Stochastic Processes and its Applications*, 1994, *49*, 207–216.

Shephard, N. and M.K. Pitt, "Likelihood Analysis of non-Gaussian measurement time series," *Biometrika*, September 1997, *84* (3), 653–667.

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

Tanner, M. and W. Wong, "The calculation of posterior distributions by data augmentation," *Journal of the American Statistical Association*, June 1987, *82* (398), 528–540.

Tierney, L., "Markov chains for exploring posterior distributions," *Annals of Statistic*, 1994, *22* (4), 1701–1728.

van der Gaag, M., "Measurement of Individual Social Capital." PhD dissertation, Rijksuniversiteit Groningen 2005.