www.ees.uni.opole.pl
ISSN paper version 1642-2597
ISSN electronic version 2081-8319
Economic and Environmental Studies
Vol. 15, No. 2 (34/2015), 231-247, June 2015

An evaluation of simulated network in the hierarchical spatial autoregressive model. Homophily based interactions matter for the consumption behaviours

Edyta ŁASZKIEWICZ
University of Lodz, Poland

Abstract: The aim of the paper is an examination wheatear the social spillover of consumption behaviours differs for the homophily and heterophily based structure of interactions. As an approximation of agents' behaviours households' expenditures with the reference person attributes were used. The simulated structure of networks with individual density, calculated using the data from Polish Social Diagnosis 2011, was applied. Three versions of matching agents in the adjacency matrix was tested: 1) non-random with the k-nearest neighbours algorithm and closeness measured by the Jaccard index, 2) random with five different spatial conditions, 3) partially random with random drawing from the limited number of individuals selected using the value of the Jaccard index. The difference in the social spillover was tested as the difference between the estimated parameters for social interactions from the hierarchical spatial autoregressive model (HSAR). An additional factors as: spatial heterogeneity and dependence, household attributes, ect. was controlled in the HSAR model. The microdata from the 2011 Polish Households Budget Survey was used in this research. The results prove that homophilously structured social networks supported the spillover of healthy food consumption among Polish households in 2011. In contrast, for heterophily based relations spreading of habits did not occur. The crucial role for social spillover plays the geographical proximity of households.

Keywords: consumption behaviour, networks, interpersonal similarity, homophily, spatial multilevel modelling
JEL codes: D12, C21, C49.

1. Introduction

There are three different mechanisms explaining the similarities between people connected in the social networks: confounding, social contagion and homophily (Żak and Zbieg, 2012). The

Correspondence Address: Edyta Łaszkiewicz, Department of Spatial Econometrics, Faculty of Economics and Sociology, University of Lodz, Poland. Rewolucji 1905 Street 37, 90-214 Lodz, Poland. E-mail: elaszkiewicz@uni.lodz.pl.

spreading of social impact in the network is defined as the social contagion, while the confounding is described as a simultaneous effect on individuals due to an external factor existence. Finally, the similarities in the behaviours might occur because of the tendency to associate with similar individuals (what affects the structure of relationships). Such mechanism is known as the homophily.

The homophily based interactions means that agents are more likely to interact with those who have similar or the same characteristics (e.g. gender, age, socio-economic status, workplace). As the human behaviour depends on what others do, we might expect the more similar agents are, the more they are susceptible for each other's behaviour. In the case of the consumption decisions the spillover of behaviour might results an appearance of the new product consumption as well as a changing of expenditures for the present goods consumption. Hence, the possible way of evaluating the importance of the homophily based interactions would be comparing the growth of expenditures which corresponds with the effect of social interactions for different structure of relationships. The predominant role of the homophily based interactions might be proved when the social interaction effect is higher for such structure of relations than for the other, non-homophily.

The aim of this research is to examine the impact of the homophily based interactions on consumption decisions. Different structures of relationships were compared to check whether the homophily improves the spillover of the consumption behaviours. The Polish Household Budget Survey microdata for the year 2011 was taken for this purpose. The influence of the homophily based interactions was tested for households' expenditure behaviour. Moreover, information about the current state of Polish society from the Social Diagnosis (e.g. the number of friends with who's Poles have typically contact and the well-connected friends who live in the same place) were used to specify more detail the potential structures of the relationships.

The rest of the paper is organized as follows. In Section 2 the role of the homophily based interactions in the social network analysis was described. Section 3 is for a basic description of the databases, while in Section 4 the hierarchical spatial autoregressive model was introduced. In Section 5 the description of the approach to generating the network structures was provided. After this, the empirical results for the healthy food consumption models were presented and discussed. Finally, the conclusion follows.

2. Homophily based interactions

In the social network analysis, it is common to assume that connections between people form the network structures and beliefs, opinions, behaviours of a network members are influenced by the others' beliefs, etc. (Wasserman and Galaskiewicz, 1994: xiii). Such social impact was widely proved for the smoking and drinking habits (Collins et al., 1985; Rosenquist et al., 2010), purchases of automobiles (Grinblatt et al., 2010) and food consumption behaviours (e.g. Sorensen et al., 2007), among others.

The crucial role in evaluation the social influence plays the specification of relationships in the network. As a different mechanism of linking people might be applied, the structure of relationships varies. That might affects the empirical results as well as conclusions, which differs due to the applied networks. Because the connections between agents are associated with the existence of interpersonal influence, conceptualization of the network needs to answer for the question about the mechanism which constitutes the interpersonal relations.

One of the most natural way of matching agents is due to the social proximity, like i.a. common culture, ethnic background or socio-economic status (see e.g. Rogers and Bhowmik, 1970). It is because people generally link with those who are like themselves in a free-choice situation (McPherson et al., 2001). Hence, the connectedness based on the social proximity follows the more general assumption which is the existence of homophily. According to the homophily rule, the more similar people are, the higher is the probability that they interact. Social proximity, calculated mostly as a social distance constitutes the elements of the social network (adjacency) matrix. Each pair of agents are connected if they share the common attributes. The more similar features have agents, the lower is the social distance between them and (by using the inverse function of distance) the more connected they are.

The approximation of homophily by the social (also geographical) proximity has a long history in the empirical researches. In one of the earliest the similarities in sex, age, education level, race, ethnic background, occupation (Bott, 1928; Loomis, 1946) as well as psychological closeness (Richardson, 1940) were studied. After this, Lazarsfeld and Merton (1954) provided a division into status and value homophily. The first term is for the similarities in the socio-demographic profile, while the second one refers to the common beliefs, aspirations, ect. Alternatively, McPherson et al. (2001) proposed to separate the situation when the opportunities for forming relations are limited toward the similar others (structural homophily) from the choices made under condition of unlimited possibilities of linking (preference homophily).

Following the similarity-attraction paradigm which stays behind the homophily effect, linking with similar others might guarantee reducing the possibility of conflicts in a group or improves a self-affirmation. The results from the Byrne and Nelson (1965) experiment showed the linear relation between the proportion of attitudinal items and the degree of interpersonal attraction. Hence, we might expect the more homophily is the network structure, the more consumer is affected by others' behaviours.

However, consumer might be inspired not only by the similar others but also by those who has the complementary traits (heterophily effect). Different others might be attractive because they have those characteristics which are gratified by agent. In the context of consumption behaviour, heterophily might improves spillover of habits when consumers (agent) would like to be as dissimilar others. The possibility of the interaction which drives the spillover is then due to the attractiveness of unique attributes for both sides of relation.

In the large body of researches, homophily was examined as a factor which influences i.a.: friendship formation (de Marti and Zenou, 2011) and opinion formation (Centola et al., 2005), speed of learning (Golub and Jackson, 2012), diffusion dynamics over social network (Yavaş and Yücel, 2014), adoption of health behavior (Christakis and Fowler, 2007; Centola, 2011). Less attention has been put into the impact of homophily based interactions on spreading the consumption habits and the potential supporting the spillover of behaviours by heterophilous connections. Despite, we might expect that while the social interactions modifies the food consumption habits, some forms of interactions improve the social spillover of behaviours more.

3. Databases

In this study two databases with a microdata from the 2011 were used: Polish Households Budget Survey (conducted by the Central Statistical Office) and Social Diagnosis (Social Monitoring Council and a team appointed by a Council of experts). Let us now provide a basic information about both of them.

3.1. Household Budget Survey

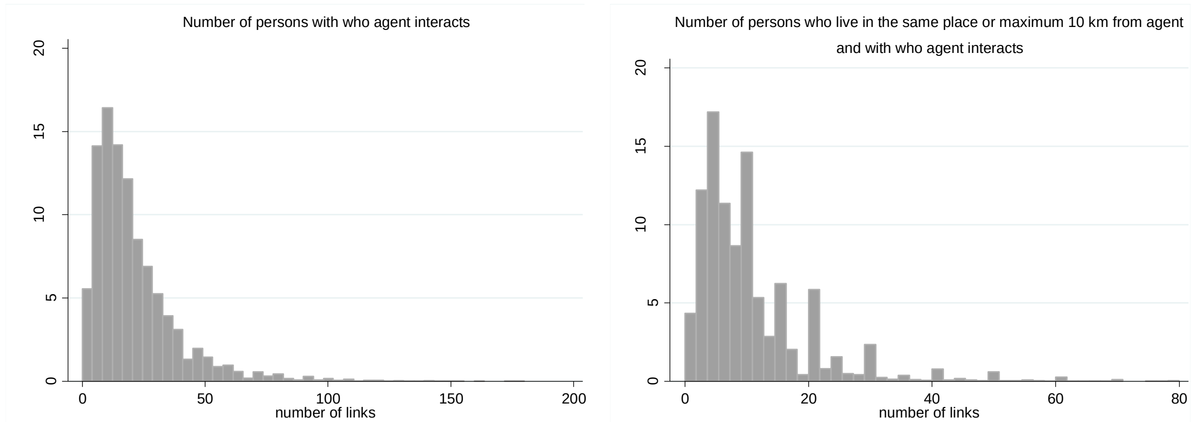
From the full sample consists of 37 375 households only those were taken who declared non-zero expenditures (16 households were eliminated) for three category of food products which might be jointly named as a healthy food: fruits, vegetables and fish. An area survey point (1 551 spatial units) was used to defines a community level. Additional, the information about the assignation of households to the voivodship was taken into account.

According to the database, in 2011 the mean monthly expenditures for the healthy food were equal to 140 PLN (5.38% of the household income). However, there was a high difference between households in the expenditures for above products (standard deviation was equal to 92.58 PLN). The highest mean expenditures for the healthy food was found in the biggest Polish cities with the population over 500 thousand (145.65 PLN). Also the significant differences between voivodships were observed (see Łaszkiewicz et al., 2014 for the more detailed statistics).

3.2. Social Diagnosis

Two variables with the number of persons with who agent interacts were taken from the Social Diagnosis (2011) database to specify the individual density of network. The distribution of both, calculated for 26 249 respondents, were shown in Figure 1. The mean density of network was equal to 21 persons, while the standard deviation was high and equal to 17 persons. In addition, the mean number of persons with who agents interacts and who live in the same (or close) place was much more lower and equal to 10 (with the standard deviation as high as 9.4 persons).

Figure 1. Distribution of individual network density



Source: own calculations based on data from Social Diagnosis 2011.

After matching the individual density of network from the Social Diagnose with the data from the Household Budget Survey, the mean density of network, calculated for all sample did not change.

4. A brief description of the HSAR model

With the consistency to the previous studies (see Łaszkiewicz et al., 2014) the hierarchical spatial autoregression model (HSAR) proposed by Dong and Harris (2014) was applied. It gives the ability to control for both contextual (as the effect of living in the same environment) and interaction effects, which are separately measured and therefore not confounded. Additional reason to use it is the multilevel spatial structure of our data with the individual level as households and community level. The HSAR model can be written as:

$$\begin{aligned} &\ln y_{ik} = \rho \sum_{i=1}^{N} w_{ij} \ln y_i + \sum_{m=1}^{M} \beta_m x_{mi} + \lambda \sum_{k=1}^{K} v_{kk'} \theta_k + \mu_k + \varepsilon_{ik}, \\ &\varepsilon_{ik} \sim N(0, \sigma_\varepsilon^2), \quad \mu_k \sim N(0, \sigma_\mu^2), \end{aligned} \tag{1}$$

where: y_{ik} is a monthly expenditures of household i localized in the community k, x_{mi} is the explanatory variable m, μ_k is the spatially uncorrelated random effect for the k community, θ_k is a term indicating the spatial component of random effect for the k community, ε_{ik} is the error term for the household i, w_{ij} is the element of the adjacency matrix **W**, $v_{kk'}$ is the element of the spatial proximity matrix **M**, σ_μ^2 is the variance of random effects, σ_ε^2 is the variance of error term, ρ is the estimated parameter of the social interaction, while λ is the estimated parameter of the spatial interaction (between communities localized in the same region).The total number of households in the sample is N=37 359 and the number of communities is equal to K=1 551. The $K \times K$ block-diagonal spatial matrix $\mathbf{M}=[v_{kk'}]$ was used to allow not only for the social but also spatial interactions. The elements of the **M** matrix were specified using a binary function with the value one if two communities are located in the same region (voivodship), or zero otherwise.

The following explanatory variables were chosen: logarithm of household income (*log_income*), household size (*h_size*), age of the reference person, sex of the reference person (1 if male), household biological type and class of locality. The last two factors were incorporated using the dummy variables. Further description and basic statistics for explanatory variables were provided by Łaszkiewicz et al. (2014).

The HSAR model was estimated using Bayesian Markov Chain Monte Carlo method. The detailed discussion about the estimation procedure for the HSAR model were provided in Dong and Harris (2014) work and it is omitted in this paper. The inferences were based on one MCMC chain that consist of 10 000 iterations with a burn-in period of 5000. The model was coded using the R language.

5. Design of simulated networks

Because the real structure of connections between agents is unknown, the simulated networks were applied. The structure of network was represented by the NxN adjacency matrix $\mathbf{W}=[w_{ij}]$. Three crucial steps to achieving the elements of the **W** matrix are: 1) decide about network density, 2) specify the rules of matching each pair of agents as connected or disconnected, 3) apply the weighting function to diversify the level of closeness in the network. The last step might be omitted by using a binary function which assigns 1 for connected agents or 0 otherwise. In such case we assume that agent is influenced in the same way by all others connected with him. Except above, the Jaccard similarity index was used as the weighting function because it might makes the structure of relations more realistic. Let us now describe the procedure of generating the networks.

5.1. Density of networks

In each structure of relations individual level density of network was achieved by using the data from Social Diagnosis 2011. By the individual density of network we mean the number of agents with who the individual *i* interact (denoted as n_i). After a series of ANOVA analyses we achieved five factors which diversify the density of individual network: age, sex, socio-economic group (*socioecon*), region, class of locality (*locality*). The variable age was aggregate into 5-years groups.

Table 1. Factors which influence the individual density of network according to the ANOVA

Variable	Partial SS	F	Prob > F
age	355 257	116	0.00
sex	3 246	12	0.00
socioecon	19334	17	0.00
region	40229	10	0.00
locality	42057	30	0.00
R-squared = 0.085			
age × sex × socioecon × region × locality	2 422 309	549	0.00
R-squared = 0.304			

Dependent variable is the number of persons with who respondent interacts.
Source: own calculations based on data from Social Diagnosis 2011.

The results from the ANOVA are presented in Table 1. For the model with no interactions between variables the coefficient of determination was low (R^2=8.5%), however all variables were significant. It means there are significant differences in the mean density of individual network between the groups separated by using the mentioned factors. However, when a set of the variables achieved by using the interactions between factors were used, the coefficient of determination in ANOVA increased to 30.4%. As the same attributes appear in the Household Budget Survey, they were used to match each household reference person with the mean number of people with who they contact. As the result the diversification of the network density for each household reference person was achieved.

5.2. Connectedness in networks

Three different strategies of linking agents were adopted: random, based on Jaccard index (non-random) and partially random. In the random approach for each agent i they were drawn n_i number of agents who potential influence the consumption behaviour of i. In contrast to Burgess et al. (2011) who also used the simulated network, five types of restrictions in drawing process were applied:

1) connected agents are drawing from all sample,
2) connected agents are drawing only from those who live in the same region and if the number of links is not enough, additional agents are drawing from the rest of the sample,
3) draw firstly n_{1i} agents from those who live in the same community and then draw additional n_i - n_{1i} agents from the rest of the sample,
4) draw firstly n_{1i} agents from those who live in the same community and then draw additional n_i - n_{1i} agents from different communities localized in the same region,
5) draw n_{1i} agents from those who live in the same community and if the number of links is not enough, additional agents are drawing from those who live in the same region.

The n_{1i} is defined as the number of persons who live in the same place or maximum 10 km from agent i and with who agent i interacts. Again, the data from Social Diagnosis were taken to obtain the value of n_{1i} and match it in the same way as for the network density. Moreover, in all generated networks the binary weighting function was applied so no difference in the influence of the connected agents was assumed. Finally, for each type of restrictions the network structure was drawn 10 times.

Finally, fixed density of individual network was tested for the condition 5 with values of density as: 5, 10, 15 and 18. The last value was taken as the median of the number of people with who agent interacts (from the whole sample). For each value of density one or two networks were drawn randomly.

In the second approach agents were matched non-randomly using the closeness measure. For each pair of agents the Jaccard index (Jaccard, 1912) was calculated, as:

$$J(i,j) = \frac{\left|S_i \cap S_j\right|}{\left|S_i \cup S_j\right|}, \tag{2}$$

where: S_i and S_j are the sets of *i*'s and *j*'s attributes. According to Eq. 2 the number of attributes common for agent *i* and *j* was divided by the total attributes taking into account, what guarantees $0 \leq J(i,j) \leq 1$. Seven different characteristics which describe the reference person of the household was taken into account: sex, age, disability, education level, socio-economic group, region, type of locality. If two agents are the same in all attributes, the value of Jaccard index is equal to 1. Next, using the measure of closeness those agents who were found as the most similar were lined. This procedure is the equivalent of the *k*-nearest neighbours procedure:

$$\begin{aligned} &\forall i,j,l \in \{1,\ldots,N\} : w_{ij} = J(i,j) \Leftrightarrow \forall j \in KNN(i,N), \forall l \in N - KNN(i,N), \\ &J(i,j) \geq J(i,l), \end{aligned} \tag{3}$$

where: $KNN(i, N)$ is the set of n_i number of agents from the full sample N, who potentially influence the behaviour of agent *i*.

In the third approach the structure of networks were set as partially random (mixed). Firstly, for each agent *i* only those agents, for who $J(i,j) \geq 0$ were selected. This means that the possibility of linking has a pair of agents who are at least quite similar. Then, the random drawing with no restrictions (as describes in the condition 1) were adopted. In contrast to the random approach it gave the ability to control the probability of matching agents. Despite, some form of randomness in the linking process was achieved, what was omitted in the second approach. By using the partially random approach 5 different network structures were drawn.

Each time when the network structure was generated the time (month) when the household declared the expenditures or in other words when the consumption behaviour occurred was taken into account. This let us control for the time consequences, e.g. agent *j* consumption behaviour might influences on the agent *i* decisions only when agent *j* decided about the consumption earlier than agent *i* (in the previous months or in the same, according to the database).

We obtained 62 networks with different structure of connections and as the result different degree of the homophily. The first approach should give us the most heterophily networks (see Easley and Kleinberg, 2010: 89). The exceptions are the networks with fixed individual density of network for which specific form of homophily which is a geographical homophily might be observed. The network obtained using non-random approach should be more homophily, while the network structures generated according to the last approach should be characterized by different degree of the homophily.

6. Empirical results and discussion

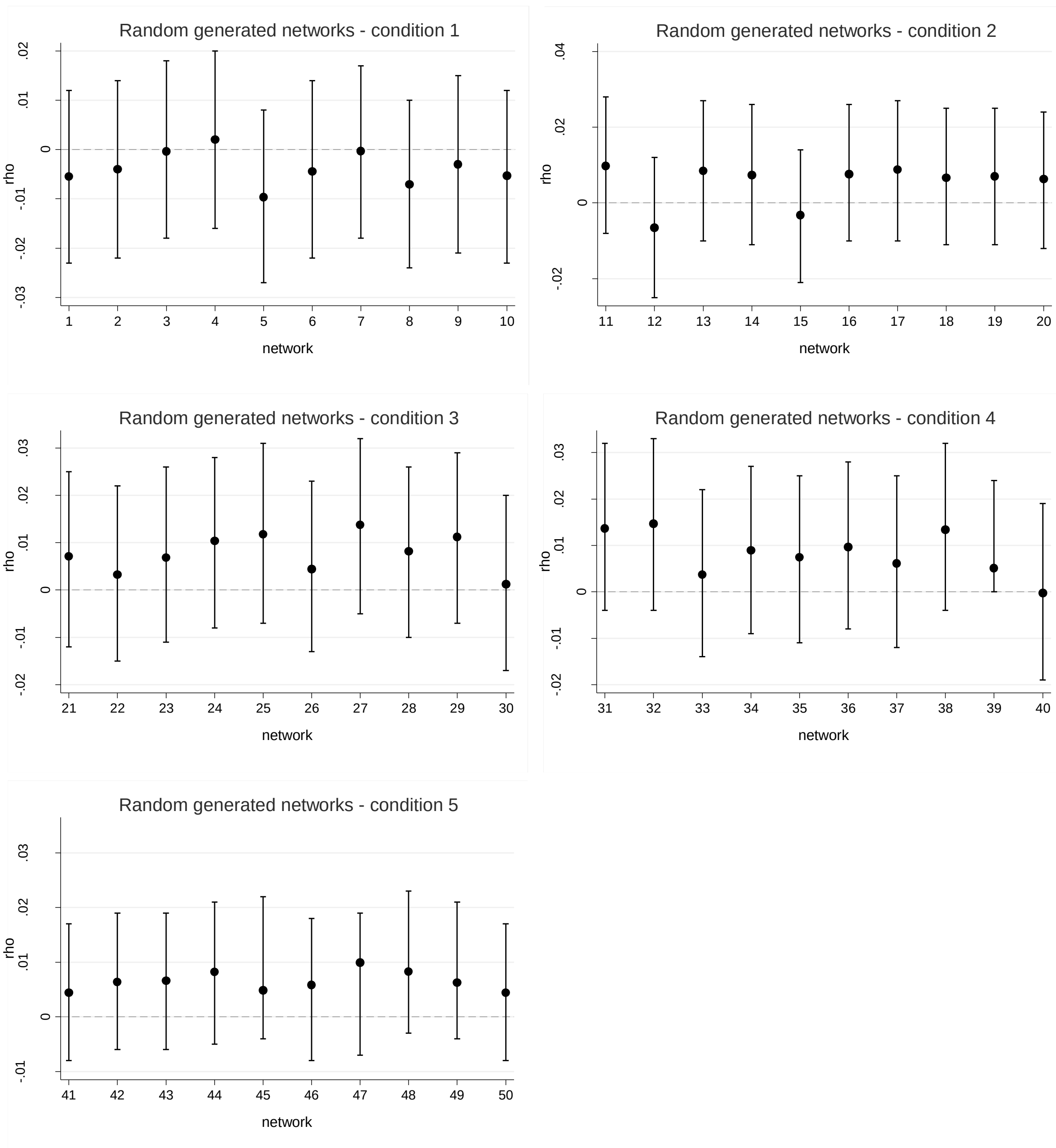
As an examination whether the heterophilous interactions support the social spillover of healthy food consumption, Figure 2 presents the estimated parameter for social dependence from the HSAR models. For each type of random generated networks the parameter for social interactions was found as insignificant, because the value of the 95% credible intervals contain the zero. This indicates that the interactions between dissimilar agents do not support the spillover of the consumption habits.

This results might suggest that households tend to interact with both those who are similar and differs but when connected agents have different features, the spillover of consumption habits does not occur. The main limitation of random generated networks is that in the real world individuals select a group of others with whom they interact more deliberately. Hence, the level of similarities obtained by using the random social networks might be lower from this observed in the real interpersonal links. As pointed by Burgess et al. (2011) matching agents is not just a simple random process.

In this study, the simulated networks were applied as an approximation of the non-homophily based interactions, so the potential lower level of a likeness (in comparison with the real structure of interpersonal relations) was ignored. Moreover, for 40 from 50 networks a simple random process of matching was modified by using the geographical limits in the drawing process.

Different restrictions which were imposed on the process of drawing agents' connections, do not affect significantly the posterior mean for social interactions parameter. For each heterophilous network the 95% credible intervals for ρ in the HSAR model contain the estimates of ρ from the model estimated using different adjacency matrix. Surprisingly, geographical concentration of interpersonal relations do not increase the chance for the social spillover of consumption habits. As the distribution of individual attributes is rather characterized by the geographical variability, the networks generated in accordance with the conditions 2-5 should be more homophilous. Hence, the results of random approach suggest that the interactions based on the homophily do not support the spillover of habits.

Figure 2. Estimated parameter for social interactions using heterophilous networks



Posterior mean assigns as a point, 2.5% and 97.5% credible intervals assign as the capped spikes.

Source: own calculations in R Cran.

One of the possible explanation why the estimates for the social interaction parameter do not increase after applying the spatial limitation in the drawing process is that the individual density of networks which were taken from the Social Diagnose does not fit the real density of social networks. Especially, the total number of connected others might be too large as only with a part of all people with who agent interacts the frequency of contact is enough intensive to giving a chance to participate in a daily consumption decisions, like buying the healthy food.

To verify whether density of individual social network affects our results, the additional 8 networks with fixed density was tested for the presence of social dependence (Table 2). For each of them the random matching of agents with spatial limitation as in the condition 5 was applied. The estimates for social interactions parameter was found as significant in each of the HSAR models. For networks with the lowest density of contacts the social spillover of consumption habits was also the lowest. The significant increase of the ρ parameter was observed as the effect of network density growth. The highest estimates for social interactions parameter was noticed when the mean number of connected others (form the Social Diagnosis) was used.

Table 2. Estimated parameter for social interactions using fixed density of individual network and its geographical homophily

Type of network	Posterior mean	2.5% CI	97.5% CI
random (as in the condition 5, with fixed individual density):			
5 links *per* agent (1)	0.07	0.05	0.09
5 links *per* agent (2)	0.08	0.06	0.11
10 links *per* agent (1)	0.16	0.12	0.19
10 links *per* agent (2)	0.16	0.12	0.19
15 links *per* agent	0.17	0.13	0.22
18 links *per* agent	0.22	0.18	0.26

Source: own calculations in R Cran.

Results showed that interactions within small communities with a limited number of persons supports the spillover of consumption behaviour. In particular, not all others with who individuals interact affect the consumption of healthy food but only a small number of them who lives in the same place. Sharing the common location seems to play the crucial role for social spillover of habits. It is because in a daily consumption of food products, households take into account opinions of those with who they have a daily, face to face contact.

The interpretation of this in the context of the homophily rule might be dual. On the one hand the geographical proximity of individuals might be a factor affecting degree of homophily. It is when spatial distribution of individuals' attributes is unequal and the preferences of linking with similar others might be successfully realized within the same community. On the other hand, geographical homophily might be the dominant form of the homophily. Such situation is when individuals tend to interact more with those others who are closer in the geographical meaning.

This is consistent with the First Law of Geography (Tobler, 1970): "Everything/everyone is related to everything/everyone else, but near things/individuals are more related than distant things/individuals". Also Zipf's law supports the explanation of the results: "It takes more energy to connect to those who are far away than those who are readily available" (McPherson et al., 2001). Although the links based on the geographical proximity might be weak (see Sudman, 1988), it is possible that they supports social spillover of daily habits more than interregional ties.

Finally, networks with a status homophily were tested for the presence of social spillover of the healthy food consumption. The results (Table 3) showed the significant spillover effect for homophilously structured social networks.

Table 3. Estimated parameter for social interaction using structural homophily of networks

Type of network	Posterior mean	2.5% CI	97.5% CI
non-random (KNN algorithm for Jaccard index > 0)	0.08	0.05	0.11
partially random:			
network 1	0.01	0.00	0.02
network 2	0.01	0.00	0.02
network 3	0.02	0.004	0.03
network 4	0.01	0.00	0.02
network 5	0.01	0.00	0.02

Source: own calculations in R Cran.

The 95% credible intervals does not contain zero both for non-random network and partially random structures of relations. However, the estimates for ρ parameter was found as significantly higher when the non-random (the most homophilously) network was used. It is worth to mentioned that significance for social interactions parameter was obtained using individual density of networks. Hence, we might expect that for the lower number of ties, the estimates for ρ parameter would be higher.

The results proved the homophily based interactions significantly increased adoption of consumption habits. However, the networks for which the status homophily was used in the matching agents support the transfer of behaviour less than the networks based on the interactions within small communities. This might suggests that for the social spillover of daily consumption habits the ties based on the geographical proximity are more important than links which occur due to the similarity of individuals' attributes. Finally, it is impossible to identify the mechanism responsible for the success of similarity based interactions.

7. Conclusion

Two main conclusions are from this study. First, homophilously structured social networks supported the spillover of healthy food consumption among Polish households in 2011. In contrast, for heterophily based relations spreading of habits did not occur. It is similar to e.g. Centola (2011) who proved the positive effects of homophily on adoption of health behaviors and innovations. Second, despite homophily based interactions are preferred for the adoption of habits, it seems the more important for social spillover is the geographical proximity of households. The highest value of estimates for social interactions parameter was found for the network with interactions within small communities. Moreover, the results demonstrate that small number of ties is better than larger for improving the spillover of daily consumption habits.

The limitation of this study is that the real structure of the interpersonal connections is not known. The only possibility to obtain the social network was by assuming the rules which constitute the mechanism of linking people. Further research based on different database should to be done to confirm the effect of social spillover.

Literature

Bott, H. (1928). Observation of play activities in a nursery school. *Genetic Psychology Monographs* 4(1): 44-88.

Burgess, S.; Sanderson, E.; Umaña-Aponte, M. (2011). School ties: An analysis of homophily in an adolescent friendship network. *The Centre for Market and Public Organisation* 267. Available at: http://www.bris.ac.uk/cmpo/publications/papers/2011/wp267.pdf. Accessed 15 July 2014.

Byrne, D.; Nelson, D. (1965). Attraction as a linear function of proportion of positive reinforcements. *Journal of personality and social psychology* 1(6): 659.

Centola, D. (2011). An experimental study of homophily in the adoption of health behaviour. *Science* 334(6060): 1269-1272.

Centola, D.; Willer, R.; Macy, M. (2005). The Emperor's Dilemma: A Computational Model of Self-Enforcing Norms1. *American Journal of Sociology* 110(4): 1009-1040.

Christakis, N. A.; Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine* 357(4): 370-379.

Collins, R. L.; Parks, G. A.; Marlatt, G. A. (1985). Social determinants of alcohol consumption: the effects of social interaction and model status on the self-administration of alcohol. *Journal of consulting and clinical psychology* 53(2): 189.

Council for Social Monitoring (2011). Integrated database. Available at: www.diagnoza.com. Accessed 2 August 2014.

Dong, G.; Harris, R. (2014). Spatial Autoregressive Models for Geographically Hierarchical Data Structures. *Geographical Analysis* 47(2): 173–191.

Easley, D.; Kleinberg, J. (2010). *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge: Cambridge University Press.

Golub, B.; Jackson, M. O. (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics* 127(3): 1287-1338.

Grinblatt, M.; Keloharju, M.; Ikäheimo, S. (2008). Social influence and consumption: evidence from the automobile purchases of neighbors. *The review of Economics and Statistics* 90(4): 735-753.

Jaccard, P. (1912). The distribution of the flora in the alpine zone. 1. *New phytologist* 11(2): 37-50.

Łaszkiewicz, E.; Dong, G.; Harris, R. (2014). The Effect Of Omitted Spatial Effects And Social Dependence In The Modelling Of Household Expenditure For Fruits And Vegetables. *Comparative Economic Research* 17(4): 155-172.

Lazarsfeld, P.F.; Merton, R.K. (1954). *Friendship as a social process*. In: Berger, M.; Abel, T.; Page, C.H. (eds.). *Freedom and control in modern society*: 18-66. Princeton, NJ: Van Nostrand.

Loomis, C. P. (1946). Political and occupational cleavages in a Hanoverian village, Germany: A sociometric study. *Sociometry* 9: 316-333.

Ma, L.; Krishnan, R.; Montgomery, A. (2010). *Homophily or Influence? An Empirical Analysis of Purchase within a Social Network*. Available at: http://community.mis.temple.edu/seminars/files/2011/02/krishnan-homophily-influence.pdf. Accessed 15 July 2014.

Marti de, J.; Zenou, Y. (2011). *Identity and social distance in friendship formation* 13. Stockholm University, Department of Economics.

McPherson, M.; Smith-Lovin, L.; Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology* 27: 415-444.

Richardson, H. M. (1940). Community of values as a factor in friendships of college and adult women. *The Journal of Social Psychology* 11(2): 303-312.

Rogers, E. M.; Bhowmik, D. K. (1970). Homophily-heterophily: Relational concepts for communication research. *Public opinion quarterly* 34(4): 523-538.

Rosenquist, J. N.; Murabito, J.; Fowler, J. H.; Christakis, N. A. (2010). The spread of alcohol consumption behavior in a large social network. *Annals of Internal Medicine* 152(7): 426-433.

Sorensen, G.; Stoddard, A. M.; Dubowitz, T.; Barbeau, E. M.; Bigby, J.; Emmons, K. M.; Peterson, K. E. (2007). The influence of social context on changes in fruit and vegetable consumption: results of the healthy directions studies. *American Journal of Public Health* 97(7): 1216.

Sudman, S. (1988). Experiments in measuring neighbor and relative social networks. *Social Networks* 10(1): 93-108.

Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography* 46: 234-240.

Wasserman, S.; Galaskiewicz, J. (1994). Advances in Social Network Analysis: Research in the Social and Behavioral Sciences. SAGE Publications, Inc.

Yavaş, M.; Yücel, G. (2014). Impact of Homophily on Diffusion Dynamics Over Social Networks. In: Bye, R.T., Zhang, H. (eds.). *Proceedings 27th European Conference on Modelling and Simulation. ECMS Webjørn Rekdalsbakken*. Available at: http://www.scs-europe.net/dlib/2013/ecms13papers/socint_ECMS2013_0177.pdf. Accessed 27 May 2014.

Żak, B.; Zbieg, A. (2012). Symulacja upodabniania się postaw w sieci relacji społecznych. *Nauki o Zarządzaniu* 4(13): 151-168.

Acknowledgements: This work was financial supported by grants ETIUDA1 from the Polish National Centre for Science, UMO-2013/08/T/HS4/00512.

Ocena symulowanej sieci kontaktów w hierarchicznym modelu autoregresji przestrzennej. Interakcje społeczne oparte na homofilii mają znaczenie dla rozprzestrzeniania się zachowań konsumpcyjnych

Streszczenie

Celem artykułu jest określenie, jaki wpływ na rozprzestrzenianie się zachowań konsumpcyjnych ma dobór sieci powiązań społecznych bazujących na homofilii i heterofilii. Jako miarę zachowań konsumpcyjnych przyjęto poziom wydatków gospodarstw domowych, zaś jako agenta uznano osobę odniesienia danego gospodarstwa. Sieci społeczne zbudowano zakładając, że każdy agent posiada indywidualną gęstość sieci kontaktów, wynikającą z jego cech. Powiązanie między gęstością sieci kontaktów a cechami uzyskano przy wykorzystaniu danych pochodzących z Diagnozy Społecznej (2011). W badaniu przetestowano trzy rodzaje sieci powiązań wspomagających lub ograniczających rozprzestrzenianie się zachowań konsumpcyjnych: 1) nielosowe z algorytmem *k* najbliższych sąsiadów, 2) losowe z pięcioma dodatkowymi typami ograniczeń, 3) częściowo losowe z losowaniem powiązań z ograniczonej liczby potencjalnych połączeń – stosując jako ograniczenie wartość indeksu Jaccarda. Występowanie efektu rozprzestrzeniania zachowań konsumpcyjnych zweryfikowano przez porównanie wartości oszacowań parametrów, uzyskanych w drodze estymacji wielopoziomowego modelu przestrzennej autoregresji (HSAR). Aplikacja modelu HSAR pozwoliła na jednoczesną kontrolę dodatkowych efektów, takich jak: przestrzenna heterogeniczność oraz zależności przestrzenne. Dane dotyczące wydatków gospodarstw domowych uzyskano z bazy danych indywidualnych Badania Budżetów Gospodarstw Domowych (2011). Uzyskane wyniki wykazały, że struktura powiązań bazująca na homofilii wspiera efekt rozprzestrzeniania się zachowań konsumpcyjnych. W przypadku sieci powiązań społecznych, bazujących na niepodobieństwie cech agentów, efekt taki nie zachodzi. Dodatkowo, zauważono iż kluczową rolę dla rozprzestrzeniania się zachowań konsumpcyjnych odgrywa bliskość geograficzna gospodarstw domowych.

Słowa kluczowe: zachowania konsumpcyjne, sieci społeczne, przestrzenne modelowanie wielopoziomowe