Regional
Innovation and Entrepeneurship
Research Center

WORKING PAPER SERIES

2018-002

Knowledge networks in regional development.
An agent based model and its application

Tamás Sebestyén
Regional Innovation and Entrepreneurship Research Center
MTA-PTE Innovation and Economic Growth Research Group
University of Pécs, Faculty of Business and Economics

Attila Varga
Regional Innovation and Entrepreneurship Research Center
MTA-PTE Innovation and Economic Growth Research Group
University of Pécs, Faculty of Business and Economics

Regional Innovation and Entrepreneurship Research Center
Faculty of Business and Economics
University of Pécs

H-7622, Pécs, Rákóczi street 80.

Phone: +36-72-501-599/63190
E-mail: rierc.center@ktk.pte.hu
Web: http://hu.rierc.ktk.pte.hu/

WP 2018-2Knowledge networks in regional development. An agent based model and its application

Tamás Sebestyén
Regional Innovation and Entrepreneurship Research Center
MTA-PTE Innovation and Economic Growth Research Group
University of Pécs, Faculty of Business and Economics

Attila Varga
Regional Innovation and Entrepreneurship Research Center
MTA-PTE Innovation and Economic Growth Research Group
University of Pécs, Faculty of Business and Economics

Abstract
Interventions targeting the support of interregional knowledge networks have become an increasingly important part of modern regional development policy. However our knowledge is limited with respect to the potential economic effects of concrete policy interventions. In order to examine these effects we developed and agent based model of network formation. We discuss the model in detail and provide the results of a simulation exercise the purpose of which is to illustrate the potential use of the model within a broader economic impact assessment framework.

Knowledge networks in regional development. An agent based model and its application[1]

1. Introduction

Interventions targeting the support of interregional knowledge networks have become an increasingly important part of modern regional development policy (McCann, Ortega-Argilés 2013). Recent debates within the realm of smart specialization policies also call the attention to the importance of learning from external sources, which is supposed to be an important tool in the hand of policymakers (Foray 2015, McCann, Ortega-Argilés 2015).

Besides anecdotal evidences (see e.g. Foray et al. 2011), there are a growing number of studies in the literature, which underline the potential effectiveness of these policies. Varga and Sebestyén (2013, 2017) compare Central and Eastern European (CEE) less developed regions with regions of old EU member states from the perspective of their capacity to absorb knowledge from research networks in two broad scientific fields supported by the EU Framework Programs (FPs). They found that knowledge accessible from FP networks affects regional innovation of CEE regions positively and significantly whereas the corresponding impact on regions of old member states was not significant. This finding clearly marks that a policy, which targets the improvement of network embeddedness is a viable option in these lagging regions.

Wanzenböck and Piribauer (2017) support the above findings by showing that lagging regions with lower levels of local knowledge resources have better chances to exploit the positive innovation effects of FP program participation. Closely related to these results Maggioni et al. (2017) emphasize that FP projects act more as possibilities for knowledge barter for core European regions whereas they act more as a one-way channel of knowledge diffusion from the core to lagging regions. Uhlbach et al. (2017) also conclude that participation in Framework Program projects can compensate lagging regions for the lack of locally available knowledge resources. De Nonic et al. (2017) arrive at the same conclusion by examining a panel dataset of regional patent data by showing that innovation performance in lagging regions is enhanced by increasing collaboration with knowledge intensive regions.

Despite this evidence on the role of external knowledge channels in the case of enhancing innovation in lagging regions, our knowledge is limited with respect to the potential economic effects of concrete policy interventions targeting network formation. In order to gain insights into these effects, one needs specially developed economic models, which capture the effects of network policies (Varga et al. 2017). In order to examine these effects we need a tool, which is able to capture the complex cumulative dynamics of network formation, including the feedback mechanisms between regional economic and innovation-related variables and interregional knowledge links. Agent based modeling provides a suitable tool to capture complex dynamic processes where the interactions between the system elements (agents) are important.

[1] The authors wish to thank Anna Csajkás for professional research assistance and Orsolya Hau-Horváth for her contribution to a fromer versions of the model presented in this paper. We also wish to thank Manfred Paier and Thomas Scherngell for useful comments.

The literature provides some model concepts which can serve as a good basis for policy impact modeling with respect to knowledge networks. The SKIN model is a widely used tool, which captures the complexity of knowledge lying behind the innovation process (Ahrweiler et al., 2004, Pyka et al., 2007; Gilbert et al., 2001, Korber and Paier, 2011, 2013). Although this model is able to capture cooperative knowledge creation, its main purpose is to model the whole innovation process and as a result it proves to be too complex to be integrated into an economic impact analysis model on the one hand and its scope is wider than the focus on link formation on the other hand. Some other models focus on the behavior and motivation of firms in the innovation process where link formation is again of secondary importance (e.g. Pyka and Saviotti, 2002; Beckenbach et al., 2007; Heshmati and Lens-Cesar, 2013).

Our goal therefore is to set up a model of network formation, which can be used in policy impact analysis. This model has to fulfill at least two goals. First, it needs to be simple as it is part of a complex economic impact model in which not only the direct effects of network policies on collaboration intensities or innovation activities can be traced but also their wider economic effects on the regional and national levels. Second, it must be able to capture, at least partially, the complex dynamics behind network formation, and its main focus must be network formation while less emphasis is put on other aspects of the innovation process. In addition to the previous two points, the model needs to be easy to empirically applied, which means that in order to run reasonable policy impact scenarios, we need to bring the model down to the data which is another point why too much complexity is problematic.

The paper is structured as follows. Section 2 provides a basically intuitive description of the model setting. Formal details of the model are summarized in the appendices. Section 3 then discusses the methods used to fill the model up with data, which is a two-step process consisting of the econometric estimation of a gravity model on the one hand and careful calibration of remaining model parameters on the other. Section 4 shows the results of a simulation exercise the purpose of which is to illustrate the potential use of the model. The paper is closed by some concluding remarks and we also note some important areas for further model development.

2. The model setup

In this section we describe the basic principles, which govern our model, and provide a formal description of the dynamics behind agent behavior. In the first step, we introduce our concept of the social space in which the agents are assumed to move in order to form connections with others. Then, we introduce the gravity and counter-force principles of the model. Finally agents' motion is set out and finally we introduce the way the heterogeneity of agents are accounted for in the model.

2.1 The social space

The basic idea behind our model is to place the agents in a social space. Their closeness or distance in this space corresponds to the probability and willingness for them to form links. The social space is a straightforward concept originating from pairwise network connections. If we assume a network of agents, which are linked together by selective, and possibly weighted ties, it is easy to construct a matrix of network distances, where the elements of the matrix represent the length of the shortest path[2] between any two agents within the network. These distances can be regarded as social distances, describing how far the agents are from each other in the social sense.[3] In principle, these network distances are independent of the conventional Euclidean space concept, but with the help of appropriate algorithms one can map the agents into Euclidean space of arbitrary dimension. In the present model setting we use a two-dimensional social space concept. This means that agents are placed in the two dimensional Euclidean space and we assume that their Euclidean distance within this space represent their social distance.[4] Then, agents can move within this space which then allows them to get closer or farther from each other. The main variable of an agent in our model is its position in the social space. As this social space is assumed to be two-dimensional, a pair of coordinates $(x^1_{i,t}, x^2_{i,t})$ fully describes the position of agent i in a given period t. Appendix A introduces the solution we applied to scale down network distances to a two-dimensional space.

2.2 The gravity principle

The idea behind agents' motion in the social space is borrowed from physics, and builds on the gravity principle: there is an attraction force between any two bodies, which positively depends on the mass of the bodies and negatively on the distance between them. The force of gravity is able to set otherwise static bodies into motion.

In our context agents are assumed to search for partners to collaborate with. In other terms, agents' behavior is basically determined by a desire to find connections with others. In the social space context this means that they would like to get closer to those other agents with whom they want to intensify collaboration. It is straightforward to say then that these agents are driven by attraction forces towards each other and this attraction force depends on the desirability of the other partners. Building on the gravity principle borrowed from physics, this desirability is then determined by the sizes of the two agents in question and their distances. As a result of this reasoning, we define an attractiveness measure $A_{i,j,t}$ among any two agents i and j in a given period t. In general, this attractiveness is determined by the sizes of the two agents ($M_{i,t}$ and $M_{j,t}$), and their distance ($D_{i,t}$):

$$A_{i,j,t} = f(M_{i,t}, M_{j,t}, D_{i,t}). \tag{1}$$

2.3 The counter-force principle

The gravity principle reflects agents' desire to collaborate with others. On the other hand agents also face costs of making new connections, which serve as a counter-force to the attraction force of gravity. These costs may come reflect transaction costs

---

[2] By shortest path between node A and node B we mean the minimum number of links through which B can be reached from A. If the links of the network are weighted, and the weights are interpreted as intensities or proximities, then the inverse of these weights are used to calculate shortest paths. See e.g. Wassermann and Faust (1994) or Barabási (2016) for more details.

[3] Two agents who are linked by a strong tie are socially proximate whereas two other agents who are indirectly linked together with many ties of small intensities are socially distant.

[4] The choice of the two dimensional Euclidean space is due to visual considerations: using this low dimensional representation, we are able to graphically represent agents' motion in the model, which helps in the interpretation of the results and predominantly understanding how the model works. An area for further developing the model would be to increase the dimensionality and check if the model performance improves.

like travelling or the opportunity cost of maintaining links (time).[5] Following the analogy with physics, if the force of gravity moves agents closer to each other, this counter-force keeps agents in place, obstructing its movement toward its goal. The intuition behind the use of the counter-force is that the benefits from new connections must exceed the cost of link formation in order to observe a new link to be formed. In the model we implement this cost in the form of a counter-force, pointing exactly to the opposite direction compared to where agents initially would like to move. As a result, the counter-force restricts the ability of agents to move away from their initial positions.

2.4 Agent heterogeneity

An important element of agent based models is that they are able to capture diversity of the agents in an explicit way. In our model we implement agent heterogeneity in several ways. First of all, the very nature of the model gives rise to heterogeneity in the sense that they are placed at different positions of the social space and their pairwise attractiveness measures can be different partly due to this positioning and partly due to heterogeneity in the individual variables used in the gravity equation. Also the coefficients for the mass variables in the gravity equation (proxied by R&D expenditures) are heterogeneous as they depend on other relational variables (cognitive proximity).[6]

Second, we assume that agents are restricted in their cognitive capacities with respect to evaluating and 'following' all other agents. It is reasonable to think that agents are not able to take account of all possible partners, only a subset of them. We implement this intuition with a parameter AP, which describes the extent of agents' capacities to take account of other agents.[7] This solution implements additional heterogeneity in the model as the ordered list of partners (based on attractiveness) is different across agents and as a result, all agents are going to 'follow' a different set of other agents.

Third, we assume that the speed of agents can be different. This agent-specific speed is linked to the size of the agent. This heterogeneity may reflect the fact that finding cooperating partners may be more or less desirable for different agents. Some agents may find it more prompting to find partners than others. We assume that this heterogeneity is reflected by an agent-specific speed parameter S_i driven by agent size, but the direction of this relationship is left to be determined through calibration.[8]

Fourth, we assume that agents are different with respect to the cost of their link formation. This assumption corresponds to differences in capacities in link formation, which may result in different costs of forming new links. This heterogeneity is

[5] See for example Bala and Goyal (2000), Jackson and Wolinsky (1996) or Carayol and Roux (2009) for some stylized models of network formation with a cost factor.

[6] See section 3.3 for the details on the estimated gravity equation.

[7] If there are N agents in the model, parameter AP takes integer values between 1 and $N-1$ and shows the number of other agents one can keep pace of. In other terms, we assume that agents only take into account the first AP most attractive partners. If $AP = N-1$ then they have full information whereas if $AP = 1$ then they only 'follow' the most attractive agent.

[8] See Appendix A for the details of the implementation of speed heterogeneity.

implemented through idiosyncratic counter-force strengths BF_i and in an analogous way to heterogeneous agent speed.[9]

3. Model estimation and calibration

In this section we focus on the methods of fitting this model to empirical data. First, we introduce the data used for estimation, then we separately describe the two main steps of empirical fitting: (i) the econometric estimation of the gravity model as the backbone of the model and (ii) the calibration of the remaining model parameters in order to improve the fit of the model. Finally, the main results of the estimation/calibration process are summarized.

3.1 Database

The empirical information we use to fit the model is basically linked to the gravity model in equation (1). All other parameters are calibrated, which means that they are optimized in order to get simulated network connection intensities as close to the observed ones as possible. Following from this, the data we use correspond to the variables appearing in equation (2), which is the empirical counterpart of the general gravity equation in (1).[10]

In this paper we define agents as regions: a sample of 181 NUTS2 regions of the European Union is applied for 12 years between 2001 and 2012. This is the longest possible interval on which all the data detailed below are available. The first variable we use is a proxy for network connections among the regions. There are many possibilities to trace cooperation between individuals, firms, institutions, regions or countries like co-patenting or co-publication.[11] In our study we build on information drawn from data on EU Framework Programs (FPs).[12] Basically, we take all these FP projects and assume a link between two regions if institutions/firms from these two regions participated in the same project. This way we get a weighted network between our sample regions for every year in the sample where the weights reflect the number of FP projects the institutions/firms of the two regions cooperated in. In order to avoid large fluctuations in the observed network connections, we use a 5-year moving average of the raw network matrices obtained from FP data.[13]

The second variable is the size of the agents, which is proxied by their total expenditure on research and development measured in real terms. The source of this data is the Eurostat database and we also use 5 year moving averages in order to harmonize these series with the collaboration data.

---

[9] See Appendix A for the details of the implementation of counter-force heterogeneity.
[10] See section 3.3 for the details on the specification of the gravity model.
[11] For networks of co-publication, see for example Abbasi et al. (2011, 2012), Beadury and Clerk-Iamalice (2010), Hopp et al. (2010) or Rumsey-Wairepo (2006). Co-oatenting networks are examined by Fischer et al. (2005), Maggoni et al. (2011) or Cassi and Plunket (2015).
[12] In these programs institutions and firms from different regions engage in cooperative research projects of different lengths and the collaborative nature of these projects allows us to infer on cooperation intensities between regions.
[13] These large changes are most likely in the case of less central regions – the majority – where one big project may significantly influence the cooperation intensities from one year to the other in the starting and ending years of the project.

The third variable in the gravity model is the cognitive proximity of the agents/regions. In this respect, we employ the technological overlap index proposed by Cantner and Meder (2007). This method takes into account the portfolio of patents in two regions with respect to the 8 main IPC classes and defines the index of overlap on the basis of similarity between these portfolios. This index lies between 0 and 1, the former referring to complete dissimilarity/distance whereas the latter showing complete similarity/proximity with respect to the patent portfolios which can be used as a reflection of cognitive similarities in the regions' knowledge bases. As before, the raw data is transformed into 5-year moving averages to avoid large swings in the data.

Finally, the gravity equation uses a second type of distance measure, the distance of agents in the two-dimensional social space. With respect to data use, this variable has a double nature. During the simulation of the model, these distances are endogenous, so that the dynamics of the model determine them. However, when the model is started, these distances are calculated from initial positions, which are based on the data.[14] Social distances are simple Euclidean distances of two regions in the two-dimensional social space.

Table 1 below gives a brief summary of the data and data sources used for model estimation and calibration.

Table 1 – Data and data sources used for estimation and calibration

Variable	Data used	Source
$COOP_{i,j,t}$	Joint Framework Program projects between regions i and j in year t	Own calculations on the basis of the administrative database of FP 5, 6, 7 (DG RTD Dir A)
$RD_{i,t}$	Total R&D expenditures of region i in year t measured in constant 2000 price and PPS	Eurostat
$CP_{i,j,t}$	Technological overlap index between regions i and j in year t, using IPC main classes	Own calculations on the basis of Cantner and Meder (2007)
$SD_{i,j,0}$	Euclidean distance of regions i and j in the initial position of the model	Own calculations based on MDS mapping of network data into 2 dimensions

3.3 Estimation of the gravity equation

After determining the initial positions of the regions in the two dimensional social space, the next step is the estimation of the gravity model in equation (1). The data used for this estimation is summarized in Table 1. As described there, the data we use has a panel structure: the observation units are the region-pairs, on which the cooperation intensities ($COOP_{i,j,t}$), the social distances ($SD_{i,j,t}$) and cognitive proximities ($CP_{i,j,t}$) are identified. After experimenting with different regression settings, the following fixed effects panel model was estimated:

$$\ln\left(COOP_{i,j,t}\right) = a_0 + a_1 \cdot \ln\left(RD_{i,t}\right) \cdot CP_{i,j,t} + a_2 \cdot \ln\left(RD_{j,t}\right) \cdot CP_{i,j,t} + a_3 \cdot \ln\left(SD_{i,j,t}\right) + a_{4,t} \cdot TD_t + a_{5,i,j} \cdot FE_{i,j} + \varepsilon_{i,j,t} \tag{2}$$

[14] See Appendix B for details on this.

where $COOP_{i,j,t}$ is the FP cooperation intensities between regions i and j in year t, $RD_{i,t}$ is total R&D expenditures of region i in year t, $CP_{i,j,t}$ is cognitive proximities of regions i and j in year t and $SD_{i,j,t}$ is the social distance between regions i and j in year t as a result of the MDS algorithm described in Appendix A. TD_t is a time dummy, $FE_{i,j}$ is an observation-specific fixed effect and $\varepsilon_{i,j,t}$ is an observation and time-specific error term.

Table 2 – Regression results from estimating the gravity model

Fixed-effects, using 390960 observations included 32580 cross-sectional units, time-series length = 12. Dependent variable: $\ln(COOP_{i,j,y})$		
Variable	Coefficient	p-value
a_0	−3.65407***	<0.0001
a_1	0.0475673***	<0.0001
a_2	0.0475673***	<0.0001
a_3	−0.300823***	<0.0001
$a_{4,2}$	0.162147***	<0.0001
$a_{4,3}$	0.381809***	<0.0001
$a_{4,4}$	0.556079***	<0.0001
$a_{4,5}$	0.69274***	<0.0001
$a_{4,6}$	0.790328***	<0.0001
$a_{4,7}$	0.742078***	<0.0001
$a_{4,8}$	0.758434***	<0.0001
$a_{4,9}$	0.654642***	<0.0001
$a_{4,10}$	0.593571***	<0.0001
$a_{4,11}$	0.508909***	<0.0001
$a_{4,12}$	0.314529***	<0.0001
LSDV R-squared	0.876869	
P-value(F)	0.000000	

The results of the estimation are shown in Table 2. These results indicate that all parameters (including the time dummies and the fixed effects) are significant and they show the expected signs: R&D has a positive impact on cooperation in interaction with cognitive proximity, whereas social distance has a negative impact. Time dummies are positive and show an inverted U-shaped pattern: cooperation intensities on average increase up to the sixth year of the sample and decrease after that.[15] The model provides an explanatory power of 0.88, which is quite good, and means that even the gravity model is able to capture large part of the variation in the dependent variable.

3.4 Parameter calibration

The remaining set of parameters is calibrated in order to improve the fit of the model. This means altogether 5 parameters:

- $\bar{S}$ – the general speed of agents. This parameter reflects the overall strength of agents' motivation to cooperate: higher speed means that agents are

[15] The reference period against which these time dummies are defined is the first year of the sample.

characterized by a strong motivation to cooperate with others as they move faster towards the desired cooperation partners in the social space.

- $\overline{BF}$ – the general strength of counter-force. This parameter reflects the average cost of link formation: the higher this parameter, the more reluctant agents are to move towards their desired partners in the social space.
- AP – the set of most attractive partners which agents follow in the social space. This parameter reflects the limited cognitive capacities of agents by restricting the set of other agents which they are able to keep track of.
- SR – speed heterogeneity parameter. This parameter reflects how the size of the agents is related to their speed. Through this parameter we can set whether larger agents move faster or slower and what is the strength of this relationship.
- BR – counter-force heterogeneity parameter. Analogously to the speed heterogeneity parameter, this parameter reflects how the size of the agents is related to the strength of the counter-force applied to them. Through this parameter we can set whether larger agents face higher or lower cost of link formation and what is the strength of this relationship.

Note, that although the last two parameters contribute to the heterogeneity of agent behavior, the parameters are still aggregate in the sense that we do not define a specific speed/counter-force to be calibrated for all agents, but only the relationship between size and speed/counter-force is established which require two parameters to be set for each: a shift parameter ($\bar{S}$ and $\overline{BF}$) and a slope parameter (SR and BR).

The range of the parameters on which the optimal parameter setting is searched is depicted in Table 3. These ranges were selected after iterative experiments so that the global optimum most likely lies within this range. All parameters are continuous except AP, which only accepts integer values by definition.

Table 3 – Calibration ranges of parameters

Parameter	Description	Range	Optimal value
AP	Length of agents' partner list	$[1{,}180]$	148
$\bar{S}$	Common speed	$[0{,}0.1]$	0.07
$\overline{BF}$	Common counter-force strength	$[0{,}1]$	0.596
SR	Speed elasticity on agent size	$[-1{,}1]$	-0.001
BR	Counter-force strength elasticity on agent size	$[-1{,}1]$	-0.052

These five parameters were calibrated in the following manner. Given the estimated gravity model parameters and the initial positions for the first year of the sample, we set up the model. Then, the model is shocked by the observed changes in R&D and cognitive proximity data between the first and second years of the sample. The shock sets the model in motion and arrives into a steady state where agents do not move further. Once the model reaches this new steady state, the pairwise attractiveness values (given by the shocked exogenous variables and the final positions and social distances) are calculated.[16] Then, we calculate the mean absolute difference between these pairwise attractiveness values and the observed cooperation intensities. This method yields an error measurement for the model in the second year of the sample. Then, we move on to the third year and use the same method to arrive at the error measurement for this year. Following this method up to the last year of the sample, we obtain 11 error measurements for the 11 consecutive years in the sample.[17] Finally, we simply calculate the average of these 11 error measures to capture how far the simulation gets on average over the years from the observed cooperation intensities.

The parameter calibration method targets the overall error measure and tries to minimize it by setting the five parameters in order to get the smallest error possible. As the optimization problem here is quite complex, with a highly nonlinear objective function (the model itself run for 11 years, which means 11 times a number of periods long enough to reach the steady state) and one of the parameters is bounded to integer values, we used genetic algorithm to find the best fit of the model[18].

3.5 Calibration results

In the previous sections we have described the method through which the model was fit to the data. This consists of (i) setting up initial positions, (ii) estimating the gravity model and (iii) calibrating the remaining model parameters. Table 2 shows the estimated parameters of the gravity model and the last column in Table 3 shows the final, optimized parameters, which were calibrated. The AP parameter is optimized to 148, which is a considerable amount of partners to be 'followed', given the maximum which is 180. The speed parameter does not have a clear meaning in itself. Its interpretation comes from the size of the space in which agents move. In the initial outset agents are placed in a square the sides of which are of length 2. This means that moving with the calibrated speed (0.07) on average agents cross 3.5% of the space in one step/period of the simulation. The optimal value of the counter-force parameter is close to 0.6, which means that the cost of link formation is not that high, it could be qualified as intermediate.[19] The two heterogeneity parameters are set different from zero which indicates that heterogeneity in speed and the cost of link formation is an important factor in the model, improving the fit of the model once taken into account. Both values are negative which shows that larger regions tend to move more slowly (their motivation to find new collaborations is lower, which is a reasonable conclusion as their inherent innovation capabilities are abundant an external sources of knowledge are less required) and tend to have lower counter-forces (their cost of forming links is lower which is a reasonable finding again).

---

[16] Convergence is assumed when the percentage change in the mean pairwise social distance (from the start of the simulation to the current period) remains below ±0.01%. As agents may 'fluctuate' around their steady state position, we run the model for 50 periods after achieving convergence and the final equilibrium positions and attraction values after the shock is calculated as the average over these 50 overrun periods.

[17] As the fitting method operates on the basis how the simulation model can approach the observed networks after a shock, we cannot fit the model to the first year as it serves as the starting point for the process.

[18] Appendix C provides more information about the genetic algorithm we used.

[19] It is important to note here that by setting the counter-force parameter below unity the stationary nature of the model does not disappear, only agents need more time to arrive into a new steady state and move farther from their initial positions. Even if this parameter is zero, the model keeps stationarity, but in this case agents collapse into the mass center of the system.

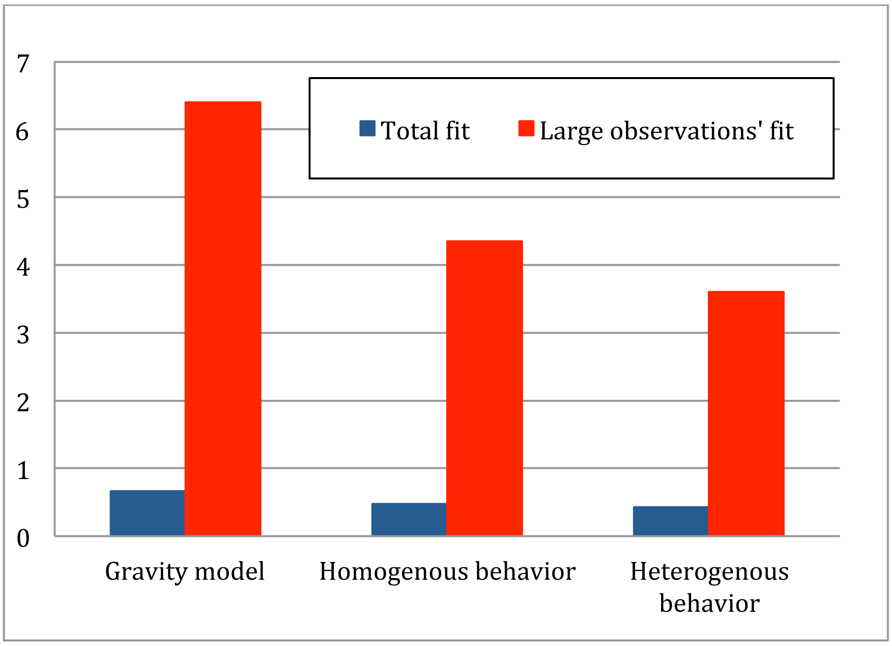

Figure 1 – Improvement in the absolute performance of different model settings

Figure 1 focuses on the absolute errors and improvements in model fit. It is visible that even the gravity model produces a reasonable fit: the average error is less than 1, which – recalling that the unit of the fit measurement is the number of joint FP projects between regions – is very small. However, in the case of large collaboration[20] intensities, the gravity model makes a much larger error. This error is reduced from more than 6 to less than 4 if we augment the gravity model with further agent based elements. Overall, the estimated gravity model provides a good fit to the data, but the calibrations build around this gravity model are able to further improve the fit of the model by a significant 35%. This improvement is even higher (reaching 44%) in the case of those observed collaborations, which are more intensive than the average.

3.6 How the simulation model works

In this section we give a brief summary of the agent-based model. In Figure 2 we show the basic logic of the model. The red boxes on the right represent the data, which is used for the model. This data determine the exogenous variables, which are the R&D intensities of agents and their cognitive proximities as well as their initial positions, which are derived from observed network distances. The initial positions of the model represent the location of the agents in the social space and their social distance. The R&D intensities, the cognitive proximities and the social distances determine the attractiveness values through the gravity equation in (2). Once a shock hits the model thorough some of the exogenous variables, agents are set in motion. Attractiveness changes leading to the motion of agents, which feeds back into the social distances through changing positions. Changing social distances again affect attractiveness, and the model keeps 'moving' for a while even in the absence of

[20] Large observation means higher than the average collaboration intensity over the whole sample. The average collaboration intensity is 2.45, which means roughly two and a half joint projects between regions.

further shocks. However, the model proves to be stationary, so after a one-time shock, motion settles down and agents reach a new steady state position.[21]

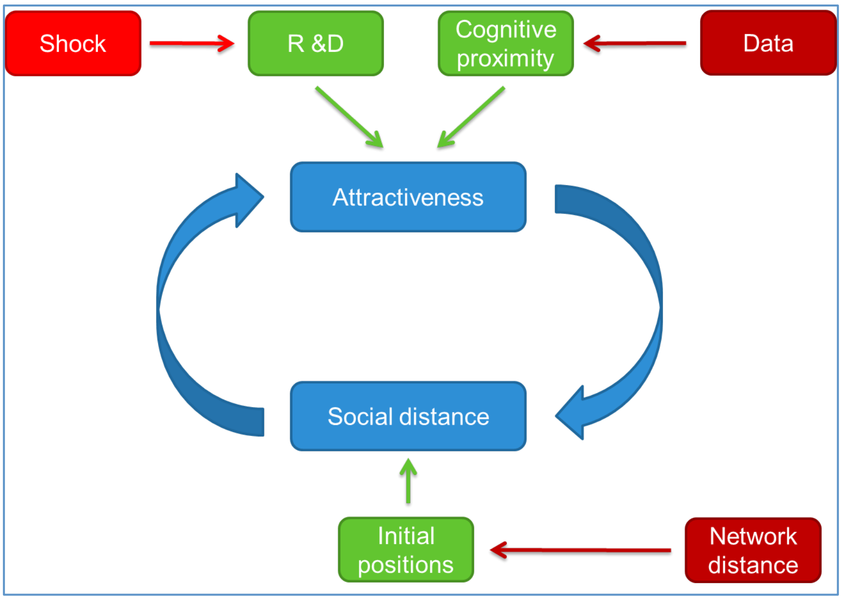

Figure 2 – The schematic setup of the agent based model

4. Network policy simulations

In this section we provide a brief discussion of some illustrative simulation exercises executed with the calibrated model. First, we carry out a policy optimization and then we implement a more complex study where this model is integrated into the GMR policy impact-modeling framework.

4.1 Policy optimization with the model

The model can be used to set out optimal policies with respect to network formation. In the following example we raise the question that what are the directions to which a given region should specialize in order to improve its network embeddedness. By specializing on a given technological field, a region can selectively improve cognitive/technological proximity with several other regions while decreasing it towards others.

On the basis of this, our exercise builds on shocking the cognitive proximity of one region with a selected other region in a positive way. Then, we check the resulting change in network structures and measure the change in network embeddedness as a result of the shock. Network embeddedness in this exercise is measured by the Ego Network Quality index, developed by Sebestyén and Varga (2013a, 2013b). This index measures the quality of knowledge accessible from a given network position by taking into account (i) the connectivity of the direct and indirect partners of a given agent and (ii) the knowledge possessed by these partners.

The simulation setting is the following. We choose Central Hungary, the capital region of Hungary, as the region of interest. This region is the most developed one in Hungary, with quite many cooperation with other regions and relatively developed

[21] Appendix D provides the technicalities behind the motion of agents in the social space.

industries. Then, we consecutively shock the cognitive proximity of Central Hungary with all other regions in our sample, and run the simulation for 10 years after the shock. The goal is to pick the region with which the increase in cognitive/technological similarity brings about the largest improvement in the ENQ index (network embeddedness).

Figure 3 shows how the ENQ index changes after a 1% shock to cognitive proximities with all the other regions. The black line shows the baseline case, when there is no shock at all. All colored lines represent the time path of the ENQ index after the shock with a given region. It can be seen that the paths are very similar, they move around the baseline, some moving above, some moving below it.

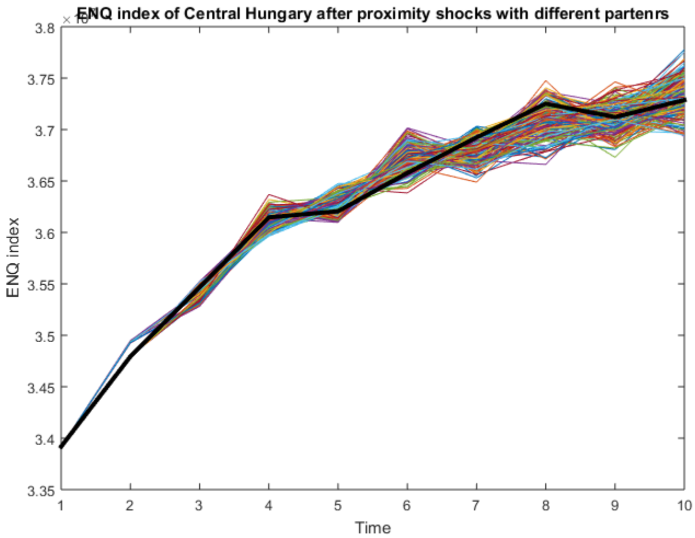

Figure 3 –Dynamics of the ENQ index after a shock to cognitive proximity with all other regions.

The task is to pick the line, which corresponds to the largest increase in ENQ. It can be hypothesized that the most benefit with respect to network position comes if those connections/partners are supported, which are relatively large. In other terms, one may argue that the target of support must be to form links with large, important regions/partners or, to further improve already existing collaboration linkages. We checked whether this kind of tendency is valid in the light of the simulation exercises. As a result of the simulation, we calculated the change in the ENQ index when Central Hungary's cognitive proximity is increased with all other possible regions in turn. Now, we examine if the resulting changes in the ENQ index correlates with the size of the partner or the size of the existing collaboration.

Figure 4 shows the scatter plots of the corresponding analysis. The panel on the left hand side shows the correlation between the change in the ENQ index (vertical axis), and the size of the partner region (horizontal axis) with which the shock is implemented. The right hand side panel shows the correlation between the change in the ENQ index (vertical axis) and the size of connection (horizontal axis), which is shocked. The main conclusion from the diagrams is that there is no correlation between either the size of the partners or the size of the already existing connection and the change in the ENQ as a result of a shock to the cognitive proximity of Central Hungary with other regions.

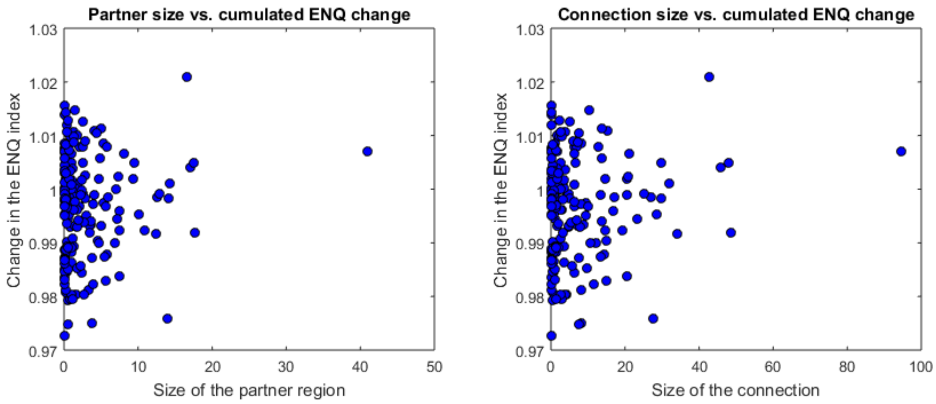

Figure 4 – Correlation of ENQ change with partner sizes and existing collaboration intensities

There is a clear policy implication from this: policies supporting network formation should not be rule-based in the sense that it supports link formation with large, central regions or strengthen already existing relationships. In contrast, they should be selective and individually designed support programs, tailored to the specific endowments and characteristics of the region in question. The role of impact modeling is therefore inevitable in this case – our model presented in this paper may be a candidate to fulfill this role.

Table 4 – The 10 best options for Central Hungary to increase cognitive proximities with

Cumulated change in ENQ			
1	Oberbayern	DE21	2.10%
2	Koblenz	DEB1	1.56%
3	Eesti	EE	1.47%
4	Basse-Normandie	FR25	1.43%
5	Burgenland (A)	AT11	1.38%
6	Dolnoslaskie	PL51	1.29%
7	Andalucía	ES61	1.26%
8	Itä-Suomi	FI13	1.20%
9	Tübingen	DE14	1.13%
10	Praha	CZ01	1.09%

Table 4 shows the 10 best partners with respect to the overall (cumulated) change in ENQ. The best partner to increase cognitive proximity with is Oberbayern in Germany, which is a reasonable result having in mind that this region is a center of German automotive industry and the Hungarian economy is tightly linked to the German economy on one hand and to its automotive industry on the other.

4.2 Economic impact modeling

In the previous section we have shown an illustrative simulation in order to underline the capabilities of the model. In this section we briefly present the results of a simulation where the innovation network model is used as an input to a regional-level economic impact model.

The model framework we use is the GMR-Europe model, which is a large-scale macroeconomic-regional model for European regions for the purposes of impact analysis for innovation policy. The model is built around standard economic modeling tools and is capable of reflecting the economic outcomes (changes in GDP, employment, prices, etc.) of different innovation policy interventions both at the regional and at the macroeconomic level. An inherent feature of the model is its detailed productivity block. This is the point where changes in the interregional knowledge network are built in the model affecting regional productivity in a complex way. The two models, the agent based knowledge network formation model and the GMR economic policy impact model is linked together through the ENQ index. By simulating expected changes in the structure of the knowledge network the AB model provides relative ENQ impacts over the simulation period which then go into the GMR model's productivity block initiating widespread economic impacts across the regions in the model. For a more detailed description on the GMR modeling framework and the GMR-Europe model the reader is directed to Varga et al. (2015).

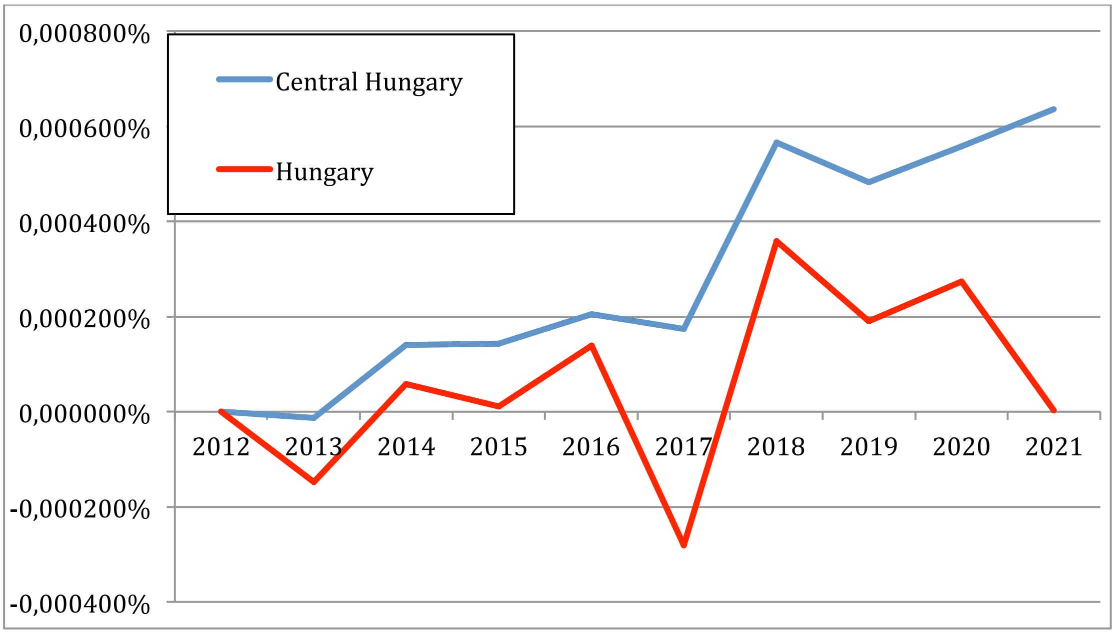

Figure 5 – The impact of increased cognitive proximity between Central Hungary and Obarbayern on Gross Value Added

Figure 5 shows the results of a simulation run with the GMR-Europe model where we used the results from the optimization exercise in section 5.1. The cognitive proximity of Central Hungary is increased with Oberbayern by 1%, which then leads to changes in the structure of the knowledge network across the regions in our sample. The resulting change in the ENQ index then sets in motion the productivity block in the GMR model, modifying productivities in the sample regions (predominantly in Central Hungary and Oberbayern), which then brings around different economic changes (increasing or decreasing employment, prices, production) all over the landscape. The figure shows how the Gross Value Added in Central Hungary and in Hungary as a whole changes as a result from the shock given to cognitive proximity. Over the ten years of the simulation there is a slight positive impact both for the region and the country, the latter being somewhat smaller.

The improvement in the network embeddedness (ENQ index) of Central Hungary has a positive impact on its productivity and as a result, brings about a positive change in the economic output of the region. Moreover, this positive effect spills over to Hungary as a whole through two channels. First, through an improvement of other regions' network position (which is a result from the complex dynamics in the network formation model) and second through dynamic economic interactions between regions such as trade and migration of production factors. The great power of the two models linked together is that it is capable to quantify the likely effects of such interventions and as a result, render them suitable for policy impact analysis in the field of innovation policies where 'soft' factors like knowledge, research cooperation and alike are of utmost importance.

5. Conclusions

Targeting interregional knowledge networks has become an important tool in innovation policy which is backed by findings in the literature showing how lagging regions can benefit from cooperation with leading innovation centers. In this paper we have set out a modeling framework, which targets capturing the dynamics of link formation in a knowledge network while it is simple enough to be integrated into a broader economic impact analysis model and also empirically tractable in the sense that the model can be brought to the data easily concerning data requirements and complexity.

The model is built around a standard gravity model, which captures a static relationship between connection intensities and agent sizes as well as distances in several dimensions. Then, this gravity model is augmented with some dynamic elements, which move agents in the social space according to attraction forces governed by the gravity model. Agents' motion in this social space is then translated into changing link formation probabilities and as a result a changing setup of the knowledge network structure.

The paper shows the calibration method of the model and argues that augmenting the standard gravity model with some elements of complex dynamics can improve the empirical properties of the model reducing fitting errors. This improvement is proved to be larger in the case of those connections, which are "large" in the sense that lay at the fat tail of a highly asymmetric distribution. This means that the augmented gravity model is not only able to capture average tendencies in the sample used to fit the model but it reflects even more precisely the dynamics of link formation in the case of the most important agents which is a significant result as these agents constitute the backbone of the network and their interconnectedness determine the overall structure and performance of the knowledge network to a large extent.

Finally we have shown a simple illustrative simulation where the calibrated model was used to run an optimization experiment: we calculated that in the case of Central Hungary (capital region of Hungary) which possible partner region yields the highest change in network embeddedness in terms of the Ego Network Quality index. Then, the results of this simulation were integrated in a broader economic impact model, the GMR-Europe model, which showed how the wider economic impacts on the regional and macroeconomic levels can be traced with the help of our model setting. Also, these simulations have shown that we cannot find significant relationship between the change in the improvement in network position and the size of or the former connection intensity of possible partners. This means that there is no general receipt for innovation policies targeting network formation but policy needs to conduct a careful analysis of possible impacts in the case of every region. The help of policy impact models is thus proved to be inevitable.

Although the proposed modeling framework seems promising, there are several limitations and avenues for further development in this field. First, the simulation exercise showed here aggregated agent behavior to the regional level. Although there may be some justification for this, the basic principles of agent based modeling is built on the assumption that the disaggregation level of these models must be as close to the real decision making units as possible. On the other hand, the model framework is easily applied in a context where agents are institutions (firms, universities, etc.) or even individuals. Without any modification to the model structure, the data background must be richer in this case. Second, the choice of the dimensionality in the case of the social space is arbitrary in this case (we used a 2D representation as it provides a simple, intuitive and easily visualized interpretation of agent motion). However, this reduction in dimensionality may result in some unintended results with respect to agent motion and link formation dynamics as it restricts agents' abilities to significantly differentiate their link formation strategies towards other agents. Currently we are working on increasing the dimensionality of the model and searching for an optimal choice in this respect, but this process is severely limited by computational capacities raising some technical issues to be resolved in the near future.

References

Ahrweiler, P., Pyka, A., & Gilbert, N. (2004) *Simulating knowledge dynamics in innovation networks (SKIN)*. Volkswirtschaftliche Diskussionsreihe. Institut für Volkswirtschaftslehre der Universität Augsburg.

Bala, V., Goyal, S. (2000): A Noncooperative Model of Network Formation. Econometrica, vol. 68(5), pp. 1181-1230.

Barabási, A-L. (2016): *Network Science*. Cambridge University Press.

Beckenbach, F., Briegel, R., & Daskalakis, M. (2007) Behavioral Foundation and Agent Based Simulation of Regional Innovation Dynamics. *Papers on agent-based economics*. Univ., Fachbereich Wirtschaftswiss., Fachgebiet Umwelt-und Innovationsökonomik.

Carayol, N., Roux, P. (2009): Knowledge flows and the geography of networks: A strategic model of small world formation. Journal of Economic Behavior & Organization, 71(2), pp. 414-427.

Cassi, L., Plunket, A. (2015): Research Collaboration in Co-inventor Networks: Combining Closure, Bridging and Proximities. *Regional Studies*, 49(6), pp. 936-954.

Conn, A.R., Gould, N.I.M., Toint, Ph.L. (1991): A Globally Convergent Augmented Lagrangian Algorithm for Optimization with General Constraints and Simple Bounds. *SIAM Journal on Numerical Analysis* 28, pp. 545–572.

Conn, A.R., Gould, N.I.M., Toint, Ph.L. (1997) A Globally Convergent Augmented Lagrangian Barrier Algorithm for Optimization with General Inequality Constraints and Simple Bounds. *Mathematics of Computation* 66, pp. 261–288.

De Nonic, Orsia, L, Belussib, F (2017) The role of collaborative networks in supporting the innovation performances of lagging-behind European regions. *Research Policy* (in press).

Fischer, M., Scherngell, T., Jansenberger, E. (2005): The Geography of Knowledge Spillovers between High-Technology Firms in Europe. Evidence from a Spatial Interaction Modelling Perspective. *MPRA Paper 77786*, University Library of Municg, germany.

Foray, D. (2015) *Smart specialization. Opportunities and challenges for regional innovation policy*. Routledge, London and New York.

Foray, D., David, P., Hall, B. (2011) *Smart specialization. From academic idea to political instrument, the surprising career of a concept and the difficulties involved in its implementation*. MTEI-WORKING_PAPER-2011-001.

Gilbert, N., Pyka, A., Ahrweiler, P. (2001) Innovation networks-a simulation approach. *Journal of Artificial Societies and Social Simulation* 4, 1-13.

Goldberg, D.E. (1989) *Genetic Algorithms in Search, Optimization & Machine Learning*, Addison-Wesley.

Hesmati, A., & Lenz-Cesar, F. (2013) Agent-based simulation of cooperative innovation in R&D. *Research Evaluation* 22, 15-29.

Jackson, M.O., Wolinsky, A. (1996): A Strategic Model of Social and Economic Networks. Journal of Economic Theory, 71(1) pp. 44-74.

Korber, M., Paier, M. (2011) Exploring the Effects of Public Research Funding on Biotech Innovation: An Agent-Based Simulation Approach. *Proceedings of the Eighth International Conference on Complex Systems" Unifying Themes in Complex Systems* 2011. 599-613.

Korber, M., Paier, M. (2013) Effects of Competence Centres on Regional Knowledge Production: An Agent-Based Simulation of the Vienna Life Sciences Innovation System. In: Scherngell, T. (ed.) *The Geography of Networks and R&D Collaborations*. Berlin-Heidelberg-New York: Springer, pp. 353-371.

Maggioni, M., Uberti, T., Nosvelli, M. (2017) The “Political” Geography of Research Networks: FP6 whithin a “Two Speed” ERA. *International Regional Science Review* 40, 337-376.

Maggioni, M., Uberti, E., Usai, S. (2011): Treating Patent sas Relational Data: Knowledge Transfers and Spillovers across Italian Provinces. *Industry and Innovation*, 18(1) pp. 39-67.

McCann, P., Ortega-Argilés, R. (2013) Modern regional innovation policy. *Cambridge Journal of Regions, Economy and Society*. doi:10.1093/cjres/rst007

McCann, P., Ortega-Argilés, R. (2015) Smart Specialization, Regional Growth and Applications to European Union Cohesion Policy. *Regional Studies* 49, 1291-1302. DOI: 10.1080/00343404.2013.799769

Pyka, A., Gilbert, N., Ahrweiler, P. (2007) Simulating knowledge-generation and distribution processes in innovation collaborations and networks. *Cybernetics and Systems: An International Journal* 38, 667-693.

Pyka, A., & Saviotti, P. (2002) *Innovation networks in the biotechnology-based sectors. Innovation Networks: Theory and Practice*, Cheltenham, Edward Elgar, 75-107.

Sebestyén T., Varga, A. (2013a) Research productivity and the quality of interregional knowledge networks. *Annals of Regional Science* 5, 155-189.

Sebestyén, T., Varga, A. (2013b) A novel comprehensive index of network position and node characteristics in knowledge networks: Ego Network Quality. In: Scherngell, T. (ed.): *The geography of networks and R&D collaborations*. Springer, series: Advances in Spatial Science, pp. 71-97.

Uhlbach, W-H, Balland, P-A, Scherngell (2017) R&D Policy and Technological Trajectories of Regions: Evidence from the EU Framework Programmes. *Papers in Evolutionary Economic Geography* 17.22, pp. 1-21.

Varga, A., Sebestyén, T. (2013) Innovation in Central and Eastern European Regions: Does EU Framework Program participation lead to better innovative performance? GRINCOH Working Paper 3.02
http://www.grincoh.eu/media/serie_3_knowledge__innovation__technolog/grincoh_wp_3.02_varga-sebestyen.pdf

Varga, A., Sebestyén, T. (2017) Does EU Framework Program participation affect regional innovation? The differentiating role of economic development. *International Regional Science Review* 40, 405-439.
DOI: 10.1177/0160017616642821.

Varga, A., Járosi, P., Sebestyén, T., Szerb, L. (2015) Extension and application of the GMR-Eurozone model towards the CEE regions for impact assessment of smart specialisation policies. *GRINCOH Working Papers* DOI: 10.13140/RG.2.1.5152.6567

Varga, A., Sebestyén, T., Szerb, L., Járosi, P. (2017) Economic impact assessment of smart specialisation policies with the GMR-Europe model. *RIERC working papers* 2017.4

Wanzenböck, I., Piribauer, P (2017) R&D networks and regional knowledge production in Europe: Evidence from a space-time model. *Papers in Regional Science*. DOI: 10.1111/pirs.12236

Wasserman, S., Faust, K. (1994): *Social Network Analysis: Methods and Application*. Cambridge University Press.

APPENDIX A: Heterogeneity parameters

The model is equipped with *agent-specific speed parameters*. In order to avoid overwhelming the calibration process with many separate parameters to be determined, agent speed is linked to agent size, which is exogenously given. The number of parameters to be calibrated is crucially lower this way and independent of the number of agents. Let us denote the size of agent i by W_i, which is centered on 1. Then the agent-specific speed is defined as:

$$S_i = (\bar{S} - SR) + SR \cdot W_i$$

where $\bar{S}$ is a common constant defining the average speed of agents and SR is a parameter defining roughly the elasticity of agent speed upon its size. It is easy to see that if $SR = 0$ then $S_i = \bar{S}$ so that agent speed is homogenous and corresponds to the common factor $\bar{S}$. If SR is different from zero then the agent with the average size ($W_i = 1$) moves at speed $\bar{S}$, whereas other agents move faster or more slowly. If $SR > 0$, then larger agents move faster, if $SR < 0$, then larger agents move more slowly. Note, that the agent-specific parameter S_i is static, so that it is not recalculated in each period of the model. As a result of this method, only two parameters have to be calibrated ($\bar{S}$ and SR) instead of a separate speed parameter all agents.

We define the *agent-specific counter-force* strengths in an analogous way:

$$BFP_i = (\overline{BF} - BR) + BR \cdot W_i$$

where $\overline{BF}$ is a common constant defining the average counter-force strength of agents and BR is a parameter defining the elasticity of counter-force strength upon agent size. It is easy to see that if $BR = 0$ then $BFP_i = \overline{BF}$ so that counter-force strength is homogenous and corresponds to the common factor $\overline{BF}$. If BR is different from zero then the agent with the average size ($W_i = 1$) has counter-force strength $\overline{BF}$, whereas other agents have higher or lower strength. If $BR > 0$, then counter-force is stronger for larger agents, if $BR < 0$, then counter-force is stronger for smaller agents. BFP_i is also static in the model. Again, only two parameters have to be calibrated ($\overline{BF}$ and BR) instead of a separate counter-force parameter all agents.

APPENDIX B: Deriving initial positions

An important element of our model is that the concept of network distances (pairwise distances of agents in the network on the basis of direct and indirect links and their weights) is scaled down to a 2 dimensional space. Put it differently, in order for the model to be useful, we need the 2 dimensional mapping of the agents represent network distances as close as possible. There are several ways to visualize networks in two dimensions, and most of these methods build on some similarity/dissimilarity measure between agents and try to define coordinates for the agents so that their distance in the two dimensional space fits with their dissimilarities. The most widely used method is MDS (Multidimensional Scaling), which defines a stress function between the dissimilarity measures and the Euclidean distances in the target space and then this stress function is minimized to achieve the best fit. The method can be used to reduce the dimensionality of any data to an arbitrary level, but for reasons of visualization the two-dimensional representation is the most common.

In our context, we also use MDS to arrive at the initial positions of agents in the model. The starting point is the Framework Program collaboration matrices, which determine a weighted network on the sample of regions. Then, we calculate network distances using this weighted network links, which defines a distance between any two pairs of regions. These distances serve as the dissimilarity inputs to the MDS. The stress function is simply the sum of squared differences between the pairwise network distances (observed) and the Euclidean distances in the two dimensional space. Minimizing this stress function yield the coordinates for the regions which represent quite well the original observed network distances.

Figure A.1 below represents the achieved fit of the initial positions through MDS. The picture only shows the data for 2012, but qualitatively similar pictures emerge for all other years. It is clear that there is a strong positive correlation between the observed network distances and the Euclidean distances calculated from the agent positions after the MDS. However, the method is biased, as the Euclidean distances tend to be larger than the observed network distances but never smaller.

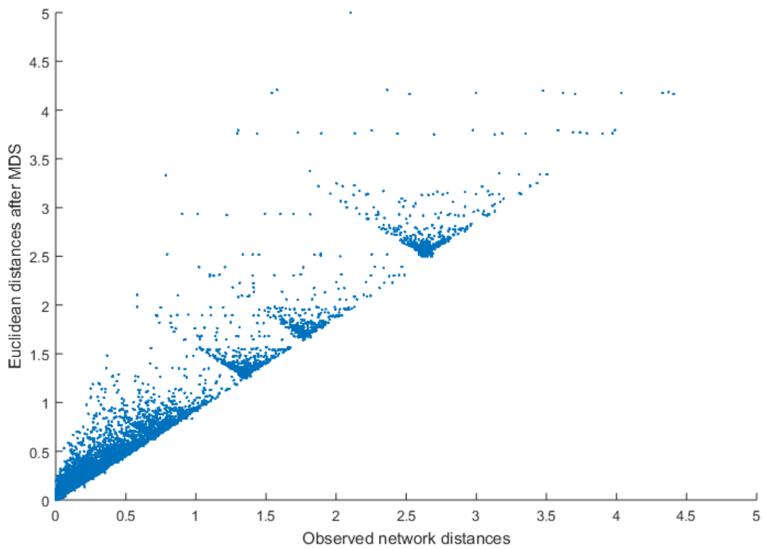

Figure A.1 – The result of MDS: correlation between observed network distances and Euclidean distances in the two dimensions after MDS, for 2012 data

In the sample there are several regions (and possibly different regions in different years), which are isolated. As their distance from others is infinity, MDS cannot handle them. In order to keep these regions in the sample we assumed that their distance from all other regions equal the largest observed distance in the given year.

APPENDIX C: The genetic algorithm

We apply a genetic algorithm to calibrate model parameters. This algorithm starts from a random population of the model parameters, i.e. a set of population members each of them is described by a random 5 element vector of the model parameters. Then, in each step the algorithm generates a new population of parameter values, based on the following principles:

- The fitness of each element in the population is calculated according to the objective function.
- The best performing elements are selected as parents.
- Worst performing elements are selected as elite and passed to the next step.

- Parents create children by mutation (random changes to elements in the parameter vector of the parent) or crossover (combining the elements in the parameter vector of the pair of parents).
- The population is replaced by the elite and the children from the current population.

The process continues until the characteristics of the population members converge to a sufficient level. The process was checked for robustness with different random initial populations. These tests showed that irrespective of the initial setting, the algorithm converges to the same optimal solution, which thus proves to be a global optimum. See for example Goldberg (1989), or Conn et al. (1991, 1997) for more details.

APPENDIX D: Agent motion

The basic principle of agent motion consists of two elements. First, the direction of agent motion is determined by the interplay between attraction- and counter-forces (i.e. their vector sum). Second, given this direction of motion, agents move according to a constant, possibly idiosyncratic speed parameter S_i.

Now assume that the position of agent i at the outset is $(x_{i,0}^1, x_{i,0}^2)$. These coordinates yield the pairwise social distances $SD_{i,j,0}$. Given the exogenous variables RD_i and CP_i, together with the parameters a_0, a_1, a_2 and a_3 in equation (2) we can determine the initial attraction values $A_{i,j,0}$ according to equation (1). On the basis of these attraction values, we can determine the desired position of each agent as follows:

$$y_{i,1}^z = \sum\nolimits_j \frac{A_{i,j,0}}{\sum_k A_{i,k,0}} \cdot x_{j,0}^z \tag{B1}$$

where $z \in (1,2)$. Here, the pair $(y_{i,1}^1, y_{i,1}^2)$ describes the position where agent i would like to arrive, so this is the direction of the attraction force. In the next step, we calculate the counter-force position, given this attraction force:

$$b_i^z = 2 \cdot x_{i,0}^z - y_{i,1}^z \tag{B2}$$

where $z \in (1,2)$. It is easy to see that this formulation provides a point in space which pins down a vector exactly in the opposite direction as the desired position. In other terms, both points $(y_{i,1}^1, y_{i,1}^2)$ and (b_i^1, b_i^2) lie at the same distance from point $(x_{i,1}^1, x_{i,1}^2)$, but exactly in opposite directions. Now define the agents' target position, taking into account both the attraction and counter-forces as

$$l_{i,1}^z = x_{j,0}^z + \left(y_{i,1}^z - x_{j,0}^z\right) + BFP_i \cdot (b_i^z - x_{j,0}^z) \tag{B3}$$

where $z \in (1,2)$. This equation states that the target location is basically the vector sum of attraction and counter-forces. One can see that given (B2) and assuming $BFP_i = 1$ we have $l_{i,1}^z = x_{j,0}^z$ so that the counter-force calibration in (B2) and the target position in (B3) really mean imply that without shocks to the model agents remain in their initial positions. The possibly idiosyncratic parameter BFP_i lies in the interval between 0 and 1, and allows to set the strength of the bacfkorce in the model.

If $BFP_i = 1$, then counter-force is of full strength, if $BFP_i = 0$, then it does not play a role at all. If $BFP_i < 1$ then agents move even in the absence of shocks.

Now assume that in period 1 there is a shock to the model. This changes the attraction values and provides a new desired position analogously to equation (B1). Together with the counter-force, the direction of motion is determined as the vector sum of the desired position and the counter-force position, adjusted with parameter BFP_i as in (B3). Now, as attraction values have changed ($A_{i,j,1} \neq A_{i,j,0}$), desired positions $y_{i,2}^z$ have also changed and $l_{i,2}^z \neq x_{j,1}^z = x_{j,0}^z$. As a consequence, agents start to move. Their movement is given by the direction $(l_{i,2}^1, l_{i,2}^2)$, together with a possibly idiosyncratic speed parameter S_i. Given the direction (target position) in (B3), the final positions of agents are given by:

$$x_{i,2}^z = x_{i,1}^z + \frac{S_i}{L_{i,2}}\left(l_{i,2}^z - x_{i,1}^z\right) \tag{B4}$$

where $L_{i,2} = \sqrt{\left(l_{i,2}^1 - x_{i,1}^1\right)^2 + \left(l_{i,2}^2 - x_{i,1}^2\right)^2}$ is the Euclidean distance between the current position and the target position of agent i. The intuition behind this formulation is that agents do not jump directly to their target position, but only proceed towards this position. The distance they move is determined by their speed S_i which gives a fixed distance in the social space and each agent moves across this distance in one model step – given that the forces of the model require them to move.

Once agents arrive to their new positions $(x_{i,2}^1, x_{i,2}^2)$, their social distances ($SD_{i,j,2}$) change, which has an effect on the pairwise attraction values ($A_{i,j,2}$). This generates new desired positions and new target positions together with the counter-force and agents move again. The same logic can be iterated until the model reaches its new steady state after the shock.

In sum, we can write the following general laws for the motion of agent between any two time periods. The desired positions are:

$$y_{i,t+1}^z = \sum\nolimits_j \frac{A_{i,j,t}}{\sum_k A_{i,k,t}} \cdot x_{j,t}^z \tag{B5}$$

The target positions are:

$$l_{i,t+1}^z = x_{j,t}^z + \left(y_{i,t+1}^z - x_{j,t}^z\right) + BFP_i \cdot (b_i^z - x_{j,t}^z) \tag{B6}$$

And finally, the actual positions are:

$$x_{i,t+1}^z = x_{i,t}^z + \frac{S_i}{L_{i,t+1}}\left(l_{i,t+1}^z - x_{i,t}^z\right) \tag{B7}$$

with the appropriate modification to $L_{i,t+1}$. Note, however, that equation (B2) with the determination of the counter-force is not put into this general setting as the counter-force is calculated only once, in the initial position. In all subsequent periods these values are used.