WORKING PAPER SERIES

2020-013

Dynamics of collaboration among high-growth firms: results from an agent-based policy simulation

Anna Varga-Csajkás
MTA-PTE Innovation and Economic Growth Research Group

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

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

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

H-7622, Pécs Rákóczi str. 80.

Phone: +36-72-501-599/23121

Dynamics of collaboration among high-growth firms: results from an agent-based policy simulation

Abstract

In this study, we introduce an agent-based model of innovation-related cooperation, which is appropriate for simulating network formation among high-growth firms (gazelles) and their partner organizations. In the model, agents represent firms or universities placed in the two-dimensional abstract social space where they are moving towards each other to find cooperation partners. In line with the gravity principle, we assume that attractiveness between two agents is affected by geographical, organizational, institutional and social proximity as well as by the mass of the two organizations. For the empirical underpinning of the model, we used survey data on the Hungarian high-tech gazelles' egocentric network that contains information about innovation-purpose cooperation in general, covering different types of formal and informal links between organizations. Part of the agent-based simulation parameters has been determined by regression analysis, the result of which shows that the geographical, social and technological distance has an impact on innovation-related cooperation. Our results are fundamentally consistent with the existing literature since we found a positive relationship between proximity and cooperation probability. For illustrative purposes, we conducted a simulation exercise that demonstrates the potential use of the model for ex-ante policy evaluation. It pointed out that a successful entrepreneurship policy, that increases the number of new firms, entering a sector that is in line with the smart specialization strategy of the region could significantly increase the density of the innovation network.

Keywords: proximities, innovation network, high-growth firms, agent-based model

Highlights

- We introduced an agent-based model which is appropriate for modeling the dynamics of network formation among high-growth firms and their partners.
- For the empirical underpinning of the model, we use unique survey data on the Hungarian high-growth firms' egocentric network, which contains information about different types of innovation-related cooperation.
- The results of policy simulation show that a successful entrepreneurship policy, that increases the number of new firms, entering a sector that is in line with the smart specialization strategy of the region could significantly increase the density of the innovation network.

JEL: C21, C63, D85, O20, O30

1 Introduction

Multiple studies have proven that geographical proximity facilitates cooperation, hence the knowledge transfer among economic actors. Marshall (1920) was the first to highlight the positive external effects of concentration of economic activity in space, but the role of geographical proximity is still a widely researched topic (Inoue et al. 2019). However, other types of proximity could also enhance cooperation and knowledge flow between organizations. According to Boschma (2005) spatial proximity is neither a necessary, nor a satisfactory condition for inter-organizational learning, and other forms of proximity may function as substitutes or as complements to the geographical one. During the last decades, it has been demonstrated in many empirical studies that other types of proximity are conducive for successful partnerships and knowledge flows (Autant-Bernard et al. 2007; Cantner and Meder 2007; Broekel and Boschma 2012; D'Este et al. 2013; Marrocu et al. 2013; Cassi and Plunket 2015; Caragliu and Nijkamp 2016; Hansen 2015; Usai et al. 2017; Capone and Lazzeretti 2018; Gui et al. 2018).

This study focuses on the network of Hungarian high-growth enterprises (gazelles) where the links between the organizations represent innovation-related cooperation. We consider innovation-related cooperation in general, covering different types of links between organizations, like information exchange, R&D cooperation, tender cooperation, innovation-purpose education or other forms. While explaining the emergence of ties, we do not concentrate only on geographical but technological, social and institutional proximity as well. Technological or cognitive proximity may facilitate the establishment of relationships as a certain set of common knowledge is required for collaboration. However, this set must be different enough to gain some novelty from cooperation. The significance of social proximity lies within the fact that personal relationships assist the emergence of trust and reduce the risk of opportunistic behavior. Institutional proximity means similar rules and norms, that promote cooperation, hence the flow of knowledge. Balland et al. (2014) emphasizes, that the relationship between proximity and knowledge networking should be analyzed in a dynamic manner. Accordingly, not only proximity affects the relationships, but also the networks re-affect the proximity dimensions. This mechanism cannot be caught by a simple regression equation, but an agent-based simulation method could be appropriate.

Although there is a rich body of literature on proximity dimensions and their effect on knowledge networks, most of the papers deal with only a few specific types of formal cooperation. Since the EU Framework Programs and co-patents are well-documented forms of knowledge relationships they can be used effectively in quantitative analysis (Scherngell and Barber 2009, Varga and Sebestyén 2017). In contrast, observing informal cooperation is more difficult and it requires primary data collection (Capone and Lazzeretti 2018). Despite the argument for the necessity of dynamic analysis of proximity and knowledge networks (Balland at al. 2014), the static approach is still dominant. In the last decade, however, there are examples for dynamic analysis as well (Balland 2012; Gui et al. 2018) and some of them use simulation methods for modeling network formation (Sebestyén and Varga 2019). Our paper bases on the work of Sebestyén and Varga (2019) who originally developed their model for the European NUTS 2 regions' Framework Program collaboration network. We borrow the mechanism from this model, but it is applied in a different setup: we use this approach in an inter-organizational context, and we initialize and calibrate it with survey data that allows us to investigate a wide range of formal and informal cooperation.

Recently, the smart specialization strategy has become the primary driver of regional innovation policy in Europe (Foray 2015; McCann and Ortega-Argilés 2015). This strategy aims to develop new specialties that build on the region's existing knowledge and competencies. Prioritization is a crucial step of this process, which means the selection of a few distinctive areas of specialization. According to Foray (2015) priority areas should be chosen through an entrepreneurial discovery process, when entrepreneurial actors explore innovative domains. However, the experience of recent years suggests that its practical implementation was more difficult than expected (Hassink and Gong 2019), and it was concluded that entrepreneurial discovery is not necessary at the priority area selection stage (Foray 2019). This new finding increases the demand for tools supporting policy planning during the prioritization process. It may imply that models, that are constructed to simulate the effect of choosing different specialization domains could play an important role in the future. Our current model contributes to this issue.

In order to empirically test the model, we use survey data on Hungarian high-tech gazelles and their partners. This survey provides information on the innovation-related cooperation patterns of these firms, their location and their activities, the content of these links and some information on their length. Using this information on the sample of high-tech gazelles in Hungary, we estimate a standard gravity equation, which provides the basis of the model. Then, a calibration procedure is used to pin down other parameters of the model which determine agents' motion in the social space. Given the estimated and calibrated model parameters, we show an illustrative simulation exercise that demonstrates the potential use of the model framework for ex-ante policy evaluation in the context of smart specialization.

2 Gazelles and networking

Birch and Medoff (1994) pointed out that a few rapidly growing firms created the majority of jobs; therefore, they named them as gazelles. However, the majority of the firm population composed of "small mice" and "elephants". The former ones are small firms unable or do not wish to grow, while the latter ones are big companies with slower growth. Due to their few numbers and noteworthy economic significance, gazelles shall be taken into consideration and put under scrutiny. Although there is no generally accepted definition for gazelles, such separations are carried out based on some growth indicators that are related to the number of employees and sales revenues. The factors behind the growth of the gazelles have been investigated in multiple studies, and some of them included networking as an explanatory variable.

It is generally assumed that networking is beneficial for small and medium-sized firms because they are provided with access to knowledge and other resources. This positive relationship can be empirically demonstrated by different measuring methods. Schoonjans et al. (2013) for instance found that the participation in formal business networks has a positive effect on the added value and assets of the firm, however, it triggers no significant increase in employment. Havnes and Senneseth (2001) demonstrated only one positive correlation out of three indicators regarding networking and growing. The volume of sales and number of employees was not bigger in the case of companies with higher networking index, but the expansion of the market was higher that contributes to the growth in the long term. Zeng et al. (2010) carried out an analysis of the impact of networking on innovation in the case of the Chinese small and medium-sized enterprises. Their findings were that the greatest positive impact on innovation performance was the cooperation between the firms, however, collaboration with intermediating and research institutions also has a positive effect.

Lechner and Dowling (2003) carried out a qualitative analysis of the egocentric network of fast-growing enterprises. They evaluated the importance of different kind of networks on the different stages of the firms' development. They found that each firm establishes its own relational mix that facilitates expansion the most. According to their results, knowledge, technology and innovation networks are important at all levels of development as opposed to some other type of networks. All in all, networking does not contribute to all of the growth indicators, but there is no doubt that it plays an important role in the development of gazelles.

In the past decade, multiple pieces of research were carried out to unearth the characteristics of the Hungarian gazelles (Papanek 2010; Csapó 2011; Némethné 2010; Békés and Muraközi 2012; Szerb et al. 2017). The most recent papers clearly indicate how difficult it is to standardize the Hungarian rapidly growing firms. It is hard to forecast which enterprises may become gazelles based on the firms' reports as they occur in all sectors and regions (Békés and Muraközi 2012). Moreover, the Hungarian gazelles often lack positive features such as innovation, export-oriented mentality or better competitiveness that are related to rapid growth, in accordance with the international literature (Szerb et al. 2017).

3 Empirical data

The data collection was performed in three rounds between 2014 and 2016. The aim of the first round was to reveal the properties of the rapidly growing firms in Hungary, the second round identified co-operative partners of the gazelles and then these primary partners were asked to specify their partner organizations. In the network thus revealed, foreign and domestic firms, educational institutions and intermediary organizations were included. The results of the survey were supplemented with other organizational data, which were collected from firm reports (http://e-beszamolo.im.gov.hu), and from the organizations' websites. A firm is considered a gazelle if it meets the following two conditions:

- the average annual growth rate of net sales revenues exceeds 20 percent over a three-year period
- at least 5 employees work in each given year.

As the goal of the survey was to measure the domestic high-growth enterprises, two additional properties were needed to be sampled: Hungarian-based firm with a minimum of 75% Hungarian ownership. In the database provided by Opten Informatics Ltd., 4037 firms met this definition. From this population, 404 firms were sampled during the layered sampling performed according to agglomeration areas. This sample was reduced according to two aspects. On the one hand, firms were filtered out that did not report any innovation cooperation, and the other hand, those firms that did not belong to the "high-tech" or "knowledge-intensive" sectors. As a result, a sample of 80 high-tech gazelles was generated. In the second round, 55 of the 80 firms finally gave valuable responses. The respondents identified 94 organizations that we call the primary partners. In the third round, these partners were questioned, and 53 of them gave a valid answer. The respondents reported a total of 183 partners, who form the group of secondary partners of the gazelles.

As a result of the survey, we theoretically get a directed graph, but we will treat the network of gazelles as an undirected graph. If one party states that there is a relationship between them, then the direction and the strength of the relationship is not interpreted, we only record that the tie is established. The gazelle's network is demonstrated by a binary symmetric matrix, where the elements represent the existence of the relationships of the organizations.

The set of partners were restricted to Hungarian firms and higher education institutions since the necessary additional information was not available for all types of organizations. In the case of universities, partners were given at different organizational levels (institute, department, faculty, university), which we aggregated to the level of the university, so the data could become comparable. Thus, a total of 207 agents remained in the examination, of which 102 form a weakly connected component, while the other organizations are located in smaller separate groups or are isolated, as shown in Figure 1. For technical reasons, we restricted the sample for the connected part of the network, so we applied the model for 102 agents. It was necessary because the social distance is a basic concept of the model and we are not able to interpret the distance between unconnected agents.

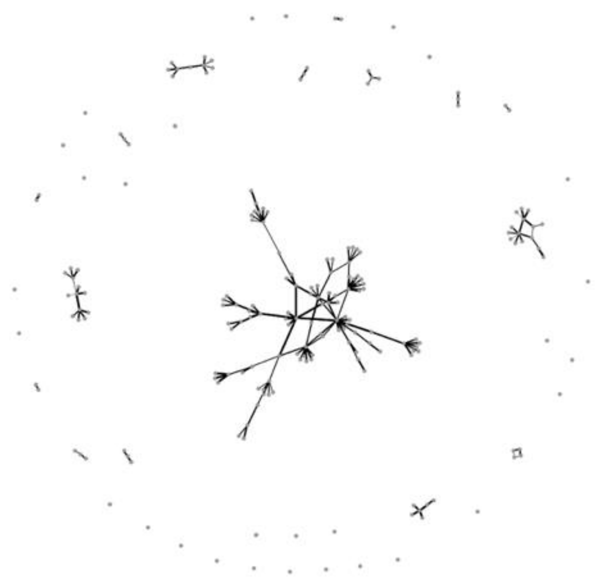

Figure 1: The innovation network of Hungarian high-tech gazelles

The results of the survey reveal that in most cases the content of the relationship was the information exchange, there was a smaller number of cases of R&D cooperation or tender cooperation, while only a couple of respondents indicated that innovation-purpose education was the content of their cooperation.

4 The agent-based model

Agent-based modeling is one of the potential techniques for modeling network formation besides random graph models (Erdős and Rényi 1959, Watts and Strogatz 1998, Barabási and Albert 1999) and strategic models of network formation (Jackson 2005). The first one takes a probabilistic view on network formation and this kind of model can explain many phenomena observed in a network topology. Strategic models start from individual incentives for link-formation that leads to the emerging network. While taking into account individual choice, these models remain stylized. The main advantage of agent-based models over the other two types is its ability to be empirically calibrated and validated which makes it appropriate for ex-ante policy simulations. The SKIN model (Gilbert et al. 2001; Ahrweiler et al. 2004; Pyka et al. 2007) is a well-known agent-based model that contains network formation. It was the base of many empirically calibrated studies that include the whole innovation system but the network of actors is a secondary interest (Korber and Paier 2013) like in other agent-based innovation models (Pyka and Saviotti 2002; Heshmati and Lens-Cesar 2013; Paier et al. 2017).

The current agent-based model builds on the work of Sebestyén and Varga (2019) whose original model was developed for the European NUTS 2 regions' knowledge network. The major elements of the model are the social space where agents are located and the gravitational force that drives their motion. Moreover, agents have heterogeneous attributes that also affect their behavior thereby the emerging network.

Agents are placed in the social space according to their position in the initial innovation network. The distance between them is measured by the length of the shortest path which can be regarded as social distance. These multidimensional network distances can be represented in two dimensions with the help of an appropriate algorithm therefore we use Multidimensional Scaling in order to get these 2D positions. From the initial positions, agents start to move towards each other according to their pairwise attraction to find cooperation partners.

Together with the social distance, the size and the proximity of actors are assumed to affect their mutual attractiveness which expresses their willingness to cooperate. The attraction force is determined by the gravity equation which contains the mass of the agents ($M_{i,t}$ and $M_{j,t}$) and the pairwise distance between them in geographical ($GD_{i,j,t}$), technological ($TP_{i,j,t}$), social ($SD_{i,j,t}$) and institutional ($IP_{i,j,t}$) respect.

$$A_{i,j,t} = f(M_{i,t},\ M_{j,t},\ GD_{i,j,t},\ TP_{i,j,t}, SD_{i,j,t},\ IP_{i,j,t})$$

According to the gravity principle, the mass has a positive and the distance has a negative impact on the attraction force between two agents. The only endogenous variable on the right-hand side is the social distance. During the simulation, the attraction force changes only if the positions of agents in the social space change. If the attraction force between two organizations reaches the threshold, they will link up.

When agents chose their target position, they take into account only a subset of the potential partners because of cognitive limitations. It is represented by the length of the partner list that expresses the number of potential partners that agents follow when they chose their target positions. Besides the attraction force, there is a counterforce that ensures the stationarity of the model. It represents the cost of forming and maintaining a relationship.

The target position is not achieved automatically but agents start to move toward that point with an agent-specific constant speed. We assume that the speed of agents is linked to the size of them. Besides, we assume that agents are different with respect to the cost of their link formation, which also depends on the agents' size. It results that in both cases two parameters are necessary to determine the agent-specific value: a common counterforce/speed parameter and an elasticity parameter. These are determined through a calibration process.

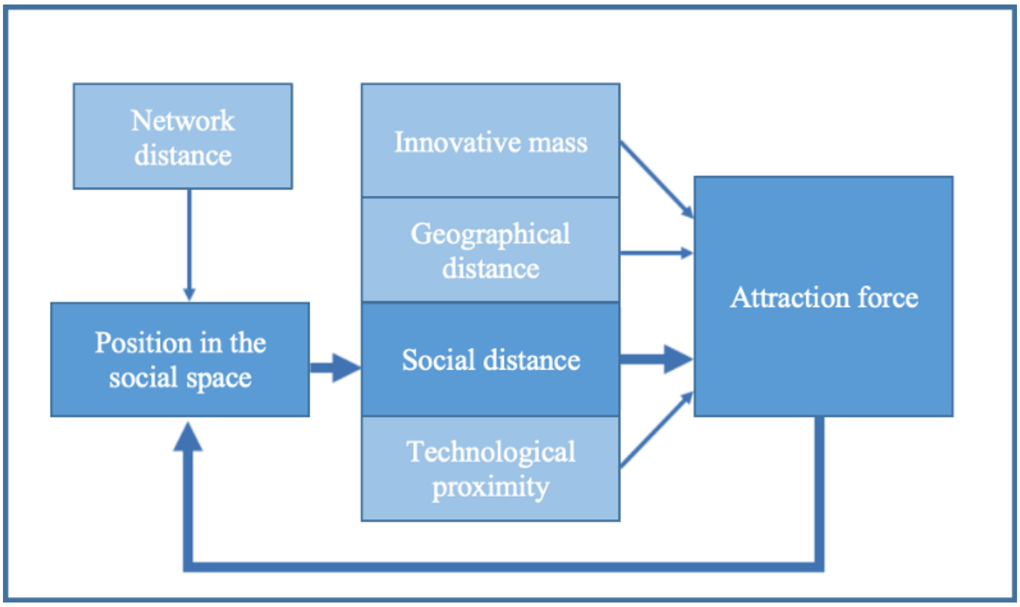

Figure 2: How the simulation model works

We can sum up the working of the simulation model as follows: the initial positions of the model represent the location of the agents in the social space and their social distance. The innovative mass, the geographical distance, the technological proximity, and the social distance determine the attractiveness values through the gravity equation. If any of these variables change, agents start to move, which feeds back into the social distance through changing positions. The modified social distances again affect attractiveness, and the system keeps moving for a while. Later on, it settles down, because of its stationary property.

5 The gravity equation

Regression analyses were conducted in order to determine the parameters of the gravity equation. Newton's gravitational law originally describes the attraction force between physical bodies, but the principle of gravitation can also be found in the social processes. It became an analytical tool of economics by the explanation of international trade, as it is a clear idea that the volume of international trade is positively influenced by the size of the two countries, while the distance between them has a negative influence. This analogy can be useful also in other contexts, such as the emergence of innovation-related cooperation. This principle indirectly appears in most of the studies examining the proximity dimensions, but in some articles, a gravitational equation is explicitly specified. For example, Maggioni and Uberti (2009) and Scherngell and Barber (2009) both examined the R&D cooperation among European regions with the help of an econometric model based on a gravity equation.

The dependent variable is the innovation-related cooperation that has two possible values. It is equal to one if there is a relationship between the two organizations and zero if they are not connected directly.

Mass: the gravitational force is higher if the mass of the two bodies is bigger. In this case, we interpret the number of existing links of a node as a proxy of mass. We suppose that the higher the number of partners the more attractive will be the agent.

Geographical distance: Geographical distance (GeoDist) is captured by the Euclidean distance between the organizations' headquarters. In the case of high-tech gazelles and their primary partners, the address of the headquarters is known from the responses. The secondary partners have not been questioned, but either the postal address or the website had to be given by the nominator, so their addresses could be specified as well. Using a geocoding program, we identified the latitude and longitude coordinates, from which we calculated Euclidean distance, and this was included in the regression model.

Technological proximity: Technological or cognitive proximity is usually measured by the overlap in patent portfolios (Cantner and Meder 2007, Cassi and Plunket 2015), but in some studies, this dimension is captured by the similarity of the economic activity of the two agents

(Usai et al. 2017). Since there were many organizations in the sample that did not have a patent, we chose the latter solution. Technological proximity (NACE1-4) is expressed by four dummy variables in our study. The value is 1 if the two organizations are in the same category according to the 1,2,3 or 4-digits NACE (Nomenclature statistique des activités économiques dans la Communauté européenne) codes. Accordingly, if all 4 digits in the NACE code of the primary activities of the two organizations is the same, then the technological proximity is very strong between them, and it is the weakest if only the first digit is the same. The reference group is when they are different even on the one-digit level. This measurement method provides an opportunity to demonstrate the proximity paradox, accordingly to which the technological proximity has a positive impact on cooperation, but only to a certain extent. After that point, it may hinder innovation-related cooperation.

Social distance: Social distance is measured in accordance with a position in a social network emerging on the basis of earlier innovation cooperation (Autant-Bernard et al. 2007; Balland 2012; Usai et al. 2017). It is possible because the date of the establishment of relationships is revealed by the database. If a tie is established in 2014 or before, the two organizations are considered having a common history. These former relationships are captured by a symmetrical unweighted matrix, whose elements are equal to one if there was a relationship between the two organizations even in 2014 and it is equal to zero if there was no cooperation between them. From this network, we calculated the geodesic distance per pair, which is the length of the shortest path between the two edges. These network distances cannot be interpreted directly in the conventional Euclidean space, so we use Multidimensional Scaling (MDS) in order to get 2D positions in the social space. The distance between these positions will enter the model as social distance.

Institutional proximity: An interpretation of the institutional proximity is, that organizations with the same status are closer to one another (Ponds et al. 2007; Balland 2012; Cassi and Plunket 2015; Usai et al. 2017). In the current study, belonging to the same organization type is considered as institutional proximity and it is measured by a dummy variable. We consider two organization types: firms and universities. The value of the variable is one if both agents are firms or universities and it is zero if they are different.

There are two main approaches in the literature to reveal the relationship between different proximity dimensions and innovation-purpose cooperation. If the number of cooperative projects is known, a Poisson or binomial count data model can be used (Hoekman et al. 2009). When, however, the dependent variable cannot be counted, but only the existence of the connection or its intensity is known, then discrete choice models are applied (Autant-Bernard et al. 2007; Paier and Scherngell 2011). In our case, the unit of observation is the organization-pair and the dependent variable is the relationship between the two organizations, so we have a binary choice model. The likelihood of the emergence of a relationship is explained by a binary logit model:

$$Conn_{i,j} = P_i = \frac{1}{1 + e^{-z_{i,j}}}$$

$$z_{i,j} = \beta_0 + \beta_1 \cdot \left(InnMass_i + InnMass_j\right) + \beta_2 \cdot GeoDist_{i,j} + \beta_3 \cdot TechD1_{i,j} + \beta_4 \cdot TechD2_{i,j} + \beta_5 \cdot TechD3_{i,j} + \beta_6 \cdot TechD4_{i,j} + \beta_7 \cdot SocDist_{i,j} + \beta_8 \cdot SameOrg_{i,j}$$

As a starting point, we included only the mass variable, then we introduced the different proximity and distance dimensions after each other. Table 1 presents a summary of the regression results.

	(1)	(2)	(3)	(4)	(5)
number of observations	10404	10404	10404	10404	10404
constant	−5.4136*** (0.135484)	−4.94132*** (0.153972)	−5.23328*** (0.167774)	−2.47979*** (0.210755)	−2.50442*** (0.284908)
$Mass_i$	0.262912*** (0.0180924)	0.25913*** (0.0182462)	0.263822*** (0.0186455)	0.231947*** (0.0204122)	0.232636*** (0.0211122)
$Mass_j$	0.262912*** (0.0180924)	0.25913*** (0.0182462)	0.263822*** (0.0186455)	0.231947*** (0.0204122)	0.232636*** (0.0211122)
$GeoDist_{ij}$		−0.40340*** (0.0778016)	−0.3685*** (0.0782538)	−0.42813*** (0.078039)	−0.42820*** (0.0780462)
$TechD1_{ij}$			1.54396*** (0.23086)	−0.0546157 (0.261105)	−0.0608071 (0.265484)
$TechD2_{ij}$			1.95864*** (0.450445)	1.88331*** (0.497809)	1.8787*** (0.499292)
$TechD3_{ij}$			−0.315881 (0.73058)	−1.12480 (0.753739)	−1.11038 (0.762052)
$TechD4_{ij}$			0.594376***	0.588778***	0.585361***

			(0.205043)	(0.216449)	(0.218058)
$SocDist_{ij}$				−1.16957*** (0.088498)	−1.16983*** (0.0885152)
$SameOrg_{ij}$					0.026437 (0.205724)
Adjusted McFadden R^2	0.151275	0.164662	0.184875	0.321731	0.320787
Log-likelihood	−888.7306	−873.6646	−703.6292	−703.6375	−703.6292

*Table 1: Regression results from estimating the gravity model. Dependent variable: innovation-related cooperation between agent-pairs. *** significant at the 0.001 significance level, ** significant at the 0.01 level, * significant at the 0.05 level.*

The values of the coefficients were determined by a maximum likelihood estimate and below them, in brackets, the standard errors are shown. The absolute value of the coefficients is not informative in the logit model, but its sign and significance can be interpreted similarly to the estimated results of the ordinary least squares method. We have chosen model (4) for further investigation. It has the highest R-square (0.32) which is considered moderate explanatory power. The results show that geographical, social and technological distance has an impact on innovation-related cooperation. As expected, we found that the closer they are in the sense of different dimensions, the higher the chance for cooperation between them. Organizational proximity was the only investigated dimension that was not significant in our analysis. It should be noted that the regression was conducted on a selective sample thus, our econometric results may be biased, therefore results could not be generalized. Nonetheless, it helps us to determine the parameters of the agent-based model.

6 Model calibration

A calibration procedure is used to specify other parameters of the model which determine agents' motion in the social space. This calibration process takes a pre-survey situation as the starting point (the collaboration patterns of which we have some information) and then search those parameter combinations which drive the collaboration patterns the closest to the observed one. Since only cross-sectional data were available, we had to construct a hypothetical starting point for the calibration. To this end, we placed the agents in the social space according to their relationships that existed already in 2010.

description	parameter	range	optimal value
common speed parameter	$\bar{S}$	[0;0,1]	0.076802
common counterforce parameter	$\overline{BF}$	[0;1]	0.64178
length of partner list	AP	[0;101]	5
speed elasticity parameter	SR	[-1;1]	0.010456
counterforce elasticity parameter	BR	[-1;1]	-0.084521

Table 2: The values of the calibrated parameters

The calibration results show, that agents take into consideration only the five most attractive potential partners when they choose their target position in the social space. The speed elasticity parameter is positive, which means that bigger agents move faster. In contrast, agents' size influences negatively the counterforce, which indicates that link formation is more costly for smaller agents.

7 Policy simulation

In the following section, we show that such a model can be useful in the prioritization stage of the smart specialization strategy (Foray 2015; McCann and Ortega-Argilés 2015). The experience of recent years suggests that although entrepreneurial discovery is a key component of S3 strategy building, it is not necessary for the prioritization phase (Foray 2019). This new finding increases the demand for tools supporting policy planning, therefore such models, that are constructed to simulate the effect of prioritization could play an important role in the future. With the help of the current agent-based model, one is able to simulate the likely effect of a policy intervention that supports new business formation in a potentially prioritized field. We concentrate on the question that how prioritization of a certain area effects the density of the innovation network. To this end, we compare two different scenarios to investigate how the density of the innovation network after a policy intervention differs from the one without intervention.

The policy simulation starts from the observed position in 2016, and it predicts the number of ties in the future. In the baseline scenario, only the 102 observed organizations are included. In the policy scenario, we suppose that after a successful policy, three new firms enter the prioritized sector. These new firms are located in Budapest and each of them has one partner from the same sector and the same region. We have chosen architectural and engineering activities as a priority area because it is one of the intelligent technologies of the region of Budapest and Pest county according to the National Smart Specialization Strategy of Hungary

(National Innovation Agency 2014). With this simulation, we can evaluate whether choosing this field was a proper decision in the sense of its network effect.

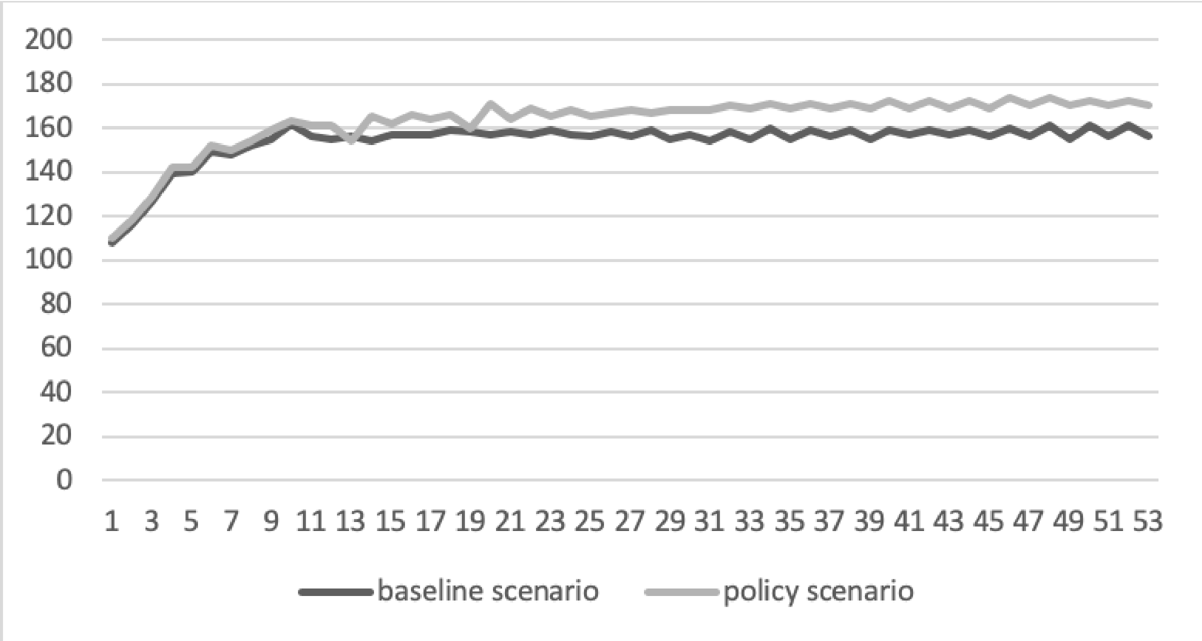

Figure 3: The number of ties in the network of the Hungarian high-tech gazelles

From Figure 3, we can see that the entry of three new firms increases the number of connections in the long run which is not surprising because the number of potential ties is higher in the policy scenario. Table 3 shows that not only the number of links rises but the density of the network as well. The simulation confirms that selecting architectural and engineering activities as priorities looks a favorable choice, because it increases the density of the innovation network, that leads to more opportunity for knowledge spillover.

	at the beginning of the simulation	**at the end of the simulation**
baseline scenario	N = 102	N = 102
	E = 108	E =157
	E_{max} = 5151	E_{max} = 5151
	D = 0,0210	**D = 0,0305**
policy scenario	N = 105	N = 105
	E =111	E = 171
	E_{max} = 5460	E_{max} = 5460

	D = 0.0203	**D = 0.0313**

Table 3: The characteristics of the innovation network in the baseline and the policy scenario

8 Conclusions and the limitations of the study

In the current study, we have simulated the relationship between the proximity dimensions and the formation of an innovation network. Part of the agent-based simulation parameters has been determined by regression analysis, the result of which shows that the geographical, social and technological distance has an impact on innovation-related cooperation. As expected, we found that the closer they are in the sense of different dimensions, the higher the chance for cooperation between them. Organizational proximity was the only investigated proximity dimension that was not significant in our analysis. This study contributes to the literature at three points. First, in contrast to most of the studies investigating the innovation-related cooperation and the role of proximity, we have taken into account not only formal but informal cooperation as well. Second, we introduced an agent-based model which is appropriate for modeling the dynamics of network formation on the bases of different proximity dimensions. Third, we demonstrated how policy simulation can help decision-making in the context of smart specialization paradigm. The policy simulation points out that a successful entrepreneurship policy, that increases the number of new firms, entering in a sector which is in line with the smart specialization strategy of the region could significantly increase the density of the knowledge network.

The limitation of the study is that only cross-sectional data were available therefore we had to construct a hypothetical starting position for the calibration. In addition, we conducted the regression analysis on a selective sample thus, our econometric results are probably biased. With representative sample and panel data, we could gain generalizable results and more relevant policy simulations.

References

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

Autant-Bernard C, Billand P, Frachisse D, Massard, N (2007) Social distance versus spatial distance in R&D cooperation: empirical evidence from European collaboration choices in micro and nanotechnologies. *Papers in Regional Science*, *86*(3), 495-519.

Balland PA (2012) Proximity and the evolution of collaboration networks: evidence from research and development projects within the global navigation satellite system (GNSS) industry. *Regional Studies*, *46*(6), 741-756.

Balland PA, Boschma R, Frenken K (2014) Proximity and innovation: From statics to dynamics. *Regional Studies*, *49*(6), 907-920.

Békés G, Muraközy B (2012) Magyar gazellák. A gyors növekedésű vállalatok jellemzői és kialakulásuk elemzése [Hungarian gazelles: what makes a high-growth firm in Hungary?]. *Közgazdasági Szemle (Economic Review-monthly of the Hungarian Academy of Sciences)*, *59*(3), 233-262.

Birch DL, Medoff J (1994) Gazelles. In: Solomon L, Levenson A, editors. Labor Markets, Employment Policy and Job Creation. Economics and Education. pp. 159–167.

Boschma R (2005) Proximity and innovation: a critical assessment. *Regional studies*, *39*(1), 61-74.

Broekel T, Boschma, R (2012) Knowledge networks in the Dutch aviation industry: the proximity paradox. *Journal of Economic Geography*, *12*(2), 409-433.

Cantner U, Meder A (2007) Technological proximity and the choice of cooperation partner. *Journal of Economic Interaction and Coordination*, *2*(1), 45-65.

Capone F, Lazzeretti L (2018) The different roles of proximity in multiple informal network relationships: evidence from the cluster of high technology applied to cultural goods in Tuscany. *Industry and Innovation*, *25*(9), 897-917.

Caragliu A, Nijkamp P (2016) Space and knowledge spillovers in European regions: the impact of different forms of proximity on spatial knowledge diffusion. *Journal of Economic Geography*, *16*(3), 749-774.

Cassi L, Plunket A (2015) Research collaboration in co-inventor networks: combining closure, bridging and proximities. *Regional Studies*, *49*(6), 936-954.

Csapó K (2011) The Characteristics and Development Possibilities of Fast-Growing Small and Medium-Sized Companies in Hungary. *ERENET PROFILE*, 16.

D'Este P, Guy F, Iammarino S (2013) Shaping the formation of university–industry research collaborations: what type of proximity does really matter?. *Journal of economic geography*, *13*(4), 537-558.

Erdős P, Rényi A (1959) On random graphs. I. Publicationes Mathematicae, 6, 290–297.

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

Foray D (2019) In response to 'Six critical questions about smart spezialisation'. *European Planning Studies*, *27*(10), 2066-2078.

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

Gui Q, Liu C, Du D (2018) International knowledge flows and the role of proximity. *Growth and Change*, *49*(3), 532-547.

Hansen T (2015) Substitution or overlap? The relations between geographical and non-spatial proximity dimensions in collaborative innovation projects. *Regional Studies*, *49*(10), 1672-1684.

Hassink R,Gong H (2019) Six critical questions about smart specialization. *European Planning Studies*, *27*(10), 2049-2065.

Havnes PA, Senneseth K (2001) A panel study of firm growth among SMEs in networks. *Small business economics*, *16*(4), 293-302.

Hesmati A, Lenz-Cesar F (2013) Agent-based simulation of cooperative innovation in R&D. *Research Evaluation, 22*, 15– 29. doi:10.1093/reseval/rvs032

Hoekman J, Frenken K,Van Oort F (2009) The geography of collaborative knowledge production in Europe. *The Annals of Regional Science, 43*(3), 721-738.

Inoue H, Nakajima K, Saito YU (2019) Localization of collaborations in knowledge creation. *The Annals of Regional Science*, *62*(1), 119-140.

Jackson MO (2005) A survey of models of network formation: Stability and efficiency. In G. Demange, & M. Wooders (Eds.), Group formation in economics: Networks, clubs, and coalitions (pp. 11–49). Cambridge: Cambridge University Press.

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 T. Scherngell (Ed.), The geography of networks and R&D collaborations (pp. 353–371). Berlin: Springer.

Lechner C, Dowling M (2003) Firm networks: external relationships as sources for the growth and competitiveness of entrepreneurial firms. *Entrepreneurship & regional development*, *15*(1), 1-26.

Maggioni MA, Uberti, TE (2009) Knowledge networks across Europe: which distance matters?. *The Annals of Regional Science, 43*(3), 691-720.

Marrocu E, Paci R, Usai S (2013) Proximity, networking and knowledge production in Europe: What lessons for innovation policy?. *Technological Forecasting and Social Change*, *80*(8), 1484-1498.

Marshall A. (1920/1890). Principles of Economics An Introductory Volume. 8, Macmillan and Co., London.

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

National Innovation Agency (2014) National Smart Specialization Strategy: https://nkfih.gov.hu/szakpolitika-strategia/nemzeti-strategiak/nemzeti-intelligens# Accessed: 5 September 2019

Némethné P (2010) Hol szökellnek a magyar gazellák?–A dinamikusan növekvő kis-és középvállalatok néhány jellemzője. *Vezetéstudomány/Budapest Management Review*, *41*(4), 32-44.

Paier M, Dünser M, Scherngell T, Martin, S (2017) Knowledge creation and research policy in science-based industries: an empirical agent-based model. In *Innovation Networks for Regional Development* (pp. 153-183). Springer, Cham.

Paier M, Scherngell T (2011) Determinants of collaboration in European R&D networks: empirical evidence from a discrete choice model. *Industry and innovation*, *18*(1), 89-104.

Papanek G. (2010) A gyorsan növekvő magyar kis-és középvállalatok a gazdaság motorjai. *Közgazdasági Szemle (Economic Review-monthly of the Hungarian Academy of Sciences)*, *57*(4), 354–370.

Ponds R, Van Oort F, Frenken K (2007) The geographical and institutional proximity of research collaboration. *Papers in regional science*, *86*(3), 423-443.

Pyka A, Gilbert N, Ahrweiler P (2007) Simulating knowl- edge-generation and distribution processes in innovation collab- orations and networks. Cybernetics and Systems, 38, 667–693. doi:10.1080/01969720701534059

Pyka A, Saviotti P (2002) Innovation networks in the biotechnol- ogy-based sectors. Innovation networks: Theory and Practice (pp. 75– 107). Cheltenham: Edward Elgar.

Scherngell T, Barber MJ (2009) Spatial interaction modelling of cross-region R&D collaborations: empirical evidence from the 5th EU framework programme. *Papers in Regional Science, 88*(3), 531-546.

Schoonjans, B., Van Cauwenberge, P., & Vander Bauwhede, H. (2013) Formal business networking and SME growth. *Small Business Economics*, *41*(1), 169-181.

Sebestyén T, Varga A (2019) Knowledge networks in regional development. An agent based model and its applications. *Regional Studies, 53*(9)

Szerb L, Komlósi É, Varga A (2017) Gyors növekedésű vállalatok Magyarországon: Az innovatív, a rejtélyes és a virtuális gazellák. [Fast-growing businesses in Hungary: the innovative, mysterious and virtual gazelles.] *Közgazdasági Szemle (Economic Review-monthly of the Hungarian Academy of Sciences*, *64*. 476-506.

Usai S, Marrocu E, Paci R (2017) Networks, Proximities, and Interfirm Knowledge Exchanges. *International Regional Science Review*, *40*(4), 377-404.

Varga A, Sebestyén T (2017) Does EU Framework Program participation affect regional innovation? The differentiating role of economic development. *International Regional Science Review,* 1-35 DOI: 10.1177/0160017616642821.

Watts DJ, Strogatz SH (1998) Collective dynamics of 'small- world' networks. *Nature*, *393,* 440–442. doi:10.1038/30918

Zeng SX, Xie XM, Tam CM (2010) Relationship between cooperation networks and innovation performance of SMEs. *Technovation*, *30*(3), 181-194.