Department of Economics and Regional Studies
University of Pécs, Faculty of Business and Economics

WORKING PAPER SERIES

INNOVATION AND DIVERSITY IN A DYNAMIC KNOWLEDGE NETWORK

Tamás Sebestyén

2010/1.

March, 2010.

Innovation and Diversity in a Dynamic Knowledge Network

Tamás Sebestyén
Department of Economics and Regional Studies
University of Pécs, Faculty of Business and Economics
Tel.: +36-72-501-599 / 3150
e-mail: sebestyent@ktk.pte.hu

March, 2010.

Keywords: knowledge network, innovation, knowledge heterogeneity, alliance formation, cluster evolution

JEL O31, O33, D21, L19

Abstract:

In this paper we examine the evolution of network formation. We present a model in which companies in an industry can innovate alone or in alliance with others. Alliance formation is based on the cognitive distance of companies: if two companies form an alliance, their probability of success in innovation depends on their proximity in knowledge space, that is, their cognitive distance. Knowledge, on the other hand, is modelled in two dimensions: breadth and depth. The main results of our analysis are that in the present setting heterogeneity decreases among companies whilst innovation can increase and decrease also, depending on the initial parameters of the industry's knowledge endowment. The model also reveales the importance of external shocks in maintaining heterogeneity and concludes with a possible typology of cluster evolution among the dimensions of heterogeneity and innovativeness.

Innovation and Diversity in a Dynamic Knowledge Network*

1. Introduction

The evolution of economic clusters, or, in a wider sense, regional economic activity and agglomeration, has become a frequently analysed area in contemporary economics. Based on the work of Krugman (1991) and the more practice-oriented contribution of Porter (1990) the field has grown enormously in recent years. Although very diversified, the roots of this line of economic thinking goes back to classical economics, namely The Wealth of Nations by Adam Smith (Smith, 1776). The main interest in Simth's work was the accumulation of wealth and he argued that economic growth stems from the division of labour and knowledge accumulation, which are two mutually reinforcing processes.

In modern economics Robert Solow put the question of growth into the focus of analysis with his neoclassical growth model (Solow, 1956). The main contribution of this model is that, if we include solely labour and physical capital into the set of production factors, economic growth can be only temporary: per capita production can not increase in the long-term. However, if (exogenous) technological change is integrated into the model (knowledge is included in the set of production factors), it can be shown that – given several assumptions – the long run growth rate of per capita output equals that of technological level: this is where long-term economic growth stems from. Later, endogenous growth theory tried to trace the specific role for knowledge in sustainable economic growth. (Arrow, 1962, Romer, 1990, Grossman and Helpman, 1991, Aghion and Howitt, 1992, Silverberg and Lehnert, 1994)

Leading directly from here is the question of how new knowledge is created and how it diffuses in the economy. Whilst the question of how new knowlegde is created remains largely unexplained, the literature on knowledge spillover treats the latter issue quite thoroughly. Jaffe (1986) proves that innovation activity is not isolated in the economy,

*The author acknowledges invaluable help from comments and notes at the Department Seminar of the Department of Economics and Regional Studies at the University of Pécs as well as comments and discussion at the 3th International Conference on Innovation and Management held in Maastricht and 6th European Meeting on Applied Evolutionary Evolutionary in Jena. Financial support of the Faculty of Business and Economics of University of Pécs and the DIME european network of excellence is also acknowledged.

but innovating companies use knowledge generated in other points of the economy as inputs to their knowledge generation processes. This clearly shows that knowledge is diffusing. Other studies, however, revealed that these knowledge spillovers are spatially bounded (Acs et al., 1992; Jaffe et al., 1992; Anselin et al., 1997). According to their findings, companies are more efficient in exploiting knowledge coming from other companies, universities and research institutes if they are located close to these sources. Jaffe et al. (1992) show, that the localised effects of spillovers die out over time, although this process is very slow.

The reasons why knowledge diffusion is spatially bounded are generally attributed to the tacit nature of knowledge. The distinction between tacit and codified knowledge comes from Polanyi (1966), although in contemporaneous literature its meaning and use is somewhat blurred (De Carvalho et al., 2006). Codified knowledge is easily formalised, and so easily communicated over long distances without loss of information or meaning. Tacit knowledge, however, can not be formalised, and so its transfer requires direct face-to-face interactions between the sender and the reciever, which in turn needs spatial proximity among agents. Hence tacit knowledge mainly spreads locally. On the other hand, companies can save travelling and other transaction costs if they locate close to each other in order to exploit tacit knowledge coming from other companies or institutions. This logic contains the conclusion that spatial concentration (or the boundedness of knowledge spillovers) is only neccesary in those industries where new knowledge is a critical competitive factor (Audretsch and Feldman, 1996a), and where knowledge is basically tacit (Sorenson, 2005).

These findings refocus our attention on local economic activity and the dynamics of agglomerations and knowledge networks. For companies relying on new knowledge as the source of their competitive advantage, it seems clearly useful to locate close to each other, and to establish strong linkages among each other in order to gain easy and immediate access to new knowledge. The resulting networks (or clusters) show tight cooperation, quick knowledge diffusion and a high level of innovativeness. Obviuosly, clustering tendencies and advantages from clustering differ among industries as these industries differ in the extent to which access to new knowledge is important, in the tacitness of knowledge used and whether the diffusion is mediated by knowledge sharing or knowledge broadcasting processes (Cowan, 2006).

Clusters and knowledge networks have been studied widely in the literature. Here we do not provide a thorough review of the literature, but we would mention some summarising works. Clusters are basically studied empirically. Numerous case studies present succesful and unsuccessful cases of local economic agglomerations (clusters). Karlson (2008) collects contemporary work on cluster research. However, Malmberg and Power (2006) emphasise that the notion of clusters is still not clearly defined: the actual characterization largely depending on the focus of different research. Although there are numerous empirical works, modelling clusters is still work-in-progress. Apart from models of new economic geography (initiated by Krugman, 1991), which miss an explicit dynamic structure, Brenner (2004) gives a simple but interesting dynamic model of local industrial clusters.

Regarding knowledge networks, once again many empirical studies exist, mainly based on R&D cooperation databases as well as patent statistics (see Giuliani 2007, Varga and Parag, 2009, for instance). These studies are at the beginning of an important reserch avenue and they try to trace a role for network characteristics relating innovation. On the other hand, modelling networks also have a wide literature. Cowan (2006) gives a good summary of network models used in innovation theory. An important element in network modelling is simulation which stems from the fact that even very simple network models are analitically intractable. In this paper we follow the latter approach by analysing knowledge networks through network modelling and executing simulations to gain an insight into the processes of the model.

However, there is another important line in contemporary literature on innovation: that of heterogeneity and complementarity. According to this, a cluster becomes dynamic and innovative through the diversity of technologies, production processes employed and product variants produced by the companies in the industry (cluster). Using the terminology of the literature, we can say that companies operate on different knowledge bases (Pavitt, 1998; Nelson, 1998). The diversity of these knowledge bases gives real innovation potential: the combination of different elements, recognising complementarities reveal a wide space for innovation based on association. The strength of innovative clusters lies in frequent interactions among diverse knowledge bases which is mediated by the increasing number of links between companies. Hence, advantages in diversity can be exploited rapidly.

On the other hand, it is understood that it is not heterogeneity itself that contributes to innovativeness, but rather complementarities in knowledge. This leads to the recognition that innovation does not grow indefinetly as heterogeneity increases. Rather, there exists an inverted U-shaped relationship between the two (Nooteboom, 1999). Too little heterogeneity means that companies know mainly the same, and so there is no room for combinig ideas: innovation activity is low. On the other hand, if heterogeneity is too high, companies do not even understand each other: they can not communicate effectively, and so innovation ceases. In our study we take this inverted U-shaped relationship between knowledge heterogeneity and innovativeness as given. For a detailed discussion on the topic see for example Cohen and Levinthal (1990), Nooteboom (2004), Wuyts et al. (2003), Cowan et al. (2007). Boschma and Iammarino (2008) emphasize the role of related variety in innovative performance. They add to the question of heterogeneity the important insight that diverse knowlede bases need to be somewhat related, complementary, in order to be useful for innovative activity. This is in line with the considerations about possible negative effects of heterogeneity.

This argument about heterogeneity has an important implication on cluster evolution. Companies, who interact frequently and form joint research alliances, gradually lose their diversity as they learn from each other. After a while, companies know the same having absorbed everything possible from the cluster, and so diversity disappears taking the wind out of innovation's sail (Cowan et al. 2007). We can therefore expect the cluster to have a special lifecycle in which the initial phase charaterised by dynamism and innovativeness is followed by a mature and declining phase when innovation and dynamism erode (Audretsch and Feldman, 1996b, Pouder and St. John, 1996).

On the contrary, however, there is some evidence that innovation and heterogeneity can be sustained in the long run. These results are in line with the previous findings: if heterogeneity goes hand-in-hand with innovation, the two must be observed jointly, or not at all. Regarding the lifecycle mentioned above, it seems that, in order to maintain innovation, heterogeneity must be rebuilt in the cluster. This can be easily done by chanelling new knowledge into the cluster from outside (Baum et al., 2003; Cowan, 2006). However, Knott (2003) builds and analyses a model in which heterogeneity and innovation are sustained inside the industry without extra-cluster linkages.

In this study we focus on the three issues introduced so far, namely innovation, locality in networks and heterogeneity in knowledge. In order to do this we built a model of an innovation network and then executed simulations in order to gain insight into the inherent mechanisms of the model and analyse its implications. The model is a straightforward extension of the model of Cowan et al. (2007). An industry is given in which companies are characterised by their knowledge bases. Depending on the similarity/dissimilarity of these bases it can be profitable or unprofitable to form alliances in order to carry out joint R&D activities. Hence networks are formed in the industry. However, in contrast to Cowan et al. (2007) in our model knowledge is modelled in two dimensions, namely breadth and depth, following Prencipe (2000) and Ozman (2006). companies can be active in different knowledge (technological fields) and can have different levels of expertise in each field. In the main part of this paper we analyse the evolution of the resulting networks, thus focusing on dynamic aspects of the three variables under consideration (i.e. innovation, locality and heterogeneity).

Our results show that, as implied by the studies mentioned earlier, the closed network of inventors loses its heterogeneity: companies become more homogeneous over time. On the other hand, besides decreasing heterogeneity, we can obtain both more and less innovation in a network. Which case occurs depends on the properties of the initial characteristics of the knowledge endowments of the companies and the industry as a whole – an also on pure chance. This conclusion leads us to build a typology of cluster evolution which differentiates between evolutionary processes among the dimensions of change in heterogeneity and change innovativeness. Our model gives examples of two types from this typology, but contains the possibility of the other two types.

On the other hand we find path dependency and an important role for external shocks in maintaining heterogeneity in the model networks. Learning seems not to have altered our results, although it strengthens the dynamics present in the model with more intense changes in our output variables. Finally, we offer a few words on cluster emergence in which we emphasise the role of diversity in the establishment of new knowledge networks 'from zero'.

The paper is structured as follows. In Section 2 we present the model with all its important features. In Section 3 we outline the methods of our simulations. Section 4 presents the analysis of the simulation results, Section 5 summarises our results on heterogeneity and innovation and gives a typology of cluster evolution. Section 6 concludes our paper.

2. The Model

The model industry is populated by N companies. The number of companies is constant over time so we disregard entry and exit in our approach. Following Prencipe (2000) we consider companies' knowledge along two dimensions, namely the breadth and depth of knowledge. By breath we mean how many technological areas a companies' knowledge base covers and depth refers to the expertise a company has in these fields. Technological or knowledge fields can be bounded according to the technological features of the components of the final product. For example an aircraft producer needs to have some knowlegde in the field of engines, autopilot systems, hydraulics, air conditioning, etc. In these fields companies can have different levels of expertise, which largely corresponds to the degree to which development processes of the different components are carried out in-house (Prencipe, 2000).

According to these, each company is characterized by a knowledge-portfolio which covers both the breadth and depth of companies' knowledge bases. This portfolio of company n is represented by the vector $\mathbf{k}^n = (k_1^n, k_2^n, \ldots, k_w^n)$, where w is the number of possible technological fields and k_i^n represents company n's knowledge level in technological (knowledge-) field i. Higher the value of k_i^n, deeper the knowledge company n has in field i. Of course, the elements of a company's knowledge vector can be zeros, so we only assume that $k_i^n \geq 0$ for all $n \in (1, 2, \ldots, N)$ and $i \in (1, 2, \ldots, w)$.[1] Consequently, the more $k_i^n > 0$ company n has, the broader is company n's knowledge portfolio, i.e. it has competence in more technological fields.

At the outset, companies' knowledge portfolios are generated randomly, in two steps. First, every $i \in (1, 2, \ldots, w)$ technological field becomes part of company n's portfolio with probability p_e, for all i and for all n. The value of p_e is a parameter of the model.[2] Second, if technological field i is part of company n's portfolio, then a given k_i^n value is assigned to the company, choosen randomly from the set $(1, 2, \ldots, k_i^{max})$. k_i^{max}

[1] $k_i^n = 0$ menas that company n has no expertise in field i.

[2] This parameter can be interpreted as the density of technological fields companies have competence in. This is because on average companies will have $p_e w$ technological fields in their portfolio.

stands for the technological frontier in technological field i. For an easier interpretation, we set $k_i^{max} = k^{max}$ for all i. Of course the technological frontier will evolve over time with companies' innovative processes.

As a consequence of our interpretation of companies' knowledge bases we can place the companies in a w-dimensional knowledge-space. A pair of companies is then characterized by their distance in knowledge space. Handling knowledge along two different dimensions (i.e breadth and depth) raises the issue of measuring similarity-dissimilarity of companies. This can be done by measuring the angle of two companies in knowledge-space as done by Cowan et al (2006). However, this approach treats two companies as completely similar if they share the same knowledge fields but one of them is far ahead of the other according to depth of knowledge. On the other hand, in our context it is also relevant how distant two companies are in the depth of knowledge. Therefore we measure the similarity/dissimilarity of the knowledge bases of two companies by the euclidean distance of points $\mathbf{k}^n$ and $\mathbf{k}^m$ in the knowledge-space. This distance can be regarded as the cognitive distance of two companies as it refers to their similarity/dissimilarity in competences and knowledge (see e.g. Nooteboom, 2004). Thus, the distance of company n and company m according to their knowledge bases is simply defined in our model as:

$$d_{n,m} = \sqrt{(k_1^n - k_1^m)^2 + (k_2^n - k_2^m)^2 + \ldots + (k_w^n - k_w^m)^2} = \sqrt{\sum_{i=1}^{w} (k_i^n - k_i^m)^2}$$

Companies can effectively communicate with each other if they share at least some technological fields in which they operate. However, if companies are competent in exactly the same fields, they can still learn from each other if one company knows more than the other, although which company learns and which recieves knowledge is predetermined in this case. On the other hand, for the effective communication it is required that companies be close in the depth of their knowledge as well, because otherwise one of them would be so advanced relative to the other, even if in the same field, that their communication would break down.

2.1. Innovation

In our model, innovation is modelled as a random process. Innovation occours with a given probability specified above. On the other hand, companies can innovate alone and in alliances with each other. How the probability of innovation is determined is different in the two cases.

2.1.1. Separate innovation

When companies innovate alone, all they can do is to increase their knowledge level in one of their existing knowledge fields. In this case, we assume that companies innovate with probability p_0, which is a parameter of the model. If innovation occours, one field is selected randomly from the company's portfolio, and the knowledge level in this field is updated according to:

$$k_i^n(t+1) = k_i^n(t) + 1$$

where t is the time index. This forumlation represents that through innovation companies move upwards on the knowledge ladder, deepening their knowledge in a given field. Assuming that the value of knowledge to the company is represented by the knowledge level, we can easily show that the expected value of separate innovation equals the probability of this kind of innovation, i.e. p_0.[3]

2.1.2. Innovation in alliances

The other possibility for companies to innovate is to look for partners in the industry. However, if two companies form an alliance, not only innovation occurs, but they can learn from each other also. We disregard this possibility for the time being, but incorporate it into the analysis in section 4.3.

Innovation between two companies takes the following form in our model. If company n and m meet, they innovate with probability $p_{n,m}$, depending on their distance in knowledge-space: $p_{n,m} = f(d_{n,m})$. According to those mentioned in the Introduction, there is an optimal distance between companies, denoted by δ, where the probability of success is the highest. Moving farther from this distance in either direction, the probability of success falls. Thus $f(d)$ is single-peaked at δ and symmetric around δ, being monotonically increasing if $d < \delta$ and monotonically decreasing if $d > \delta$. This solution is borrowed from Cowan et al. (2007).

It seems straightforward to assume that allied companies will innovate in those areas where they both have competence. This can be acknowledged

[3] See derivations in the Appendix.

realtively easy by intuition. When R&D alliances form, the partners set the areas in which they will work together. However, a company is not interested in choosing areas where the company itself or the partner is not competent as it would radically lower the chance of successful innovation. As research fields are narrowed in this way, it becomes highly unlikely that the alliance will innovate at a field where either of the partners have no competence. However, we will relax this assumption later by allowing alliances to innovate in those fields where either of the allying companies have competence.

When deciding whether to form an alliance with another company or not, companies must evaluate the expected value of joint innovation. It is showed in the Appendix that the expected value of joint innovation is $p_{n,m}$ in the case of company n and m, i.e. it equals the probability of joint innovation.

2.2. Network formation

Given these results, we can see, which companies will form an alliance. Consider, that the maintenance cost of a partnership is c in each period.[4] In this case, an alliance between company n and company m will form, if the expected value of the joint innovation exceeds the costs of partnership formation:

$$p_{n,m} > c$$

Therefore, formation of a link is simply the function of the distance of the two companies considered.[5] Holding c constant, the number of links a company has is only dependent on the average distance between it and the other companies. This distance, in turn, depends on the parameters of the outeset of the model, namely k_{max}, w and p_e. As $d_{n,m} = d_{m,n}$ by definition, $p_{n,m} = p_{m,n}$. This means that if a partnership with company m is profitable for company n, it is also profitable for company m. So links will be stable in the sense that all alliances that form is beneficial for both partners, and so they are intersted in keeping it alive at least until the next period when knowledge bases change and companies' distances change as well.

[4] Of course, it is a simplification that the cost of maintaining a relationship is independent of the number of these relationship: these cost may increase as the number of links grow which is an interesting extension of the present model.

[5] Recall that $p_{n,m} = f(d_{n,m})$!

The probability of successful joint innovation must reflect our arguments about the optimal cognitive distance presented in the Introduction, i.e. there must be a cognitive distance at which $p_{n,m}$ is the highest for all n, m pairs. This relationship between cognitive distance and innovative performance can be formulated many ways: we follow the simplest rule for $f(d_{n,m})$, which is also used by Cowan et al. (2007):

$$p_{n,m} = 2p_0 \left(1 - \frac{|d_{n,m} - \delta|}{2\rho}\right)$$

.

This formula gives an inverted V form for the above relationship with δ being the optimal cognitive distance for innovation (with the highest $p_{n,m}$) and ρ measures the base width of the inverted V. That is, the larger ρ is, the more companies become suitable partners according to their cognitive distance.

Our model gets a dynamic character when we iterate the alliance-formation and innovative process. Given a network established according to the previous rules, some companies innovate and some not, therefore knowledge bases change which alters the cognitive distance between companies. This causes some links to be dissolved and some to be maintained, thus our network evolves as time passes by. We are primarly interested in this evolutionary process as described in what follows.

2.3. Output measures

As discussed in the Introduction we are interested in the co-evolution of innovation, locality and knowledge heterogeneity in the networks which we build according to the rules detailed above. In order to capture these characteristics we use three output measures along which our simulation results can be interpreted. These are (i) heterogeneity; (ii) innovativeness; (iii) clustering.

- **Heterogeneity.** As it was pointed out in the Introduction, heterogeneity seems to be an important factor in defining innovativeness. The heterogeneity of a population can be measured in different ways. In our context heterogeneity means diversity in the knowledge bases of the companies. As the similarity of companies' knowledge bases with regard to their joint innovation potential is linked to their distance in knowledge space (see the model description), it is straightforward to use a similar measure in order to refer to their diversity. For this reason we calculate the average distance of companies. However this in itself is not a useful measure of heterogeneity because if the knowledge space widens in response to innovation, the possible average distance increases also. To rule out this bias we relate average distance to the diameter of the knowledge space and so we obtain a measure of heterogeneity which is between zero and one (with value zero in case of total homogeneity and value one when companies are as different as possible).[6]

- **Innovativeness.** As our main focus is on innovation activity, we measure how innovative companies are in the industry. We do this by simply counting the number of innovations appearing in each period (both joint and own innovations).

- **Clustering.** It is frequently agrued in the literature that dense local interactions among economic agents contribute to innovativeness. Translating this to the language of networks this means that the presence of local clicks is favorable for innovation. To measure this kind of cliquishness we use the common measure of clustering coefficient as an output variable. The clustering coefficient measures how much a network is clustered, i.e. how much one's friends are friends of each other (Cowan, 2006). This measure is based on counting links in a network, but it reveals some additional information about the structure of the network in contrast to simply counting the links between agents (i.e. using degree as an output measure). However, clustering itself can be misleading in the case of evolving networks therefore we introduce a more convenient measure of localty in what follows.[7]

3. The simulation setting

The model described so far can not be treated analytically. This is not mainly due to its complexity but to the heterogeneity of agents, i.e. that we can not use a representative agent approach. For this reason it is a common method to use simulations in network theory. However, different kinds of simulation can be executed depending on the purpose of the researcher. Here we use simulation techniques in order to substitute for the analytical

[6] See the Appnedix for further discussion.

[7] See section 4.1.2 an the Appendix for further discussion.

analysis of our model. A first possibility in this line is to consequently take all parameter combinations and run the simulation for each combination. Of course the parameter-ranges must be bounded in this case and a minimal parameter-step must be determined, and so we can not examine literally all parameter combinations, although these restrictions can be meaningfully managed by the researcher. However, in the case of too many model parameters this method can become cumbersome. The other possibility is to run Monte Carlo simulation, i.e. to iterate the simulation with randomly generated parameters. Then it is easy to use statistical methods to analyse the correlations between parameters and output measures. This method is followed in this paper.

As a summary of our model, in Table 1 we present the input parameters and output measures being used.

Input parameters	
N	The number of companies
k^{max}	The technological frontier at the outset
w	The number of technological fields
p_e	The density of technological fields in companies' portfolios
p_0	The probability of separate innovation
δ	Optimal cognitive distance
ρ	The measure of the amount of possible partners
c	The cost of partnership
Output variables	
Heterogeneity	Average distance
Innovation	Number of innovation
Cliquishness	Clustering coefficient

Table 1: Input parameters and output measures.

With regard to the design of our simulation experiments, we proceeded as follows. For the model described earlier, 1.000 independent runs were executed. In each run the input parameters were generated randomly (Table 2 summarises the intervals in which the parameters were allowed to vary). Each run was executed for 300 consecutive periods, i.e. the link-formation process and the consequent update of knowledge bases were iterated 300 times during every run. In each period the three output statistics were calculated and recorded, and so we have a record of the dynamic evolution of networks; moreover we have 1.000 such records. This method of simulations yields the opportunity to analyse the effect of different input parameters on the evolution of networks. We turn to the analysis in the next section.

N	$[20, 100]$	p_0	$(0, 1)$
k^{max}	$[1, 20]$	δ	$(0, 20)$
w	$[2, 20]$	ρ	$(0, 20)$
p_e	$(0, 1)$	c	$c = p_0$

Table 2: Intervals of input parameters during the simulations.

4. Discussion of simulation results

4.1. Descriptive statistics and dynamic analysis of output measures

As a first glance at the simulation results, it is interesting to see what the overall dynamics of networks look like. For this reason we posed the question as to whether there are networks at all and, if there are, do they change over time, or remain stable. We sorted the specific runs according to the change in link number and the initial characetristics of the network. That is, we checked, if during a simulation run (i) the network was empty at the beginning or not, (ii) the network changed during the 300 periods or not.[8] Table 3 presents these descriptive results from the simulations.

	Initially empty	Initially not empty
Evolving	75	869
Stable	56	0

Table 3: Basic characetristics of the simulated dynamic networks.

As it is obvious from the table, vast majority of the runs produced evolving networks, which were not empty at the beginning. Only 5,6% of the networks were stable (i.e. no links dissolved or formed during the runs), and these networks were in every case empty networks. This points to the fact that stability of really existing networks (i.e. there exists links among the nodes) is a marginal characteristic of our model which is detected only in those cases when input parameters do not allow for the establishment of linkages at the beginning of network evolution.[9] For this reason we can leave aside the analysis of this kind of network. More interesting is the case of initially empty networks which evolved over time. In these cases input parameters are not able to generate link formation among

[8] A network being empty means that there are no links among the nodes, while change in the network means that at least some links dissolve over time and some new links form.

[9] Whether links form at the beginning or not, depends on the values of input parameters. For example high average cognitive distance among companies associated with a low value of optimal cognitive distance may lead to empty networks as it is not optimal for companies to form alliances.

companies. However, totally probabilistic innovative activity of separated companies lead to an appropriate change in knowledge bases, which, in turn, generates alliances. This early finding is interesting, as some answers to the question of emergence of innovative clusters may lie behind this result. In section 4.4 we briefly return to this issue. In what follows, we analyse the charactersitics of the majority of our experiments, i.e those 869 networks which were not empty at the beginning and evolved over time.

4.1.1. Innovation

First, we look how innovation evolves over time in our experiments.[10] We took those 869 runs where an (initially not empty) network evolved over time and calculated the *average change* in innovation for all runs. That is, the changes in innovation between all periods were recorded and the average of these changes was calculated. Thus we got a value for all 869 cases – which show how, on average, innovation evolved through time. The average of these values is -0,122 which suggests a decreasing trend in innovation. However, this mean is accompanied by a standard deviation of 0,171 which results in a 140% relative standard deviation in absolute value. Standard hypothesis testing shows that in sipte of this high relative deviance the mean is significantly different from zero. It is interesting to look at the histogram of the values of average change on Figure 1.

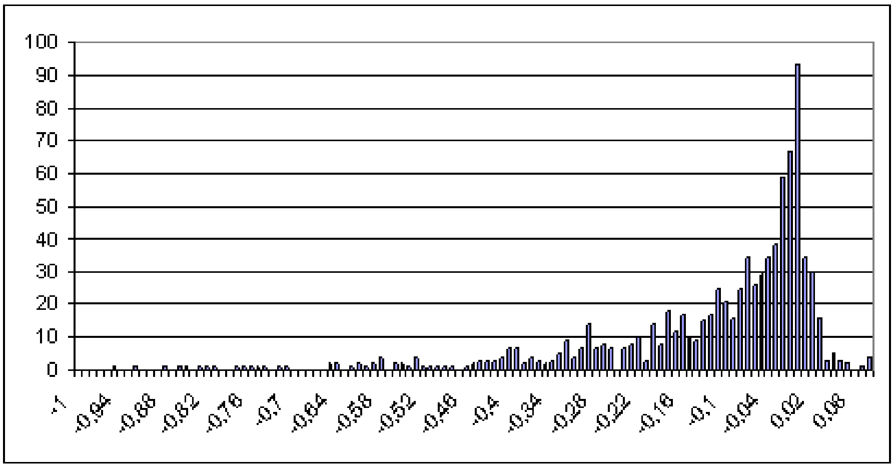

Figure 1: The distribution of average change in innovation.

Average change in innovation follows a rather skewed distribution. The most likely event is to have an average decrease between -0,01 and 0. However, 12% of the networks show increasing innovativeness – a

[10]By innovation we mean the number of innovations observed in the industry, as described in the previous section.

considerable amount, although not the majority. The same results hold, if we consider relative change in innovation, i.e. if we relate innovation activity in the last period to that of the first. On average the networks show 10% weaker innovative activity in the last period than in the first,[11] showing that innovation decreases on average. The relative standard deviation is only 48% in this respect which points to the fact that on average we can seriously speak about decreasing innovation.

To sum up, our analysis of the dynamics of innovation reveals interesting results. First, it seems that innovation decreases in the majority of our experiments, although this decrease is only a slight one. On the other hand, a remarkable portion of cases show increasing innovativeness through time.

The effect of model parameters on innovation

It is interesting to see if the parameters of the model are able to explain if innovation increases or decreases through time. In order to answer this question we carried out a structured regression analysis. First a binary logistic regression model was built in order to evaluate the effect of different model parameters on the probability of observing increasing innovation. Then, a standard linear regression model was built up for the two subgroups (those observations in which innovation increases or decreases over time) in order to analyse the effects of model parameters on the extent of increase or decrease in innovativeness. The results of the regression analyses are presented in Table 4.

The results are interesting. First, only p_e, δ and ρ have significant effect on the probability of increasing innovation. However, the explanatory power of this model is very low, which is also reflected by its capacity of prediction: there is a huge bias towards decreasing innovation. Whilst the model correctly forecasts decreasing innovation, only 3% of the cases when innovation increased were classified properly. These values warn us to be cautious with the results. It seems that model parameters have no clear effects on the probability of getting increasing innovation over time.

Given that innovation increases or decreases over time (the third and fourth column of Table 4, respectively) the effects of different parameters are as follows. The effect of N is of opposite direction depending on the evolution of innovation. If we consider networks with increasing innovation

[11]This means that in the last period the number of innovations were 90% compared to that of the first period on average.

Parameter	Effect on the probability of increasing innovation	Effect on average change in innovation if innovation increases over time	Effect on average change in innovation if innovation decreases over time
N	0,002	0,256**	-0,428**
w	-0,008	0,104	-0,051
k_{max}	-0,028	-0,440**	0,259**
p_e	-1,712**	0,267**	-0,250**
p_0	-6,12	0,118	-0,038
δ	0,067**	0,165	-0,086**
ρ	-0,099**	-0,088	-0,145**
R^2	0,127	0,376	0,335

Table 4: Regression results for innovation. In the case of linear regressions (columns 3 and 4) standardized coefficients are presented. ** means significance at the 0.01 level, * means significance at the 0.05 level.

(column 3) the effect of the number of companies is positive, whilst in the case of decreasing innovation (column 4) its effect is negative. This result is intertesting if we compare it to standard interpretations in cluster literature. There we can find the argument that more companies lead to more intense competition which in turn results in more innovation (see Porter, 1990 for instance). This argument is detected in our model in the types of networks where innovation increases over time. However, the same is not true when innovation follows a decreasing trend. In this case more companies lead to a more intense decrease in innovation over time. On the other hand, our results show that increasing possibility of alliance formation (increasing number of companies) do not neccesarily lead to more innovation in this setting. Innovativeness is also influenced by the knowledge setting of the industry as a whole.

The effects of k_{max} and p_e show the role of this knowledge setting. Given w, k_{max} gives the initial size of the knowledge space whilst p_e affects the overlap among companies' initial knowledge portfolios. The two effects are of opposite direction: increasing k_{max} increases average distance among companies (given all other parameters), whilst increasing p_e decreases average distance among companies (given all other parameters). Regarding these considerations the results in Table 4 show that a higher average distance at the outset (i.e. high k_{max} together with a low p_e) leads to a low average change in innovation in absolute value. In other words, the effect of initial average distance (knowlegde heterogeneity) is different whether innovation increases or decreases over time. If innovation increases over time, higher average distance lowers the pace of increase. If innovation decreases over time, higher average distance lowers the pace of decrease. This result points to our assumption of optimal cognitive distance. It seems that industries being close to this optimal distance on average at the beginning show little change in innovation in either direction. On the contrary, if average distance is far from the optimal, the change in innovation is higher, but its direction is not determined by the value of average distance.

This finding is important from the perspective of culster evolution: essentially the same initial conditions can lead to different outcomes regarding long term evolution. Whether innovation increases or decreases over time seems to be determined by chance in our model (see our previous analysis of the logistic model of innovation). Given this direction a small average distance among companies at the outset can lead to maintained innovativeness or declining innovation.

The effects of δ and ρ are also interesting. They seem to have some effect on the overall direction of innovation, while no effect if innovation increases but negative effects if it decreases. This asymmetry may come from the fact that the sample of increasing innovation is not large enough to generate significant results. If innovation decreases, this negative effect of δ and ρ is quite easy to interpret. Higher values of δ require larger distances in knowledge space for innovation to be maintained. If average distance decrease on average (as it does – see section 4.1.3.) this means that companies are farther away from optimal cognitive distance thus innovation decreases. A smaller value of ρ generates a smaller interval around δ suitable for partnership, thus it leads to reduced innovation.

As a final remark on innovation, although on average a slight decrease in innovation is present in the model, the increase of innovation is also possible. Which case occours seems to be explained by model parameters only to a limited extent. Regarding the effect of model parameters on the change in innovation over time our results revealed interesting points, mainly that the number of companies have negative effects on innovation while the average distance among companies at the outset have different effects depending on increasing or decreasing innovativeness in the industry. Nevertheless, caution is required when interpreting these results, because the explanatory power of the regression models are very low. This points to the fact that chance is an important driver of innovation in our model. However, we must add, that by chance we mean the randomn modelling methods used in our setting. It refers to all those factors which affect innovation but not modelled here explicitely. That is, chance does not neccesarily mean that innovation comes from the 'air'.

4.1.2. Clustering

In what follows, we apply exactly the same methodology used for innovation to analyse the evolution of clustering (network structure) over time. In those 869 runs where an initially not empty network evolved over time we calculated the *average change* in the clustering coefficient. This shows a slight decline with a mean of -0,00012, and standard deviation 0,000438 which results in 350% relative standard deviation.[12] If we take a look at the historgram in Figure 2 it becomes obvious that we can not convincingly state that clustering decreases over time. The majority of cases (569 out of 869) is found between -0,0001 and 0,0001. However, more than half of them lie in the negative range. Considering relative change in clustering we found that in the last period the clustering coefficient is only 95% that of the first period with a standard deviation of 37 percentage points, which also show a very slight and not too significant decrease.

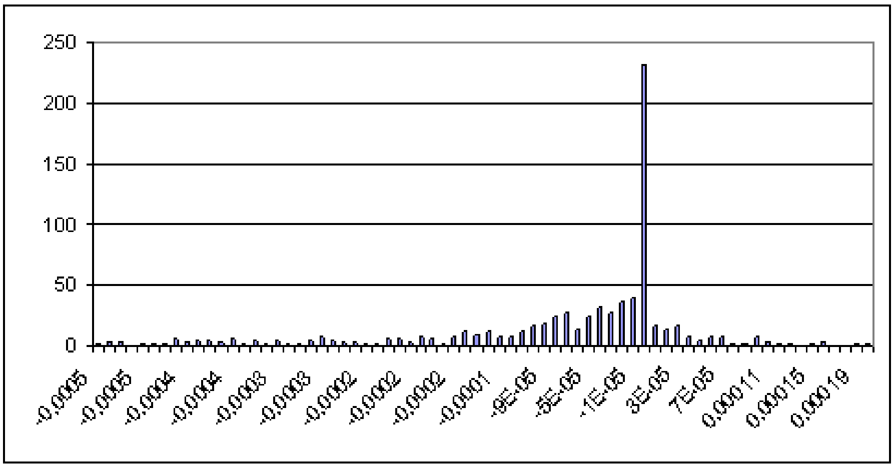

Figure 2: The distribution of average change in clustering.

The effect of model parameters on clustering

Following the same methodology as in the previous subsection, we tried to trace the role of different model parameters in the evolution of the clustering coefficient. The results are presented in Table 5. The only

[12]Note, that a small absolute value in average change is not surprising in itself, as the clustering coefficient is between 0 and 1.

difference is that networks are again differentiated along the evolution of innovation and not clustering. This is due to comparability between results.

Parameter	Effect on the probability of increasing innovation	Effect on clustering if innovation increases over time	Effect on clustering if innovation decreases over time
N	0,006	-0,034	0,215**
w	-0,001	-0,029	-0,039
k_{max}	-0,019	-0,435**	-0,083**
p_e	-0,234	0,286**	0,065*
p_0	2,107**	0,182*	0,386
δ	0,034	0,097	0,060
ρ	-0,091**	0,008	0,046
R^2	0,130	0,276	0,212

Table 5: Regression results for clustering. In the case of linear regressions (columns 3 and 4) standardized coefficients are presented. ** means significance at the 0.01 level, * means significance at the 0.05 level.

With regard to increasing or decreasing clustering only two parameters are found to significantly affect the direction of the change. p_0 stands for the probability of separate innovation and indirectly affects joint innovation also. Its positive effect shows that a higher probability of innovation in the industry, leaving all other parameters constant, increases the probability for clustering to increase over time. This is an interesting finding as here we have some solid base for reflection: the probability of innovation is given at the outset, and so, if we accept the results from the regression, increasing clustering can be detected as being caused by higher innovation probability, or, as it is frequently interpreted, greater efficiency in R&D activity. This is an important finding as it points to the fact that it is possible that higher clustering is caused by more dense innovative activity and not vica versa as it is frequently claimed.

The other parameter to affect clustering is ρ, that is, the suitable interval for joint innovative activity around optimal cognitive distance. Its negative effect is somewhat interesting, but there is a clear way to interpret it. It is interesting, because one would expect a higher ρ to increase clustering as higher ρ makes more links profitable and thus more link forms and clsutering increases. However, this reasoning leaves aside one important aspect of clustering, namely that it measures local linkages in contrast to overall linkages in the network. Clustering increases if local neighbourhoods are getting denser. Taking a reverse perspective, if ρ decreases, it is less suitable for companies to form pairs on average, but those companies who are close to each other in the knowledge space (who are relatively similar), it is still profitable to form alliances. Thus a lower ρ results in less links but these links will be concentrated in the knowledge space among companies who are close to each other. Local linkages thus gain more weight compared to more 'global' links.

Looking at the effect of model parameters on clustering (differentiating again between increasing and decreasing innovation) we see basically similar results, alhough there are some remarkable differences. First, industry size (N) only affects clustering in the case of decreasing innovation. If a network shows declining innovation, more companies lead to higher change in clustering: clustering increases more or decreases less intensely. This is not true in the case of increasing innovation. If innovation increases, industry size has no effect on clustering. The effect of the initial knowledge space (k_{max} and p_e) on clustering is more interesting. It seems that there is no ambiguity in this respect in contrast to the case of innovation. Independent of the overall direction of innovation, the more dense is the initial knowledge space (higher p_e and lower k_{max}), the higher is the average change in clustering. This points to the conclusion that clustering does not follow a linear trend in our model. Given that companies are cognitively close to each other at the beginning, a more clustered network is likely to emerge initially. A decreasing trend in clustering results in a slight decrease when clustering is high at the beginning whilst a more robust decline when clustering is lower: the trend of clustering follows a logistic pattern. There is no difference between cases when innovation increases or decreases over time.

There is another parameter, p_0, which affects the average change in clustering. The interpretation is quite clear. A higher probability of joint innovation makes it more profitable to form alliances (given the costs of alliances, c). This results in more links, either local or global, and so clustering increases over time.

A few notes on normalised clustering

It is impotrant to look at the evolution of the clustering coefficient itself, however, it is also interesting to pose the question whether a higher clustering coefficient really means that network structure becomes more local.[13] It would be neccesary to normalise somehow the clustering coefficient to capture this relative aspect of local link formation. As the clustering coefficient is basically a density measure (it counts the number of local links relative to possible local links) it is straightforward to use global density as a benchmark. If we divide the clustering coefficient by network density we get a measure of the locality of the networks: if it is greater than one, local density (clustering coefficient) is higher than overall density meaning that links tend to be local in the sense of the network space.[14]

If we use this method to refer to the localised character of the networks under consideration, we get an interesting result. In those 869 cases which we have considered so far, this ratio is significantly greater than one. This means that in our experiments networks were more dense locally than gobally, i.e. we have an overall clustered network structure. This is not that interesting in itself, as all networks in our sample have the same property on average. However, if we look at the evolution of this ratio, we find that it is increasing over time in 85% of the cases considered. Thus we can draw the conclusion that although the clustering coefficient reports a slightly decreasing trend in the coefficient, the structure of our networks become more local. This result is explained by the fact that on average the number of links decrease during our experiments. This leads to a decrease in clustering as this measure is highly correlated with link number. On the other hand, local links (links in neighborhoods) decrease at a slower pace than more distant links. This leads to the observed phenomenon of increasing locality in the networks.

Our results point to the fact that clustering coefficient is not always a good measure of the local structure of a network. In our experiments locality increases in spite of a decreasing clustering coefficient. However, it must be noted that the standard deviation of average decrease in locality is very high again, thus the detected increase in this respect is, although statistically significant, quite small.

4.1.3. Heterogeneity

Finally, we take a look at how heterogeneity evolved over time in our experimental networks. Essentially the same methods were used as before in the case of innovation and clustering. On average the change in

[13]The reason is that higher clustering coefficient can be a result of more links in the network in general which obviously leads to more links in local neighborhoods.

[14]See the derivation in the Appendix.

normalised average distance[15] was -0,00066, and so a decrease is again detected in this respect. The small absolute value is due to the fact that normalised average distance must lie between 0 and 1.[16] However, now the standard deviation is 0,00041 which results in a relative standard deviation of 61%. This is much lower than the similar values of innovation and clustering which shows that here we can talk about a remarkable downward trend in normalised distance, i.e. heterogeneity. Again, hypothese testing shows that this average change is significantly different from zero. With respect to relative change we found that in the last period the value of normalised distance was only 53% that of the first period. Using this value as a reference we can say that in the last period companies were slightly more than half as heterogeneous as before: industries in our experiments tend towards homogeneity. The usual histogram is presented in Figure 3. It can be seen that now the distribution is close to normal. There is only 32 cases where heterogeneity increases, being rather an exception than a rule.

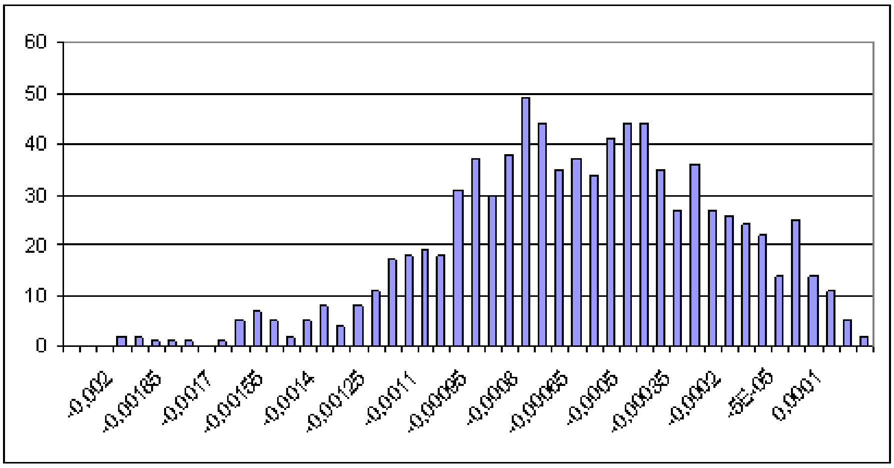

Figure 3: The distribution of average change in heterogeneity.

This shows that heterogeneity clearly falls in our model over time, that is, companies become more similar. This is due to their joint innovative activity which results in more homogeneous knowledge portfolios. This result is in line with those thoughts mentioned in the Introduction. However, there are two important extensions to previous studies based on our analysis. First, in our model the loss of heterogeneity can be consistent with increasing innovativeness as we can not detect co-evolution in heterogeneity and innovation (see section 4.2. for details). Second, there seems to be a reinforcement of the studies on social capital theory which

[15]See the Appendix for details.
[16]And the experiments show that with the exception of really few cases it is lower than 0,5.

argue that external refreshment is required for maintaining heterogeneity in a given group (e.g. Coleman, 1990; Putnam et al., 1993, Granovetter, 1973). Our model gives an example of a closed group of innovators who are not exposed to this kind of external shocks. It is clearly shown that this setting leads to a loss of heterogeneity. However, in our model innovation can increase in spite of decreasing heterogeneity which is one of the most important results here.

The effect of model parameters on heterogeneity

We conducted a similar analysis for model parameters as before. The results are presented in Table 6. Again, the different cases were classified in column 3 and 4 with respect to the evolution of innovation, in order to have a benchmark among the analyses.

Parameter	Effect on the probability of increasing innovation	Effect on heterogeneity if innovation increases over time	Effect on heterogeneity if innovation decreases over time
N	-0,043**	-0,091	-0,224**
w	-0,062	0,071	0,184**
k_{max}	0,024	0,467**	0,499**
p_e	-12,387**	-0,439**	-0,246**
p_0	13,512**	-0,464**	-0,477
δ	0,017	-0,023	-0,034
ρ	-0,062	0,109	0,037
R^2	0,671	0,649	0,599

Table 6: Regression results for heterogeneity. In the case of linear regressions (columns 3 and 4) standardized coefficients are presented. ** means significance at the 0.01 level, * means significance at the 0.05 level.

With regards to the question of increasing or decreasing heterogeneity, three model parameters are found to have significant effect. All three decrease the probability of increasing heterogeneity, i.e. all parameters act towards homogeneity. First, a more dense industry with respect to the number of companies lead to decreasing heterogeneity with a greater probability. The reason for this can be found in our definition of heterogeneity. Given the size of knowledge space (k_{max} and w) more companies lead to lower average distances at the outset.[17] Thus, a more dense industry is associated with more homogeneity in the beginning therefore there is less room for a further decrease. Fewer companies,

[17] Recall that companies are randomly distributed at the beginning.

on the other hand, leave more space for further decrease in homogeneity. Second, p_e has a negative effect on the probability if heterogeneity increases over time. This parameter affects the average distance of companies in the initial knowledge space and so its effect can be explained along the same arguments as that of N. Third, the effect of p_0 shows that a higher probability of innovation leads to less heterogenous networks. The reason behind this result is less technical than that of the other two parameters. Increasing p_0 leads to more links (as more links become profitable). Through more links companies interact more and their knowledge base is more likely to converge thus converting them more homogenous. However, we must note that in the present setting there is no learning included in the model. Companies become more similar only through joint innovation.

If we look at the effect of parameters on the value of average change in heterogeneity it is clear that the cases of increasing and decreasing innovation is basically the same. Parameters k_{max}, p_e and p_0 have the same effect regardless of the direction of innovative activity. k_{max} and p_e, as was discussed previously, define the initial level of heterogeneity in the industry. The signs of the respective coefficients show that an initially more heterogeneous industry (higher k_{max} and lower p_e) is likely to lose less from its heterogeneity than an initially more homogeneous one.[18] This result can be interpreted in two different ways. First, it can reveal a logistic shape of the evolution of heterogeneity: slight decrease at the beginning (with higher levels of initial heterogeneity) while larger decrease at the end (from lower values of initial heterogeneity). Second, and more importantly, this points to the fact that, although convergence is not ruled out in this model, there is an inherent inertia in the evolution of heterogeneity. Initially more heterogeneous industries are more likely to maintain their diversity whilst more homogeneous ones tend to homogeneity more rapidly. This reveals the typical path-dependence in cluster evolution discussed widely in the literature. If the industry attains a high level of heterogeneity, it is easier to maintain it as forces towards homogeneity are less strong. How this level of heterogeneity is achieved, however, is not modelled in our context. Initial knowledge portfolios are given at the outset.

The positive effect of w can be interpreted similarly to k_{max} as this parameter also increases heterogeneity in the initial knowledge space. However, its effect is clear only in the case of increasing innovation. The

[18] Here we took the overall declining trend in heterogeneity as given. In this case the interpretation of a positive sign in k_{max} for example is that the decrease is smaller if k_{max} gets higher.

effect of N is negative and only detectable in the case of decreasing innovation. More companies lead to larger decline in heterogeneity given that innovation decreases over time.

4.1.4. Conlcuding remarks on model parameters

So far we have analysed the effects of model parameters on the evolution of different output measures in detail. Although their effect provides precious insights into the inherent processes of our model, the relevance of these findings should not be overemphasised. If we look at the explanatory power of these regression models, it is clear that, although model coefficients are found significant, the evolution of different output measures are not always explained well by the values of model parameters. This is true for both the logistic and the linear regression models presented above. The only exception is heterogeneity, where R^2s are at an acceptable level. This result, however can not be seen as a drawback of our analysis. The small explanatory powers are consistent with the overall evolution of the different output variables. Innovation and clustering decrease only very slightly on average. A regression model which tries to trace out the role of different effects in this trend can not be succesful in such a situation. In the case of heterogeneity the picture is more favorable but again, this is in line with the finding that the overall decline in heterogeneity is much stronger.

These arguments let us conclude that in the case of innovation and clustering we can not convincingly state that these measures decrease over time. It is more convenient to say that these values are quite stable with a greater probability of decrease than of increase. On the other hand, heterogeneity is clearly decreasing, so in our model networks tend towards homogeneity with regard the knowledge portfolios of the companies.

4.2. The co-evolution of network characteristics

In the analysis so far, we have discussed the dynamic behavior of the three output measures and tried to trace out the effect of different model parameters on this evolution. In what follows, we consider the co-evolution of different output measures to find out if there is a correlation among different characteristics of our model industry. We do this in two aspects. First we look only at the binary variables of increasing or decreasing innovation/clustering/heterogeneity. This analysis can detect if there is a basic correlation among output variables. For example, does innovation increase if heterogeneity increases? Second, we consider the correlation among average changes in the three output measures. Do a higher average change in innovation correspond to a higher average change in heterogeneity?

In Table 7 the correlation coefficients are presented with respect to our first aspect, i.e. the correlation among increasing and decreasing output measures. As a point of reference we included our locality measure in the analysis also.

	Innovation	Clustering	Heterogeneity	Locality
Innovation	1,000	0,170**	0,116**	-0,267**
Clustering	0,170	1,000	-0,046	0,165**
Heterogeneity	0,116**	-0,046	1,000	-0,177**
Locality	-0,267**	0,165**	-0,177**	1,000

Table 7: Correlation coefficients of binaries showing the overall direction of change in output parameters.

The results are not at all convincing. Although significant relationships are detected, the value of the correlation coefficient is always under 0,3. A deeper analysis (directly counting co-occurences of 0s and 1s in this database) suggest the same conclusion. The significance of the results stem from the fact that vast majority of the experiments show similar patterns in the output measures (decreasing innovation, clustering and heterogeneity). This, in turn, causes many similar counts on one side but there is a missing relation on the other. For instance, taking innovation and clustering there are only 32 cases when innovation and clustering both increase over time whilst innovation increases 104 times and clustering 126 times. At the same time there are 671 cases when both measures decrease over time. The same is true for all pairs of output variables.

This leaves us with the conclusion that no robust co-evolution can be detected regarding the overall direction of output measures. This is substantially important in the case of innovation and clustering as well as innovation and heterogeneity. We cannot state in either case that the two output measures evolve in correspondance. While clustering decreases innovation can increase and decrease as well. And while heterogeneity basically decreases, innovation can increase and decrease also. Although these results seem unsatisfying at the first sight, they reveal important characteristics of our model. In our model the different output measures are not deterministically correlated. Although heterogeneity really seems to decline over time resulting in more homogeneous industries, this does not lead to less innovation per se: innovation can increase in those industries where companies become homogeneous over time.

Taking a final look on locality, essentially the same picture emerges. Irrespective of the overall direction in locality, innovation can increase or decrease over time. However, looking at the corelation between locality and clustering our argument that the two measures are basically different can be proved again. Higher clustering is not neccesarily followed by higher locality and vice versa.

The picture revealed previously changes somewhat if we do not correlate binaries but average changes in the different output measures. The results are presented in Table 8.

	Innovation	Clustering	Heterogeneity	Locality
Innovation	1,000	-0,051	0,428**	-0,014
Clustering	0,051	1,000	-0,325**	0,399**
Heterogeneity	0,428**	-0,325**	1,000	-0,054
Locality	-0,014	0,399**	-0,054	1,000

Table 8: Correlation coefficients between average changes in output variables. All cases included.

What is clear, that significant relationships become rather more robust in this case, although they are all below 0,5 in this case also. Altogether three significant relations can be detected. First, innovation and heterogeneity has a positive correlation. This is not surprising as not innovation and heterogeneity per se are correlated here, but their change over time. If we consider the decreasing trend in the two measures this correlation means that a smaller decrease in heterogeneity corresponds to a smaller decrease in innovation. That is, in those industries where heterogeneity declines at a lower pace, innovation also declines at a lower pace – or it may increase also. In industries where heterogeneity increases, innovation may not decrease but increase - however this last statement is not proved by our previous analysis of binary variables. Thus in our model, at least at the level of average change in output measures, heterogeneity seems to foster innovation. It is not clear however, which causes which. Larger decreases in innovation can lead to larger decreases in heterogeneity also.

Second, clustering is negatively correlated with heterogeneity. A larger decrease in heterogeneity corresponds to a smaller decrease in clustering. This result is very interesting and again emphasises the basic difference between clustering and locality: in spite of the correlation with clustering, there is no significant correlation between heterogeneity and locality. Our third significant relation is that of clustering and locality: this is positive and shows that a higher change in clustering corrpesponds to a higher change in locality. This result is embarassing as clustering decreases on average whilst locality increases on average. This contradiction can be resolved by the definition of our locality measure. Locality is defined as the ratio of the clustering coefficient and network density. As both clustering and density declines over time (because fewer links are formed in the networks) it is obvious that density decreases at a higher rate than clustering. Given a rate of decrease in density, the smallest is the decline in clustering, the highest is the increase in locality, which is what the results tell us.

We should stress that these results are still not too convincing. They can be interpreted as a slight co-evolution of output variables in the three cases underlined above. Things become more clear, however, if we make our usual distinction between increasing and decreasing innovation, i.e. if we treat separately the cases when innovation increases or decreases over time in our experiments. Table 9 and 10 present the respective data.

	Innovation	Clustering	Heterogeneity	Locality
Innovation	1,000	-0,428**	-0,567**	0,205*
Clustering	0,428**	1,000	-0,565**	0,419**
Heterogeneity	-0,567**	-0,565**	1,000	-0,208*
Locality	-0,205*	0,419**	-0,208*	1,000

Table 9: Correlation coefficients between average changes in output variables. Cases with increasing innovation are included.

What is obvious that the relationships become more robust in all cases. On the other hand, the directions are basically the same as before, with two differences. First, whilst innovation and heterogeneity are positively correlated if innovation decreases over time, the two variables are negatively correlated if innovation increases over time. This last result modifies our earlier finding. There we concluded that a larger decrease in heterogeneity leads to a larger decrease in innovation. This is still true, but if we treat those cases where innovation increases separately, we find the opposite results. Given that innovation increases over time, a higher change in innovation corresponds to a smaller change in heterogeneity. In those industries where innovation increases, this increase is larger if heterogeneity decreases more (given that heterogeneity is typically decreasing over time). This again points to interesting coclusions. Whilst innovation can increase if heterogeneity decrease in our model, a higher pace of innovation correpsonds to a higher decrease in heterogeneity, i.e. more homogeneous companies.

The second difference is that although more significant relationships occur, these are characterised by a small correlation coefficient, therefore we cannot treat them as real correlations. There is one exception: the correlation between innovation and clustering if innovation increases over time. This means that a higher increase in innovation corresponds to a higher level of innovation, and so the basic argument in the literature that dense local relationships are favourable for innovation can be detected here, although this relationship is valid only if innovation increases over time.

	Innovation	Clustering	Heterogeneity	Locality
Innovation	1,000	-0,160**	-0,458**	-0,026
Clustering	-0,160**	1,000	-0,309**	0,423**
Heterogeneity	0,458**	-0,309**	1,000	-0,037
Locality	-0,026	0,423**	-0,037	1,000

Table 10: Correlation coefficients between average changes in output variables. Cases with decreasing innovation are included.

Regardig our results it is still important to note that the correlation values are still under 0,5 and that no causal relationships can be detected. These correlations show only co-evolution and not the direction of effects.

4.3. Dynamics with learning

In the experiments presented above companies did not learn from each other, but innovated in those areas where both partners had competence. However, it is hardly reasonable to assume that companies do not learn from each other when they work together in alliances. In what follows, we introduce learning into or model and present basic results of simulations under the assumption of learning. However, it seems that the introduction of learning does not change our results but makes the observed relationships stroger.

In order to examine the implication of learning in our model, we proceed in two ways. The two approaches are distinctive on how learning occurs: we can assume that successful innovation is required for learning something new, and also that it is possible that it is not required. We refer to the first case as 'learning through innovation' and to the second as 'autonomous learning'. Regarding our model presented previously, learning can be handled in two different ways again. First, it can be seen as 'catching up' with allies, i.e. if there is a difference in knowledge levels at the field selected for innovation, the lagging company can catch up with the leader. In this case, however, the knowledge portfolios of companies do not become wider (no new areas are incorporated), they are fixed to the portfolio determined at the outset. Second, leraning can be seen as integrating new knowledge fields into the knowledge portfolio of a company. This means that joint innovation is not restricted to those fields where both companies have competence, but they can innovate in the union of fields belonging to either companies' knowledge portfolio.[19] This altogether defines four different types of learning, although our results show that there is no basic difference between them, so we do not treat them separately.

If we look at the descriptive statistics on the basic characteristics of the emerging networks, we see a very similar picture as presented in Table 3. The numbers are slightly different but this can not be regarded as systematic: they are due to chance inherent in our model. 87% of the experiments show evolving networks from initially not empty ones, and there is no such case where an initially not empty network remained stable over time. A slight difference is present in the cases where industries start from an empty network. 62% of our experiments remained empty in this respect, compared to the 56% in the no-learning experiments, whilst in 68% of the cases a cluster emerged from an empty network over time, in contrast to the 75% in the no learning cases. This difference is very small in order to conlcude that learning creates a less favourable environment for the emergence of clusters – which would, however, be quite counterintuitive.

Now we turn to the characteristics of the different output measures. The basic trends are similar to those of the no learning case. Innovation decreases in most of the experiments, but the average decrease is somewhat larger. On the other hand, clustering also decreases, but the average decrease is much smaller in absolute value: the slight downward trend in clustering becomes nearly unvisible if learning is incorporated (regardless of the type of learning). Heterogeneity also decreases among companies, but the average decrease is a little larger than in the no learning case. These results can be interpreted easily. Learning leads to homogeneity more

[19] Note, that technically we can not distinguish between innovation and learning in our model. Learning means that a company have new knowledge which is the same as innovation. The difference is that learning means new knowledge only to the company itself. (On the other hand, innovation is not neccesarily new to the industry as a whole: alliances may innovate something which is already known by others.) Innovation is surely new to the given allies but it is not neccesarily new to the industry.

rapdily and so companies leave the optimal interval for joint innovation (they become too homogeneous) more rapidly also. This results in a higher rate of decrease in innovation on average.

Our result on the disappearing decrease in clustering is interesting. This finding is underlined by the behavior of locality. If learning is incorporated into the model, the increase in the locality of the network becomes significantly higher. This is in line with the (nearly stagnating) clustering coefficient, but reveals that the structure of networks become more localised over time. This reveals an important relationship between learning and localisation (clustering). It is argued by others (see Cowan, 2006, for instance) that a more clustered network is favourable for learning through intense local relationships. In our setting, the reasoning goes the other direction: learning leads to more localised structures in networks. The two arguments are, however, not contradictory, but refer to a self-sustaining dynamic process: learning leads to the emergence of localised network structures, which again reinforces learning. That is, we can complement the static view between network structure and learning with a dynamic, positive feedback rule.

Our experiments with learning included in the model shows that its basic properties do not change. The overall decreasing trend in innovation is slightly strengthened which is due to the effect of learning in leading to homogeneity more rapidly. However, the decline in clustering and locality is not present which reveals a positive dynamic contribution of learning to the emergence of local structures. If we take these two findings in contrast, our model clearly shows that in such dynamic networks, either learning or innovation dominate. Learning can lead to more localised structures which reinforce learning, but at the same time the network loses its innovative potential. If learning is not present, innovation is more sustainable and networks show a less localised picture.

4.4. Cluster emergence

We have mentioned previously that a portion of our simulation runs produced clusters emerging from 'nowhere', that is the parameters of the model generated an empty network (no links formed at the beginning), but due to the seperate innovative processes of the companies knowledge portfolios evolved such a way that some links eventually became profitable to form. From that point on these networks remained 'not empty', i.e. at least some links were present in every period. It would be interesting to see if the parameters of the model have an influence on the emergence of networks. In order to analyze this aspect, we took those runs in which the networks were empty at the beginning and differentiated among stable and evolving ones. A binary logistic regression was carried out to see if the different parameters have significant effects on the emergence of clusters. The results of this regression analysis can be seen in Table 11.

Parameter	B
N	0,042
w	-0,176
k_{max}	-0,138
p_e	-0,380
p_0	1,074
δ	0,412**
ρ	0,091

Table 11: Binary logistic regression results for the effect of model parameters on cluster emergence. ** means significance at the 0.01 level, * means significance at the 0.05 level.

The results reveal that there is only one parameter, δ which has significant effect whether a cluster emerges or not from an initially empty network. This parameter defines the 'optimal cognitive distance' for joint innovation, i.e. that level of similarity/dissimilarity among companies (regarding knowledge) which is the most favourable for joint innovation. The positive effect of δ on the emergence of clusters means that if a higher cognitive distance is optimal for joint innovation, it is more likely that a network will be established among industry companies. Higher optimal distance in turn corresponds with higher optimal diversity. Thus, we can conclude that higher heterogeneity is favourable for cluster emergence which is in line with some previous studies (e.g. Karlsson, 2008). An interesting extension of the present study would be to focus on the emergence of clusters based on the knowledge properties of the industry.

5. Innovation and heterogeneity

Many aspects of our model have been analysed, although our basic question about innovation and heterogeneity remains to be summarised. As it was mentioned in the Introduction, different studies tell us different stories about the relationship between these two phenomena. In addition, our analysis draws a rather mixed picture also. In order to clarify our arguments it seems reasonable to introduce a basic typology of cluster evolution. This typology differentiates between clusters with respect to the evolution of the heterogeneity of companies and that of the innovation potential of the cluster: first, if companies become homogeneous or if there is a maintained heterogeneity in the cluster and second, if innovation is sustained or it is decreasing through time. With two classes along both dimensions we can distinguish among four types of cluster-evolution, as presented in Table 12.

	Sustainable Innovation	Decreasing Innovation
Persistent Heterogeneity	**Dynamic cluster** *Knott (2004)*	**Fragmented cluster** *?*
Decreasing Heterogeneity	**Specialised cluster** *This model*	**Declining cluster** *Cowan&Jonard (2007), this model*

Table 12: Typology of cluster evolution regarding heterogeneity and innovation.

- The first category can be labelled as 'dynamic', showing both sustainable innovation and persistent heterogeneity. These kinds of cluster are typically based on urbanisation or Jacobs' externalities (Jacobs, 1969) where innovative potential stems from heterogeneity itself. The model of Knott (2003) fits this type.

- The opposite category is that of decreasing heterogeneity (companies becoming homogenous) and decreasing innovation. This type may be named 'declining' and the model of Cowan and Jonard (2007) fits to this one. Our model presented in this paper also contains the possibility of this type.

- We shall term 'fragmented' those clusters where companies remain heterogeneous but innovation decreases over time. Although we have not found examples for this type in the literature, it is not theoretically impossible.

- The fourth version is the case of decreasing heterogeneity and sustainable innovation, which could be labelled as 'specialised' cluster. Our model also contains this type of cluster.

This typology clearly shows that there is no one simple model of cluster evolution, rather different paths can be detected. Which type are relevant to a special case depends on different factors. Our analysis shows that the initial setting of the knowledge space in the industry can be a relevant factor, however pure chance can also be important. Other studies imply that different types of cluster evolution exist in different phases of the cluster life cycle (e.g. Menzel and Fornahl, 2007).

Our model presents a simple example of the possibility of increasing and decreasing innovative potential, whilst companies in the industry become more heterogenous. This points to the complex mechanisms hiding behind cluster evolution. On the other hand it is straightforward to cite the obvious connection with different types of externalities discussed in the literature on the economics of agglomeration. Localisation economies refer to those effects which makes co-location advantageous for similar companies. In this case similar knowledge bases drive the agglomeration process: companies can easily communicate with each other as they are all specialised in the same technological fields (Johansson and Forslund, 2008; Weterings and Boschma, 2006). On the other hand, urbanisation economies (also referred to as Jacobs externalities) are quite different. In this case not similarity, but heterogeneity is the main force which drives economic agglomeration. Jacobs (1969) argues that large cities are attractive places for innovative activities because of the heterogeneous community which serves as an unfailing source of new ideas and associative innovations. On the basis of these urbanisation economies, companies feel it advantageous to co-locate with other companies not because of the ease of knowledge transfer but because of the highly vibrant atmosphere which provides extraordinary possibilities for innovation.

In this paper we proposed a model which proves that basically similar mechanisms can lead to both types of agglomerations or clusters. Which type of externalities prevail depends on the initial characteristics of the knowledge space and on chance. The model also reveals that path-dependence is an important factor in cluster evolution: either type of evoutionary path emerges at the beginnig, previous developments determine future paths.

On the other hand, our model does not deal with the possibility of increasing or maintained heterogeneity – which is an obvious drawback. Several empirical studies emphasise the presence of this type of clusters (Molina-Morales and Martínez-Fernández, 2004). One defensive point according to this consideration can be that, by including learning into our model, we can distinguish between higher and lower decrease in heterogeneity, admitting that this approach is far from maintained or even increasing heterogeneity.

6. Discussion and conclusions

In this paper we presented a simple model of network evolution based on knowledge sharing and innovation. The analysis of simualtion results revealed interesting insights into the working of the model and the evolution of knowledge networks, as detailed in the sections above. In what follows, we briefly discuss the most important findings and draw conclusions about them.

During our simulations, the evolution of three main characteristics of the newtorks were recorded: innovativeness, clustering (locality) and heterogeneity. First, we found that heterogeneity is decreasing over time in our model, i.e. companies become more homogeneous, which is due to innovation as their knowledge bases converge during innovative activity. Second, this decrease in heterogeneity does not correspond to declining innovativeness as it is frequently argued in the literature. Besides tending towards homogeneity some experiments show increasing innovativeness which points to the conclusion that innovativeness is not linked to the heterogeneity of knowledge bases of companies. In our cluster typology this model presents both the 'specialised' and the declining' cluster': a decrease in heterogeneity can accomodate both increasing and decreasin innovativeness. Third, we find that the clustering coefficient is quite stable over time with a minor decrease on average, although our special measure of locality reveals that networks become more locally structured as time passes, meaning that local links dominate global links.[20] This is, again, an important finding as it shows i) that locality in network structure can not be fully measured by the clustering coefficient and ii) that increasing local density in network structure can be reconciled with increasing and also decreasing innovativeness. Fourth, the mutual decrease in heterogeneity and increase in locality strengthens our view on the standardising forces of strong local linkages. However, as opposed to several studies (e.g. Granovetter, 1973) in our model strong locality and the resulting homogeneity do not neccesarily lead to a loss of innovativeness although in the majority of the experiments this is the case.

A further important finding in our analysis is the presence of path dependency. The initial setting of the knowledge space defines the path of network evolution to a considerable degree. This feature can be found in the case of clustering and heterogeneity also. If we consider

[20] By local here we mean geodetic locality as geogpraphy is not included in the model.

the clustering coefficient, as described in section 4.1.2., an initially more clustered networks remains relatively clustered in spite of the decreasing trend in clustering, while initially less clustered networks show a more intense decrease in clustering. Considering our relative locality measure the same tendencies are present into the other direction (locality generally increases in the experimental netwroks). We also found this feature in the case of heterogeneity: initially more heterogeneous networks, although tending towards homogeneity, lose relatively less of their heterogeneity than initially more homogeneous networks. Thus, the evolution of networks in our model reinforces initial differences in heterogeneity and clustering in spite of the overall decreasing tendency of the two.

Although offers interesting insights into knowledge-based network evolution, our model has its clear limitations. Although modelling knowledge is said to be relatively sophisticated as both the dimensions of breadth and depth are considered, our approach is still far from perfect. First, our knowledge space is orthogonal which results in the fact that all knowledge fields are equally different/similar. Therefore we can not distinguish between closer and farther substitutability between knowledge fields. Second, limiting innovation to one-unit steps upwards on the knowledge ladder is also a major simplification. A straightforward extension would be to allow for innovation to be proportional to the existing knowledge stock of the company. One further limitation is the symmetry of network formation. It is obvious that research alliances form only if the alliance is benefitial for both partners. Our model is consistent with this recognition. However, it is not neccesary that if a link is benefitial for company A it would be benefitial for company B as well – which is the case in our model. Incorporating this consideration into our model would not essentially alter our results, only the average number of links would be lower.

Furthermore, although listed in our typology in section 5, our model does not generate networks with increasing heterogeneity. This is mainly due to its closed structure: there are no outside shocks which could add new knowledge fields to the existing ones. Although innovation can in principle be infinite in the knowledge fields, the number of fields is fixed at the outset for the industry. On the other hand, this shows i) the importance of external effects in maintaining heterogeneity (though not neccesarily innovativeness); ii) that adding such extarnal shocks to the model it could be able to generate all four possible cluster types given in our typology.

References

Acs, Z.J., Audretsch, D.B., Feldman, M.P. (1992): The Real effects of Academic Research: Comment. *American Economic Review*, 82. 363-367.

Aghion, P., Howitt, P. (1992): A Model of Growth Through Creative Destruction. *Econometrica*, 60 (2), pp. 323-352.

Anselin, L., Varga, A., Acs, Z.J. (1997): Local Geographic Spillovers Between University Research and High Technology Innovation. *Journal of Urban Economics*, 42. 422-448.

Arrow, K.J. (1962): The economic implications of learning by doing. *Review of Economic Studies*, 29. 166-173.

Audretsch, D.B., Feldman, M.P. (1996a): R&D Spillovers and the Geography of Innovation and Production. *American Economic Review*, 86(4), 253-273.

Audretsch, D. B., Feldman, M. P. (1996b): Innovative Clusters and the Industry Life Cycle. *Review of Industrial Organization* 11(2), 253-273.

Baum, J.A.C., Shipilov, A.V., Rowley, T.J. (2003): Where Do Small Worlds Come From? *Industrial and Corporate Change*, 12(4).

Boschma, R., Iammarino, S. (2008): Related variety, trade variety and regional growth in Italy. *Papers in Evolutionary Economic Geography*, 08/02.

Brenner, T. (2004): *Local Industrial Clusters. Existence, emergence and evolution.* Routledge, London New York.

De Carvalho, A.J., Saur-Amaral, I., Marques, M.J. (2006): Cooperation Networks and Regional Development Case of Multisectoral Partnership for Innovation. In: Cooke, P., Piccaluga, A. (eds.) *Regional Development in the Knowledge Economy.* Routledge.

Cohen, W.M., Levinthal, D.A. (1990): Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly*, 35, pp.128-52.

Coleman, J.S. (1990): *Foundations of Social Theory.* Cambridge MA, Belknap Press of Harvard University.

Cowan, R. (2006): Network models of innovation and knowledge diffusion. In: Breschi, S., Malerba, F. (eds.): *Clusters, Networks and Innovation*, Oxford University Press, Oxford, pp. 29-53.

Cowan, R., Jonard, N., Zimmermann, J-B. (2007): Bilateral Collaboration and the Emergence of Innovation Networks. *Management Science*, 53, pp.1051-1067.

Cowan, R., Jonard, N., Zimmermann, J-B. (2006): Evolving Networks of Inventors. *Journal of Evolutionary Economics*, 16(1-2), pp. 155-174.

Giuliani, E. (2007): Towards an understanding of knowledge spillovers in industrial clusters. *Applied Economics Letters*, 14(2), pp. 87-90.

Granovetter, M.S. (1973): The Strength of Weak Ties. *American Journal of Sociology*, Vol. 78, No. 6: pp. 1360.

Grossman, G., Helpman, E. (1991): *Innovation and Growth in the Global Economy*. Cambridge: MIT Press.

Jacobs, J.(1969): *The Economies of Cities*. New York: Random House.

Jaffe, A. (1986): Technological Opportunity and Spillovers of RD: Evidence from companies' Patents, Profits, and Market Value. *American Economic Review*, vol. 76(5), pp. 984-1001.

Jaffe, A., Trajtenberg, M., Henderson, R. (1992): Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics*. 108, 577-598.

Johansson, B., U. Forslund (2008): The analysis of location, co-location and urbanization economies. In *Handbook of Research on Cluster Theory*, ed. Karlsson, UK: Edward Elgar.

Karlsson, C. (2008)(ed.): *Handbook of Research on Cluster Theory*. Edward Elgar, Cheltenham, UK, Northampton, MA, USA.

Knott, A.M. (2003): Persistent Heterogeneity and Sustainable Innovation. *Strategic Management Journal*, 24 pp. 687-705.

Krugman, P.R. (1991): Increasing returns and economic geography. *Journal of Political Economy*, 99, pp. 48399.

Malmberg, A., Power, D. (2006): True Clusters. A Severe Case of Conceptual Headache. In: B. Asheim, P. Cooke, R. Martin (eds.) *Clusters in Regional Development, Regional Development and Public Policy Series*. Routledge: London.

Menzel, M-P., Fornahl, D. (2007): Cluster Life Cycles - Dimensions and

Rationales of Cluster Development. *Jena Economic Research Papers in Economics* 2007-076, Friedrich-Schiller-University Jena, Max-Planck-Institute of Economics.

Molina-Morales, F.X., Martínez-Fernández, M.T. (2004): How Much Difference is There Between Industrial District companies? A Net Value Creation Approach. *Research Policy*, 33, pp. 473-486.

Nelson, R. (1998): Selection Criteria and Selection Process in Cultural Evolution Theories. In: Ziman, J. (ed.): *Technological Innovation as an Evolutionary Process.* Cambridge University Press, Cambridge.

Nooteboom, B. (1999): *Inter-company Alliances: Analysis and Design.* Routledge.

Nooteboom, B. (2004): *Inter-company Collaboration, Learning and Networks.* Routledge.

Ozman, M. (2006): Knowledge integration and network formation. *Technological Forecasting and Social Change*, 73(9) pp. 1121-1143 .

Pavitt, K.L.R. (1998): Technologies, Products and Organization in the Innovating company: What Adam Smith Tells us and Joseph Schumpeter doesnt. *Industrial and Corporate Change* 7(3), 433-452.

Polányi, M. (1966): *The Tacit Dimension.* RoutledgeKegan Paul, London.

Porter, M.E. (1990): *The competitive advantage of nations.* Free Press: New York.

Pouder, R., St. John, C. H. (1996): Hot Spots and Blind Spots: Geographical Clusters of companies and Innovation. *Academy of Management Review* 21(4), 1192-1225.

Prencipe, A. (2000): Breadth and Depth of Technological Capabilities in CoPS: the case of the aircraft engine control system. *Research Policy*, 29, pp. 895-911.

Putnam, R.D., Leonardi, R., Nanetti, R.Y. (1993): *Making Democracy Work: Civic Traditions in Modern Italy.* Princeton NJ, Princeton University Press.

Romer, P. (1990): Endogenous Technological Change. *Journal of Political Economy*, 98, (5), part 2, S71-102.

Silverberg, G., Lehnert, D. (1994): Growth fluctuations in an evolutionary model of creative destruction. In: Silverberg, g., Soete, L. (eds.): *The*

Economics of Growth and Technical Change, Edward Elgar, pp. 74-108.

Smith, A. (1776): *The Wealth of Nations.* W. Strahan and T. Cadell: London

Solow, R.M. (1956): A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics*, 70. 65-94.

Sorenson, O. (2005): Social Networks, Informal Complexity and Industrial Geography. In: Fornahl, D., Zellner, C., Audretsch, D.B. (eds.): *The Role of Labour Mobility and Informal Networks for Knowledge Flows.* Springer.

Wuyts, S., Colombo, M.G., Dutta, S., Nooteboom, B. (2005): Empirical tests of optimal cognitive distance. *Journal of Economic Behavior & Organization*, vol. 58(2), pp. 277-302.

Weterings, A., R. Boschma (2006): The impacts of geography on the innovative productivity of software companies in the netwherlands. In *Regional Development in the Knowledge Economy*, ed. Cooke, Piccaluga, 63-83. UK: Routledge.

Varga, A., Parag, A. (2009): Academic knowledge transfers and the structure of international research networks. In: Attila Varga (ed.): *University knowledge transfers and regional development: Geography, entrepreneurship and policy.* Edward Elgar Publishers pp. 138-159.

7. Appendix

7.1. Mathematical derivations

First we show that the expected value of innovation equals its probability, p_0. Calculate the value of the knowledge level for company n at time t as follows:

$$V^n(t) = \sum_{i=1}^{w} k_i^n(t)$$

When evaluating the value of the innovation, consider the expected value of company n's knowledge base in the next period, which depends on the expected value of the individual knowledge fields:

$$E\left[V^n(t+1)\right] = \sum_{i=1}^{w} E\left[k_i^n(t+1)\right]$$

where E is the expected value operator. The expected knowledge level in technological field i in the next period can be written as follows:

$$E_i\left[k_i^n(t+1)\right] = \left[p_0\left(\frac{1}{w^n}\left(k_i^n(t)+1\right) + \frac{w^n-1}{w^n}k_i^n(t)\right) + (1-p_0)k_i^n(t)\right]\mathbf{1}_{k_i^n>0}$$

where $\mathbf{1}_{k_i^n>0}$ is the indicator function of $k_i^n > 0$, and w^n is the number of technological fields in which company n is competent:

$$w^n = \sum_i \mathbf{1}_{k_i^n>0}$$

Thus, the expected value of the knowledge base in period $(t+1)$ is:

$$E\left[V^n(t+1)\right] = \sum_{i,k_i^n(t)>0} p_0\left(\frac{1}{w^n}\left(k_i^n(t)+1\right) + \frac{w^n-1}{w^n}k_i^n(t)\right) + (1-p_0)k_i^n(t) =$$

$$= \frac{p_0}{w^n}\left(w^n \sum_{i,k_i^n(t)>0} k_i^n(t) + w^n\right) + (1-p_0)\sum_{i,k_i^n(t)>0} k_i^n(t) = \sum_i k_i^n(t) + p_0$$

where in the last equality we used the fact that $\sum_{i,k_i^n(t)>0} k_i^n(t) = \sum_i k_i^n(t)$ by definition. So we have a very simple form for the value of innovation in this case. The value of innovation can be simply characterized by the expected growth in the value of the knowledge base:

$$E\left[V^n(t+1)\right] - V^n(t) = p_0$$

Not too surprisingly, this tells us that if innovation increases one of the company's existing knowledge type by 1 unit, and if innovation occours with probability p_0 than the expected value of this innovation is p_0.

As carried out for the autharchic innovation above, the expected value of innovation can be derived for tha case of joint innovation as well. First, calculate the expected value of company n's knowledge base in period $(t+1)$ if it establishes an alliance with company m. The main difference in contrast to the separate innovation is two-fold. First, the porbability of successfull innovation now is $p_{n,m}$. Second, innovation can occour only in those knowledge fields, where both company n and company m have competence. According to these, the expected value of company n's knowledge base in technological field i if it cooperates with company m, is:

$$E_i\left[k_i^n(t+1)\right] = \begin{cases} p_{n,m}\left[\frac{1}{w^{nm}}\left(k_i^n(t)+1\right) + \frac{w^{nm}-1}{w^{nm}}k_i^n(t)\right] + \\ +(1-p_{n,m})k_i^n(t);\ \forall i, k_i^n(t)>0, k_i^m(t)>0 \\ k_i^n(t)>0;\ otherwise \end{cases}$$

where w^{nm} is the number of technological fields in which both company n and company m has competence:

$$w^{nm} = \sum_i \mathbf{1}_{k_i^n>0, k_i^m>0}$$

From the above, by summing through i, we get the expected value of company n's knowledge base:

$$E\left[V^n(t+1)\right] = \sum_i k_i^n(t) + p_{n,m}$$

from which the value of joint innovation is:

$$E\left[V^n(t+1)\right] - V^n(t) = p_{n,m}$$

Of course, this method does not take into consideration, that in several cases companies do not have common technological fields. Instead of building this into the formulae above, we simply rule out this possibility by assuming that companies do not form alliances if they have no common technological fields. In this case, the equations above are correct, as if company n and company m do not have common fields, they do not evaluate the value of their joint work, as it is apparently zero.

7.2. The measure of heterogeneity (normalised average cognitive distance)

Let's denote the Eucledian distance of company n and company m by $d_{n,m}$ as defined in the model description. Then the average distance of company n from all other companies in the knowledge space can be written as

$$\bar{d_n} = \frac{\sum_m d_{n,m}}{N-1}$$

Averaging over n gives the average cognitive distance in the whole industry:

$$\bar{d} = \frac{\sum_n \bar{d}_n}{N} = \frac{\sum_n \sum_m d_{n,m}}{N(N-1)}$$

However, this is not independent of the size of knowledge space. If the knowledge space widens, average distance may (but not neccesarily) increase as well. To rule this effect out first we calculate the diameter of the knowledge space (as the largest possible distance between two points in this space):

$$s = \sqrt{\sum_{i=1}^{w} (k_i^{max})^2}$$

where k_i^{max} is the technological frontier in technological field i which can be different across fields and periods. To yield a normalised measure of heterogeneity we relate average distance to the diameter, thus obtaining a value between zero and one with zero desiganating a homogenous industry with similar knowledge bases and one designating the most heterogeneous industry with regards to cognitive distance:

$$AD = \frac{\bar{d}}{s}$$

7.3. The measure of locality (clustering coefficient and local density)

The clustering coefficient refers to the extent to which neighbours of a given agent in the network are neighbours to each other (Cowan, 2006). Consider agent i in the network, its neighbourhod is Γ_i. The number of possible links in this neighbourhood is $\|\Gamma_i\| \cdot (\|\Gamma_i\| - 1)/2$. If we have exactly this number of links, then this subnetwork is totally clustered as the neighbours of agent i are all neighbours to each other. If this is not the case, we can count this deviation by summing the links among agent i's neighbours. Definig $X(j,l)$ as the indicator function of j belonging to Γ_i given that $l \in \Gamma_i$,[21] we have

$$\sum_{j,l \in \Gamma_i} X(j,l)$$

[21] That is, $X(j,l) = 1$ if $j \in \Gamma_i$ and $X(j,l) = 0$ otherwise.

as the number of links in agent i's neighbourhood. Normalizing with the possible number of links we have

$$C = \sum_{j,l \in \Gamma_i} \frac{X(j,l)}{\|\Gamma_i\| \cdot (\|\Gamma_i\| - 1)/2}$$

which is 1 in the case of total clustering and 0 if agent i's neighbours do not have links with each other. By summing through i and normalizing by the number of agents in the network, we can derive the clustering coefficient of the whole network which moves between zero and one reflecting the clustering of the network.

$$AC = \frac{1}{N} \cdot \sum_i C_i = \frac{1}{N} \cdot \sum_i \sum_{j,l \in \Gamma_i} \frac{X(j,l)}{\|\Gamma_i\| \cdot (\|\Gamma_i\| - 1)/2}$$

As mentioned in the paper, the clustering coefficient may be a misleading measure of the local structure of a network if the links in a network increase ore decrease over time. If this is the case, the clustering coefficient may increase while there is no essential change in the network structure as the whole network becomes more dense: increase in clustering stems from the increase in average degree and local and global ties may form with equal possibility. In order to rule out this bias we propose a measure of 'network locality', which is simply the ratio of the clustering coefficient as defined above and the global density of the network. By constructing this measure we use the fact that the clustering coefficient is a special density measure, i.e. it relates the number of links in a neighbourhood to all possible links in that neighbourhood. Let us denote the number of links (i.e. the degree) of agent i by D_i. In this case global density can be written as

$$G = \frac{\sum_i D_i}{N \cdot (N-1)/2}$$

Our network locality meaure thus comes as

$$L = \frac{AC}{G} = (N-1) \cdot \sum_i \sum_{j,l \in \Gamma_i} \frac{X(j,l)}{\|\Gamma_i\| \cdot (\|\Gamma_i\| - 1)} \cdot \frac{1}{\sum_i D_i}$$

Previous Issues:

Varga Attila: From the geography of innovation to development policy analysis: The GMR-approach (2007/1)

Bessenyei István: Növekedési pólusok a térben és a társadalomban (2007/2)

Darvas Zsolt - Schepp Zoltán: Kelet-közép-európai devizaárfolyamok elorejelzése határidos árfolyamok segítségével (2007/3)

Varga Attila: GMR-Hungary: A Complex Macro-Regional Model for the Analysis of Development Policy Impacts on the Hungarian Economy (2007/4)

Reiff Ádám - Zsibók Zsuzsanna: Az infláció és az árazási magatartás regionális jellemzői Magyarországon, mikroszintű adatok alapján (2008/1)

Varga Attila - Parag Andrea: Egyetemi tudástranszfer és a nemzetközi kutatási hálózatok szerkezete (2008/2)

Schepp Zoltán - Szabó Zoltán: Felsőoktatás-politika és állami finanszírozás: a 2007. évi felvételi tanulságai a gazdaságtudományi alapképzésben (2008/3)

Kaposi Zoltán: Város és agrárrendszer a polgárosodás korában (1850-1914) (a mezőgazdaság változásai Nagykanizsán) (2008/4)

Barancsuk János: Néhány gondolat az árelfogadó és ármeghatározó fogalmak jelentéséről (2009/1)

Kiss Gy. Kálmán: A szövetkezeti bank megteremtésének kísérlete Magyarországon (2009/2)

Zeller Gyula: Létezik-e a Smith probléma, avagy mennyire egységesek Adam Smith nézetei? (2009/3)

Járosi Péter - Atsushi Koike - Mark Thissen - Varga Attila: Regionális fejlesztéspolitikai hatáselemzés térbeli számítható általános egyensúlyi modellel: a GMR-Magyarország SCGE modellje (2009/4)

Mellár Tamás: Felemás magyar modernizáció (2009/5)

Zoltán Szabó: Az új paternalizmus: a nem-racionális hitelfelvevői magatartás és a túlzott eladósodás néhány gazdasági viselkedéstani összefüggése (2009/6)

Erdős Katalin-Varga Attila: Az egyetemi vállalkozó: legenda vagy valóság az európai regionális fejlődés elősegítésére? (2009/7)