The coevolution of endogenous knowledge networks and knowledge creation

Citation for published version (APA):

Document license:
CC BY-NC-ND

DOI:
10.1016/j.jebo.2017.11.023

Document status and date:
Published: 01/01/2018

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
https://www.tue.nl/index.php?id=71870

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
The coevolution of endogenous knowledge networks and knowledge creation

Elena M. Tur*a,b,c,*, Joaquin M. Azagra-Caro a

a INGENIO (CSIC-UPV), Universitat Politècnica de València, Camino de Vera s/n, E-46022 Valencia, Spain
b School of Innovation Sciences, Eindhoven University of Technology, The Netherlands
c Institute for Innovation and Entrepreneurship, Department of Economics and Society, School of Business, Economics and Law, University of Gothenburg, Sweden

A R T I C L E   I N F O

Article history:
Received 20 July 2016
Received in revised form 10 October 2017
Accepted 24 November 2017
Available online 26 November 2017

Keywords:
Knowledge creation
Collaboration
Simulation model
Endogenous network

A B S T R A C T

Knowledge creation is increasingly a collaborative process, but empirical studies provide conflicting evidence on whether the relation between knowledge creation and number of collaborators is positive, negative, or nonexistent. The simulation model developed in this paper offers a deeper formal theoretical understanding and analyzes the feedback between the processes of knowledge creation and network collaboration. The model is formed by two functions, one for the formation of the network and another for the creation of knowledge, that suffice to reproduce the three coevolution scenarios described in the empirical literature. Due to the feedback mechanisms between the two functions, changes in one of the parameters deeply affect the outcome of the model, both in the amount of knowledge produced and the structure of the resulting network, as well as in the relation between them. Analyses of collaborative knowledge creation would benefit from taking into account this feedback.

© 2017 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Simulation models are useful for explaining evolutionary processes. They have been used widely to analyze knowledge networking, the process by which agents interact to create knowledge in inventor network models (Cowan et al., 2006), researcher collaboration models (Grebel, 2012), and interfirm research and development (R&D) alliance networks (Ahrweiler et al., 2004). Simulations have been used also to investigate knowledge creation by agents in a network. In those studies, knowledge is an abstract idea (Cowan and Jonard, 2003), or a concrete output of the abstract knowledge, that can be measured by number of scientific papers in the case of researchers (Borner et al., 2004), new products in the case of firms (Malerba et al., 1999), etc.

Many empirical studies have analyzed the creation of knowledge in networks (for a review, see Ozman, 2009; Phelps et al., 2012). Most empirical studies use measures of the network related to one agent (the ego network) to explain the agent’s output but ignore possible feedbacks. Thus, results are mixed and inconclusive. The simplest measure of the ego network, the number of collaborators (or degree), has been used in numerous studies to try and determine how collaboration affects performance, in terms of knowledge creation or innovative outcomes. These studies can be grouped in three main categories:

* Corresponding author at: Technische Universiteit Eindhoven, P.O. Box 513, 5600MB Eindhoven, The Netherlands.
E-mail address: e.m.mas.tur@tue.nl (E.M. Tur).

https://doi.org/10.1016/j.jebo.2017.11.023
0167-2681/© 2017 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
those that suggest that the relation between collaboration and performance is positive, those that find a negative relation, and those that do not find a significant relation between them.

The first group of studies indicates that knowledge creation is a collaborative process, that is, that the more intensely agents collaborate, the more knowledge they create (Ibarra, 1993; Ahuja, 2000; Cohen et al., 2002; Cassiman and Veugelers, 2006; Boschma and ter Wal, 2007). This kind of coevolution occurs when collaborations provide agents with resources and new information (Ahuja, 2000), at least if coming from outstanding researchers (Borjas and Doran, 2015). A second group of studies suggests that the number of collaborators is not a determinant of agents’ innovative performance (Dakhlali and De Clercq, 2004; Bell, 2005; Vega-Jurado et al., 2009; Martinez-del Rio and Cepedes-Lorente, 2013). For example, Bell (2005) finds no significant relation between the number of a firm’s formal ties and its innovativeness, because institutional ties transmit only relatively well-known information. Finally, a third group finds that collaborating can harm performance (Woolcock, 1998; McFadyen and Cannella, 2004; Laursen and Salter, 2006; Carayol et al., 2008; Molina-Morales and Martinez-Fernandez, 2009; Grimpe and Kaiser, 2010). McFadyen and Cannella (2004) suggest that the more collaborations an individual is forced to maintain, the less effort can be focused on knowledge creation. Thus, at some point, agents with too many collaborators may start to underperform compared to those with fewer partners. However, there is also evidence of the opposite idea: knowledge collaboration may be beneficial for knowledge production only after having reached a minimum threshold of collaborators (Mao et al., 2016).

These different studies consider an exogenous network structure, that is, they take the structure as given and predict its impact on knowledge creation. This is not realistic, mainly because agents choose their collaborators for specific reasons, including reputation or previous performance (Wagner and Leydesdorff, 2005; Balland et al., 2012). Thus, network formation and knowledge creation are endogenous processes. The structure of the network affects the output of agents, which determines the future structure of the network. Previous empirical studies do not account for this feedback, which may explain the ambiguous results they obtain.

This feedback has been taken into account in recent simulation studies on endogenous networks. For example, Dawid and Hellmann (2014) and Savin and Egchetokun (2016) consider companies that form strategic alliances, based on the maximization of an objective function. Those theoretical models allow switching between different strategies (e.g., single innovators, collaborative innovators and imitators) to reproduce some empirical regularities and then design strategies and policies to improve technological innovation, but do not deal with the conflicting empirical evidence on the relation between networks and the performance of agents. We will fill this gap by generalizing companies to any type of knowledge producer and technological innovation to any knowledge output, designing a simulation model of an endogenous and evolving network of agents that create knowledge. Our model does not intend to describe and replicate how knowledge producers interact to create knowledge, nor does it assume any kind of rationality for their collaborations. Instead, agents are attracted to one another for collaborations, and they create knowledge without any assumption on their strategies or objectives in doing so. Simulations of the model generate different theoretical scenarios of the coevolution of knowledge networks and knowledge creation.

Agents choose partners for obtaining knowledge, thus creating a network (Baum et al., 2010). However, few studies propose a theoretical modeling of the interaction between the network and the creation of knowledge. The works by Cowan et al. (2004) and Cowan and Jonard (2007) emphasize the complex nature of knowledge creation through collaboration. In this study we will propose a model that builds on theirs and additionally analyzes the amount of knowledge created in the network. Guimera et al. (2005) propose a mathematical model of knowledge networking, and find empirical correlations with knowledge generation, but this does not form part of the mathematical model. Chen et al. (2009) explain how connecting structural holes among networks contributes to knowledge advances, but how the knowledge base is created at every period is beyond their objective. Brummitt et al. (2015) do inquiry about the mutual relation between knowledge networks and creation, but model knowledge creation as the probability of two partners matching compatible knowledge. Hence, they do not address the economic problem of how costs and benefits of knowledge diffusion form part of knowledge production.

The main goal of this study is to call for attention on the importance of taking into account the feedback between knowledge creation and network formation when studying knowledge creation in networks. The paper is structured as follows. Section 2 presents the coevolution model of knowledge creation and knowledge networks. Section 3 presents the results of the simulations. Section 4 discusses the different results and concludes the paper.

2. The model

Let us consider a set $S = \{1, \ldots, n\}$ of agents that interact over $T$ periods of time. In each step $t \in \{1, \ldots, T\}$, they form a network and create a certain amount of knowledge $k(i,t)$. The network is represented by its adjacency matrix $\Omega_t$, where $\Omega_t(i,j)$ takes the value 1 if agents $i$ and $j$ collaborate in period $t$, and 0 otherwise. The degree (or number of collaborators) of agent $i$ in period $t$ is $d_t(i) = \sum_j \Omega_t(i,j)$. This is a basic indicator to measure the ego network, often used in the empirical literature (see Cooke and Wills, 1999, or Bell, 2005).

The network in each period is formed depending on the network and the knowledge created in the previous periods. As the network is generated in each period, links are allowed to break and form over time. Similarly, the amount of knowledge created in each period depends on the amount of knowledge previously generated, and on the structure of the network. As the process develops, agents become heterogeneous in their knowledge endowments and ego networks. A link between two
agents can be created in any period even if they did not collaborate in the past. Likewise, two agents can break an existing collaboration if the link is not updated in a later period.

2.1. Knowledge creation

In order to get a model as flexible as possible, we consider knowledge in a very abstract way as in Cowan et al. (2004). Thus, agents in our model could be researchers, inventors, firms, universities, or any other type of agent involved in knowledge creation with the possibility to collaborate.

Prior research has shown that the performance of an agent in a knowledge network can be deeply influenced by the structure of the network (De Solla Price, 1965; Guler and Nerkar, 2012). On the one hand, collaborations can provide with resources, new information or new ideas (Ahuja, 2000). This is captured by parameter $\theta \geq 0$, the positive effect of collaborations in knowledge creation. On the other hand, collaborations can be costly (McFadyen and Cannella, 2004; Ozman, 2009), with the result that a very large number of collaborations can hamper knowledge creation. Additionally, the cost of maintaining an additional collaboration increases with the number of collaborations an agent already maintains. This is captured by parameter $\gamma \geq 0$, the cost of collaborating, together with the square of the number of collaborations $d_i$.

Finally, knowledge is a cumulative process: new knowledge can be created from previous knowledge (Jaffe et al., 2000). Parameter $\alpha$ measures how much new knowledge is created from the stock of knowledge of an agent, and hence, accounts for the cumulativeness of knowledge. The length of the time window is $t$, which is the number of periods before knowledge becomes obsolete.

Eq. (1) shows the functional form for the creation of knowledge. No negative amounts of knowledge are allowed. The amount of knowledge created by agent $i$ at time $t$, $\kappa(i, t)$, depends on the structure of its ego network and on the stock of knowledge it possesses. As stated, $\theta$ is the positive effect of collaborations, $\gamma$ is the cost of collaborating and $\alpha$ is the knowledge produced from the stock.

$$
\kappa(i, t) = \frac{\theta}{t + 1} \sum_{s=t-\tau}^{t} d_s(i) - \gamma d_s(i)^2 + \alpha \frac{t-1}{t} \sum_{s=t-\tau}^{t-1} \kappa(i, s)
$$

(1)

This functional form implicitly assumes that there are only two possible sources of new knowledge for an agent: collaborations and the pool of knowledge the agent already possesses. Due to the recombinant nature of knowledge explored in previous theoretical studies (Konig et al., 2011, 2012; Sole et al., 2016), if an agent never collaborates, the possible number of new combinations of its existing knowledge are limited. At some point, the agent will be unable to continue to create new knowledge unless it starts to collaborate. Given this functional form for the creation of knowledge, $\alpha$ is necessarily bounded to $[0, 1]$.

This function is similar to that by Konig et al. (2011, 2012). In their model, the amount of knowledge created by an agent is a recombinant of the knowledge stocks of the agent and his neighbors. In line with this approach, we add a new dimension to the effect that the knowledge of neighbors has on the knowledge of an agent. In our model, the knowledge stocks of potential collaborators influence the amount of knowledge agents create, through their influence on whether or not they become collaborators. Although it does not appear explicitly in the function of knowledge creation, it implicitly affects the result.

2.2. Network formation

In a knowledge network, agents break and establish links as a result of strategic decisions (Barabasi et al., 2002; Fleming and Frenken, 2007). These strategic decisions usually depend on two main components, previous history and attractiveness of agents (Ahrweiler et al., 2004). More experienced and more successful agents are more likely to find partners (Wagner and Leydesdorff, 2005; Balland et al., 2012), so the probability of collaborating increases with the attractiveness of the agent. In our model, the attractiveness of an agent depends on the amount of knowledge it created previously, relative to the knowledge created by the other agents in the network. Moreover, previous collaboration increases the willingness to engage in joint knowledge creation (Cowan et al., 2006; Baum et al., 2010), so the probability that a link is formed is higher for pairs that have already collaborated.

We model the probability that agent $i$ collaborates with agent $j$ as a linear combination of their previous history and the attractiveness of agent $j$ (Eq. (2)), respectively weighted by $\lambda$ and $1 - \lambda \in [0, 1]$. The previous history of a couple of agents is the number of times they have collaborated in the recent past, during the time window. The attractiveness of an agent is the knowledge it has created in a period as a proportion of the maximum amount of knowledge created by itself or another agent in that period. If the agent is the maximum knowledge producer, this proportion will be equal to 1. This function follows a preferential attachment dynamic (Barabasi and Albert, 1999; Albert and Barabasi, 2002), as agents with a high number of collaborators are likely to attract even more collaborators.

$$
P(i \rightarrow j, t) = \lambda \frac{1}{t} \sum_{s=t-\tau}^{t-1} \Omega_s(i, j) + (1 - \lambda) \frac{1}{t} \sum_{s=t-\tau}^{t-1} \max_k \kappa(j, s)
$$

(2)
By taking $\lambda$ and $1 - \lambda$ in $[0, 1]$ instead of taking them in $[0, 1]$, we leave out the extreme cases where collaboration is only determined by either previous history or attractiveness. Those extreme cases can lead to some peculiar behaviors. For instance, if $\lambda = 1$, then two agents that have collaborated for $\tau$ periods will collaborate with probability 1 for all subsequent periods.

The probability that agents $i$ and $j$ collaborate is that both of them collaborate with the other (Eq. (3)). In considering probabilities we account for their being willing to collaborate but unable to do so for some reason. The probability that a link breaks, that is, that a collaboration in time $t - 1$ does not continue in time $t$, is $1 - P(i \rightarrow j, t | i \in \Omega_{t-1}, (i, j) = 1)$, the probability that it does not form in $t$ given that it existed in $t - 1$. The probability that it breaks after $s \in \{1, \ldots, \tau\}$ periods is defined likewise. The development of the model is, thus, not deterministic.

$$P(i \rightarrow j, t) = P(i \rightarrow j \cap j \rightarrow i, t)$$

(3)

Notice that agents’ actions are not the result of an explicit decision making process since they do not develop strategies that maximize some objective function, or compete over knowledge, as in the theoretical literature (see e.g. Azagra-Caro et al., 2008 or Westbrock, 2010) or previous work on the coevolution of knowledge and networks (Cowan et al., 2004; Llerena and Ozman, 2013). In this paper, we model the probability of a collaboration between two agents, not the rationale for their collaboration which is beyond the objectives of this study.

A summary of the parameters can be found in Table 1.

### 3. Simulation results

In this section we present different scenarios for several sets of parameters. For the simulations, we consider a set of $n = 200$ agents interacting for $T = 500$ periods. The results of the simulations are not qualitatively different depending on the length of the time window $\tau$. In the simulations we use a value of $\tau = 1$, although the results are similar for other window lengths. Every agent starts with 1 unit of knowledge. Every pair of agents is connected with probability $2/n$, which results in a moderately dense initial network. Similar results were obtained with other initial conditions, like binomial knowledge endowments or other probabilities for the initial network. The initial conditions were the same in all the settings but resulted in very different final configurations. At the beginning of the simulation, agents are homogeneous in their knowledge endowment and heterogeneous in their ego network, and as the process unfolds, they become heterogeneous in both the knowledge they have created and the network they are forming. All the codes and simulations are run with R (R Core Team, 2014).

#### 3.1. Positive, negative and independent coevolution

The behavior of the model is depicted graphically. Figs. 1–3 are examples of the three possible scenarios described in the empirical literature: positive, negative, and independent relation between collaboration and knowledge creation.  

When the positive effect of collaborations, $\theta$, is high enough compared to the cost of maintaining a collaboration, $\gamma$, collaborations are profitable. This results in a positive relation between the number of collaborations and the creation of knowledge (Fig. 1). Fig. 1a shows an example of a reinforcing effect in knowledge creation for the whole population. Most agents create moderate amounts of knowledge, with a low number of collaborations. A few of them, nonetheless, are more productive and attract more collaborators, and thus produce even more and are more attractive in the following steps.

As a robustness check, we run 20 simulations with the same parameter settings and plot the last period of all simulations. Since we overlap the points of all runs, spurious relations between the different scatter plots could appear (for instance, if two runs gave identical independent plots, but one of them higher and to the right of the other, the overlapped plot would look like a positive plot). Thus, the robustness check is presented as an additional plot (Fig. 1b), in addition to the snapshots of single simulations. Fig. 1b shows the last period of the simulation for 20 simulations with the same parameters as Fig. 1a. The overall aspect of both figures is similar, even though some of the agents in some of the runs in Fig. 1b reached higher

---

1. Results for other values of the window length $\tau$ are available under request.
2. There is a fourth theoretical scenario: in some cases, the coevolution is attracted to a “no collaboration, no knowledge” steady state. Since $\alpha < 1$, if no one collaborates ever, the knowledge pool will become exhausted eventually. If no one collaborates in any period, no knowledge can be created in the long run, and since no one has collaborated and no knowledge is being created, no one will collaborate in subsequent periods. Since this scenario is straightforward and does not show any interesting behavior, we will leave it out of our analysis.
Fig. 1. Positive coevolution.

(a) Whole population, final outcome (one run)
\[ \alpha = 0.9, \, \theta = 0.1 \]
\[ \gamma = 0.0001, \, \lambda = 0.2 \]

(b) Whole population, final outcome (100 runs)
\[ \alpha = 0.9, \, \theta = 0.1 \]
\[ \gamma = 0.0001, \, \lambda = 0.2 \]

(c) A single agent, dynamics (one run)
\[ \alpha = 0.9, \, \theta = 0.1 \]
\[ \gamma = 0.0001, \, \lambda = 0.2 \]

Fig. 2. Negative coevolution.

(a) Whole population, final outcome (one run)
\[ \alpha = 0.1, \, \theta = 0.5 \]
\[ \gamma = 0.01, \, \lambda = 0.2 \]

(b) Whole population, final outcome (100 runs)
\[ \alpha = 0.1, \, \theta = 0.5 \]
\[ \gamma = 0.01, \, \lambda = 0.2 \]

(c) Whole population, final outcome (one run)
\[ \alpha = 0.9, \, \theta = 0.5 \]
\[ \gamma = 0.01, \, \lambda = 0.8 \]

(d) Whole population, final outcome (100 runs)
\[ \alpha = 0.9, \, \theta = 0.5 \]
\[ \gamma = 0.01, \, \lambda = 0.8 \]

(e) A single agent, dynamics (one run)
\[ \alpha = 0.9, \, \theta = 0.5 \]
\[ \gamma = 0.01, \, \lambda = 0.8 \]
amounts of knowledge and degree. This shows that the aggregated coevolution pattern depends only on the parameters of the model, not on the initial conditions. On the other hand, the particular amounts of knowledge and degree do depend on how the simulation develops, that is to say, on the probabilities involved in the network function (Eq. (3)).

Fig. 1a and b show the results for the whole population in a single period, the last one. The dynamics for a single agent are depicted in Fig. 1c: the more knowledge it creates, the more collaborators it gets, and the more collaborators it has, the more knowledge it creates. This leads to the differentiation between those agents with low levels of both knowledge and collaborations, and those that attract most collaborations and create higher amounts of knowledge.

If the cost of maintaining collaborations, $\gamma$, is high compared to their positive effect, $\theta$, every additional collaboration is prejudicial to the agent’s performance. In this case, the amount of knowledge created is decreasing in the number of collaborations, and the observed coevolution is negative (Fig. 2). The existence of a limitation to collaboration is a necessary condition for this coevolution pattern. In some cases, the cost of collaborating can be so high that agents have too many collaborations to produce any knowledge (Fig. 2a and b). In other cases, the relation can be negative but the amounts of knowledge created can still be high enough to avoid the appearance of unproductive but collaborative agents (Fig. 2c and d). In negative coevolution scenarios, the most productive agents become the most attractive. As their number of collaborations increases, they create less and less new knowledge and become less and less attractive. At some point, they have a smaller number of collaborations which allows them to perform better, and thus to become more attractive again. These dynamics are depicted in Fig. 2e for a single agent.

Finally, the amount of knowledge created through collaborations may be similar to the amount based on previous knowledge. Due to the mixed effect of both sources of new knowledge, the performance of agents with many collaborators, and agents with small numbers of collaborators but a large stock of knowledge, is similar. Thus, the coevolution is independent, since the number of collaborators seems not to affect the creation of knowledge (Fig. 3). This last case shows a parameter setting in which agents can have a high number of collaborations. Notice that our agents can be any kind of knowledge producers. In certain academic disciplines, for instance, very high numbers of collaborators are plausible: the original human genome paper (doi:10.1038/35057062) has 249 authors, and the Higgs Boson paper (doi:10.1016/j.physletb.2012.08.021) is signed by 2891 authors.

3.2. Feedback mechanisms in the model

In this section we show the determinant role of the feedback between the two functions. Some variations in a single parameter are illustrative of the complexity of the model described. The straightforward effects of the model are that changing parameters in Eq. (1) varies the amount of knowledge created, and that changing parameters in Eq. (2) varies the network structure. Nonetheless, the model also shows the indirect effects of each equation in the outcome of the other, due to the relation between the two functions. Figs. 4 and 5 are examples of the complex behavior of the model due to the feedback between the processes of knowledge creation and network formation.

First, we show how changes in the network formation parameters affect the results of the simulation. Fig. 4 shows the effect of increasing $\lambda$ (the weight of having previously collaborated on the probability to collaborate) from 0.2 to 0.5. Notice that this is the only changing parameter, both in the knowledge creation function and in the network formation function.
When \( \lambda = 0.2 \), the probability to collaborate depends mainly on the attractiveness of agents; while when \( \lambda = 0.5 \), it depends also on whether or not the two agents have collaborated in the previous steps. With this increase in \( \lambda \), the process can change the type of coevolution case. Fig. 4a shows a switch from an independent to a negative coevolution scenario, and Fig. 4b depicts a swap from a negative to a positive coevolution. In both cases, the apparent relation between the number of collaborators and the amount of knowledge is not driven by the knowledge creation function but by the network formation process. In Fig. 4a, with a high value of \( \lambda \), only the most productive agents attract new collaborators. This leads to a dynamic similar to the one previously depicted in Fig. 2e. Likewise, in Fig. 4b the relation is positive not because agents with more collaborators create more knowledge, but because the most productive agents attract more collaborators. In this case, the positive coevolution scenario performs worse than the negative coevolution case. When agents can choose their collaborations freely, based on attractiveness rather than previous history, the overall amount of knowledge created is higher.

A similar behavior can appear for changes in parameters of the knowledge function as well. Fig. 5 shows two examples of increasing \( \gamma \), the cost of collaborating, from 0.001 to 0.01. Of course, the amount of knowledge created will be higher for lower values of this cost. Moreover, increasing \( \gamma \) can lead to a change in the structure of the resulting knowledge network. Fig. 5a shows a case where the cost of collaborating is so high that it suffocates the whole process: the amount of knowledge created is lower, and the network does not develop. Fig. 5b, on the other hand, shows a switch from a positive to an independent coevolution. In the positive coevolution, some agents have the role of “star scientists”: they create a much higher amount of knowledge than the rest, and they are more attractive to collaborators. With the increase in \( \gamma \), the reinforcement mechanism disappears and all agents create similar amounts of knowledge. They are also similarly attractive, so the degrees are more uniformly distributed in the resulting network.

### 3.3 Characterization of the scenarios

Previous sections presented snapshots of the model, without a systematic exploration of the parameter values. In this section we present a more structured analysis of the scenarios that different combinations of parameters present. The model has five main parameters: the positive effect of collaboration in knowledge creation \( \theta \), the negative effect of collaborating \( \gamma \), the effect of the stock of knowledge \( \alpha \), the weight of previous collaboration in the probability to collaborate \( \lambda \), and the length of the time window \( r \). As already commented, changing the length of the time window does not qualitatively affect
In the regression summary, the Table 2 simulations output 25 times DC2 combinations knowledge 0.5, behavior 0.9. Table 2: The simulations of the output of the simulations. We keep one more parameter fixed, γ = 0.01, and compare the scenarios that appear in the combinations of three different levels (high, medium, low) of the other three parameters. We run the simulations $R = 25$ times for every combination of parameters.

The outcome of the different runs can either follow a very similar pattern or show great variation. Fig. 6 shows the 25 simulations for two sets of parameters.\(^3\) In the first case (Fig. 6a), there is little variability in the final outcome of the simulation: the behavior of the model is clearly determined by the parameter values. In the second case (Fig. 6b), the final behavior is more dependent on the stochastic component of the process.

Table 2 summarizes the different types of coevolution that can appear for every combination of $\theta \in \{0.1, 0.5, 0.9\}, \alpha \in \{0.1, 0.5, 0.9\}$, and $\lambda \in \{0.2, 0.5, 0.8\}$. The type of coevolution is determined by the slope of a linear regression on the amount of knowledge created and the degree of agents in the last run of the simulation.\(^4\) The negative sign in the cell for $\alpha = 0.9$ and $\theta = 0.1$ in Table 2c ($\lambda = 0.8$) indicates that all simulations for that combination of parameters fitted a linear regression with a

---

3 The results for all the combinations of parameters are available on demand.

4 The results of the regressions for all the combinations of parameters are available under request.
statistically significant negative slope: that is to say, they all showed a negative coevolution. On the other hand, the positive and negative signs in the cell for \( \alpha = 0.9 \) and \( \theta = 0.9 \) of that same table indicate that some of the slopes were positive while some others were negative: this combination of parameters can show scenarios both of positive coevolution and of negative coevolution.

The main result from this analysis is the non-monotonic nature of the model. This can be seen, for example, for \( \lambda = 0.8 \) and \( \alpha = 0.5 \) (second row of Table 2c). For \( \theta = 0.1 \), the relation between knowledge and degree is positive (that is to say, the linear regression fitted to the results of all the simulations has a positive slope). Increasing \( \theta \) to 0.5 changes this relation to negative. As already discussed, this leads to some agents becoming too attractive. If this effect was monotonic, increasing \( \theta \) even further would result in more positive relations. Nonetheless, for \( \theta = 0.9 \) the relation is negative again. This non-monotonic effect can also be seen in other parameter combinations, like \( \lambda = 0.8 \) and \( \theta = 0.1 \), or like \( \alpha = 0.1 \) and \( \theta = 0.9 \).

A second result is that the case with a low value of \( \lambda \) presents the most variability of outcomes. That is to say, when previous collaboration has a small weight on the probability of collaborations, the coevolution type is not determined. This means that, from a same set of parameters, it is not possible to tell beforehand whether the coevolution will be negative or independent or positive.

This highlights the effect of feedback in the process that was already discussed in the previous section: if we pool all scatterplots of all simulations with a same set of parameters, we miss the behavior of the individual runs. Indeed, if we aggregate the degree and amount of knowledge created in the 25 simulations in a single database and run the same linear regression, the slope of the relation turns out to be negative in all the undetermined cases of \( \lambda = 0.2 \). In other undetermined cases (for example, for \( \lambda = 0.8 \), \( \alpha = 0.9 \), \( \theta = 0.9 \)), the slope turns out positive. In previous sections we have followed this conventional approach (in Figs. 1b and 2b) to represent the behavior of several runs. This result shows that this approach can, actually, be misleading. This further advises against using methods that do not take into account the feedback mechanisms of knowledge creation and network formation, when analyzing collaborative knowledge creation.

Moreover, the case with a low \( \lambda \) shows a counterintuitive result: increasing the positive effect of collaborations \( \theta \) does not necessarily guarantee a positive relation between collaboration and knowledge creation. For example, for a medium effect of the knowledge pool \( \alpha \), the simulations with both a low and a medium effect of collaboration \( \theta \) show cases with a positive coevolution, while the scenarios with a high \( \theta \) only show negative or independent coevolution cases. This non-linear effect of increasing \( \theta \) shows for all values of \( \lambda \), and it is due to the importance of network formation in determining the type of coevolution (which we have already mentioned in the previous section). In these cases, the positive coevolution dynamics are not so much due to the fact that people with many collaborators create more knowledge (otherwise the more \( \theta \) increases, the more positive that relation would be). Rather, it is due to the fact that successful people are more attractive for collaborations.

The same goes for increasing the effect of the pool of knowledge \( \alpha \). The intuitive result would be that, if most knowledge is created from collaborations as compared to the existing pool of knowledge (top right corner), the relation between collaborations and knowledge should be positive, while if most knowledge is created from the pool instead of collaborations (bottom left corner), the relation should be closer to independent. Instead, what we find is that increasing \( \alpha \) has a variety of effects, from showing scenarios with a tendency towards negative coevolution (for \( \lambda = 0.8 \)), to increasing the tendency for positive coevolution (for \( \lambda = 0.8 \) and \( \theta = 0.9 \)).

4. Discussion and conclusions

This paper presents a simulation model of the coevolution of knowledge networks and knowledge creation. Knowledge appears increasingly to be an interactive process, and a deeper understanding is needed to improve the efficiency of the performance of the system. Different coevolution scenarios can arise depending on the importance of the collaborations for knowledge creation, the role of previous knowledge, and the process of partner selection. Two rules of behavior reproduce the scenarios in the empirical literature, which correspond to different cases of apparently conflicting empirical evidence. Thus, all those different cases of knowledge creation through collaboration may originate in a single process.

The main result of the paper is the importance of feedback between the knowledge creation process and the formation of the network. Modifying the partner selection process from depending mainly on attractiveness to depending also on whether two agents have already collaborated leads to changes not only on the structure of the network, but also on the amount of knowledge agents produce overall. It can also lead to a change of coevolution type (from negative to positive, for example, or from independent to negative). Likewise, changes in the knowledge creation function, like a variation in the cost of collaborations, produces changes not only on the amount of knowledge created overall, but also on the type of coevolution and the structure of the resulting network. These feedback mechanisms must therefore be accounted for in studies aiming to analyze a process of knowledge creation in a network.

The results draw attention to several other features of the process. Remarkably, the positive relation between the number of collaborators and the performance of agents may be due to partner selection rather than knowledge creation. This has implications for researchers since it points to the importance of taking into account the endogeneity of the network when analyzing the effect of collaborations in knowledge creation.

The process of knowledge creation can be hampered by “myopic” partner selection (Carayol and Roux, 2009) based on previous history rather than attractiveness. If agents are bound to their previous collaborations, the overall performance of the system is lower; if they can freely establish new links or break existing ones, the levels of performance achieved by the
system are higher. If previous history is important because of high levels of uncertainty and instability, one solution might be to improve the legal framework in order to reduce the risk of hold-up, and thus increase the willingness to interact with unknown partners. In the case of a social context where agents have few opportunities to meet new partners and start new collaborations, encouraging agents to increase the number of their collaborations might be enough to force the creation of linkages with new partners. When the goal is to increase both knowledge creation and collaboration, this can be achieved by focusing on improving collaboration.

Despite the general belief that collaborations boost performance and knowledge creation, it is important to remember that collaborations are not costless. The cost of establishing and maintaining a collaboration can sometimes outweigh the benefits of collaborating. In a dynamic setting, it might switch the relation between knowledge creation and networking, so that the most productive agents may attract too many collaborators and become unproductive.

This paper has some limitations. First, the simulation model suggests different lines of action for different underlying processes. Selecting the right process is essential for choosing the right action to implement. In order to address this, future research would benefit from empirically validating the model. This could be done with the Werker and Brenner (2004) method, or by calibrating the parameter values with part of a panel of patent data and then checking how this calibrated model fits the rest of the data from the panel. This empirical validation may allow comparison of different knowledge creation processes. Furthermore, the model could be adapted to incorporate parameter changes through time in order to implement policy actions. Then, the policy actions suggested for the different scenarios could be tested through simulations, in a secure and costless way. However, measuring the parameters needed for every application with sufficient precision would be difficult, even more so since each parameter value is likely to be heterogeneous for each actor. This either increases the amount of data needed to calibrate the model to fit a specific problem or reduce its predictive power, if parameter value at the population level are left constant. How to solve these issues could be subject of future work.

Nonetheless, the main goal of this paper was to call into attention the importance of the feedback between network formation and knowledge creation. This has been accomplished by showing that changes in one of the two functions (the knowledge or the network) can affect the results of the other (the network or the knowledge). Taking into account the endogeneity of the process is crucial when studying knowledge creation in networks. Scholars aiming to improve the empirical literature should tackle this issue and not take the network as fixed or given to study its (one-way) effect on the performance of its agents.

**Acknowledgements**

We thank J. Olmos-Peñuela, F. Rentocchini and K. Frenken, as well as the participants to the EMAEE 2013 Conference, Eu-SPRI Early Career Researcher Conference 2013, VPDE-BRICK Workshop 2013 and DRUID Academy Conference 2014, especially Z. Babutsidze and B. Sanditov for their useful comments. The suggestions of three anonymous reviewers developed and improved the manuscript. The authors acknowledge funding by the Generalitat Valenciana (Project GV/2012/018). E.M. Tur’s Ph.D research was supported by a CSIC JAE-PreDoc fellowship co-financed by the ESF.

**References**


