



JENA ECONOMIC RESEARCH PAPERS



2013 – 022

Emergence of Innovation Networks from R&D Cooperation with Endogenous Absorptive Capacity

by

**Ivan Savin
Abiodun Egbetokun**

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

© by the author.

Emergence of Innovation Networks from R&D Cooperation with Endogenous Absorptive Capacity*

Ivan Savin^{†‡} and Abiodun Egbetokun[†]

Abstract

This paper extends the existing literature on strategic R&D alliances by presenting a model of innovation networks with endogenous absorptive capacity. The networks emerge as a result of bilateral cooperation over time between firms occupying different locations in the knowledge space. Social capital is ignored, and firms ally purely on the basis of knowledge considerations. Partner selection is driven largely by absorptive capacity which is itself influenced by cognitive distance and investment allocation between inventive and absorptive R&D. Cognitive distance between firms changes as a function of the intensity of cooperation and innovation. Within different knowledge regimes, we examine the structure of networks that emerge and how firms perform within such networks. Our model replicates some stylised empirical results on network structure and the contingent effects of network position on innovative performance. We find networks that exhibit small world properties which are generally robust to changes in the knowledge regime. Second, subject to the extent of knowledge spillovers, certain network strategies such as occupying brokerage positions or maximising accessibility to potential partners pay off. Third and most importantly, absorptive capacity plays an important role in network evolution: firms with different network strategies indeed differ in the build-up of absorptive capacity.

Keywords: *absorptive capacity; agent-based modeling; cognitive distance; dynamics; innovation; knowledge spillovers; networks*

JEL Codes: *C61, C63, D83, D85, L14, O33*

*Financial support from the German Science Foundation (DFG RTG 1411) is gratefully acknowledged. Thanks are due to very helpful comments and suggestions from Uwe Cantner, Robin Cowan, Luigi Orsenigo, Peter Winker and participants of the RSAEM seminar in Brussels. The usual disclaimers apply.

[†]DFG Research Training Program 'The Economics of Innovative Change', Friedrich Schiller University Jena and the Max Planck Institute of Economics

[‡]Corresponding author. E-mail: Ivan.Savin@uni-jena.de. Bachstrasse 18k Room 216, D-07743 Jena, Germany. Tel.: +49-3641-943275, Fax: +49-3641-943202

1 Introduction

The main aim of this paper is to examine the influence of absorptive capacity on the structure and performance effects of innovation networks that emerge from bilateral R&D collaboration. Innovation is, by nature, a highly uncertain process which involves recombination of knowledge (Dosi, 1988). Knowledge recombination is facilitated when it diffuses effectively. Networks are often perceived as an infrastructure for knowledge diffusion (Cowan, 2005). These networks usually arise out of voluntary cooperation either among firms or between firms and other economic agents. A standard result in studies of strategic alliances and networking is that firms benefit through cooperation. The benefits show up in terms of accessing complementary resources, division of labor, risk sharing, reduction of uncertainty and improved chances of innovative success through multiple search efforts (Pittaway et al., 2005; Powell, 1998). Two alternative explanations for the emergence of and benefits derived from networks can be found in literature.

From the social capital perspective, network position is considered to be very crucial, such that more central firms tend to outperform peripheral ones both in terms of successful alliances and innovativeness (Gulati, 1995; Powell et al., 1999). For this reason, alliances are thought to be largely motivated by social capital considerations (Coleman, 1988; Ahuja, 2000; Burt, 2004; Gilsing et al., 2008) and most of the empirically observed properties of innovation networks are explained by the fact that firms are seeking to increase their number of economically valuable connections. In particular, some authors argue that it is strategically important to combine both relational and structural embeddedness in networks (Moran, 2005; Rowley et al., 2000). In this regard, small world structures are thought to be particularly beneficial for innovation and the diffusion of knowledge (Schilling and Phelps, 2007).¹

From a knowledge perspective, alliances can be heavily motivated by technological fit, that is the extent to which partners potentially learn from each other (Cowan, 2005). On the one hand, what is missing from a firm's stock of knowledge and competences influences its decision to cooperate and its choice of partners. In this sense, multiple partnerships may not be necessary and a firm may stop its partnership search once it locates a technologically fit partner. On the other hand, a firm's suitability is assessed by potential partners on the basis of what is present in its knowledge base. Thus, firms' internal knowledge deficiencies and externally available complementarities play a significant role in the emergence of learning and innovation networks. In this regard, small world structures are important because they preserve the quantity and diversity of knowledge (Baum et al., 2003), thereby affecting the learning and innovation potential of alliances.

The foregoing considerations are central to the models of Cowan et al. (2007) and Baum et al. (2010) in which alliance formation is driven by its probability to succeed in terms of knowledge generation and innovation, as well as the proximity of the potential partner. The studies demonstrate that networks with small world properties and other empirically founded network characteristics such as repeated alliances and transitivity can be observed even when alliances are formed only on the basis of knowledge considerations. However, these studies treat absorptive capacity as an exogenous parameter which is similar for all firms in an industry. This simplification is motivated by the fact that it allows to focus on the nature of the innovation process and its effects on emergent network

¹With respect to a network of firms, a small world is a network in which distinct regions with dense interconnectedness (or cliques) of firms are linked by relationships (or clique spanning ties) that act as information conduits between them (Watts and Strogatz, 1998).

properties. However, this imposes the neglect of an important source of heterogeneity, that is, differences in firms' learning rates.

In this paper we also approach alliance formation from the knowledge perspective but with endogenous absorptive capacity. Firms form alliances for the purpose of knowledge sharing. Partner selection is entirely network-independent, implying the exclusion of network-based motives. The effectiveness of alliances is influenced by two factors: cognitive distance between partners and their investment allocation. Both factors determine absorptive capacity which is required to effectively deploy externally generated knowledge (Cohen and Levinthal, 1989). While the former has an inverted 'U'-shaped relationship with the learning and innovation potential of the alliance (Wuyts et al., 2005), the latter presents a trade-off in the optimal distribution of total R&D investments between the creation of own knowledge and the improvement of absorptive capacity. The higher the investment in original knowledge creation, the more attractive a firm appears as an innovation partner. At the same time, the lower the investments in the build-up of absorptive capacity, the more difficult it is to exploit external knowledge.

Taken together, the foregoing hold important implications for cooperation and partner selection. On the one hand, a firm needs to carefully balance between R&D investments made to generate inventions and to develop absorptive capacity. On the other hand, the firm needs to select partners that are neither too close to it in the knowledge space (to facilitate novelty) nor too far away (to facilitate understandability). An additional consideration is the distinction that can be made between voluntary spillovers which exist in the context of cooperation and involuntary ones that exist elsewhere. In particular, voluntary spillovers are reciprocal, thereby constituting both a benefit and a potential risk. In this regard, firms will pay attention not only to the amount of knowledge they can get from their potential partner but also to the partner's absorptive capacity. These elements were combined in our earlier model of absorptive capacity and inter-firm cooperation (Egbetokun and Savin, 2012). In that static model, the cognitive distance between cooperating partners was set exogenously. This simplification permitted a focus on the relationship between performance and cooperation strategy for a representative firm.

Building on research on alliance formation, we focus on dynamic aspects of cooperation wherein the cognitive overlap between partners increases with intensity of cooperation, either in terms of duration or frequency (Cantner and Meder, 2007). For instance, Wuyts et al. (2005) argue that the cognitive distance between cooperating firms is a negative function of their frequency of interaction. In other words, their knowledge bases become more similar as they cooperate more frequently. A similar argument was made by Mowery et al. (1998) for the duration of cooperation. *Ex post*, the knowledge overlap may be greater than its pre-cooperation level because of the mutual knowledge exchange over time. Cooperating firms may then become so close that the knowledge potential of their partnership becomes too low to permit recombinant novelty (Antonelli et al., 2010, p. 53). At this point, investments in absorptive capacity become less productive as far as the particular partnership is concerned. This may motivate the firms to invest more in own knowledge generation (inventive R&D) while reducing the absorptive R&D. In addition, when this stage is reached, the firms might reconsider their cooperation decisions and the partnership may dissolve.² Heterogeneity between the firms increases again

²This situation arises even between asymmetric firms, that is, a technological leader and a follower, because, as long as they operate within the same technological trajectory, the leading firm has little reciprocal incentive to continue the relationship except that of opportunism or expropriation, which constitute disincentives for the follower. As noted by Nooteboom (1999, p. 802), "A problem in collabo-

if they subsequently generate new knowledge either alone or in cooperation with other partners. The two firms may then be incentivised to re-establish partnership.

This paper analyses a dynamic model³ in which networks emerge as a result of bilateral cooperation between firms occupying different locations in the knowledge space. The hypothesis that we examine is straightforward: could the empirically observed properties of networks be reproduced by abstracting from social capital and focusing exclusively on knowledge considerations with endogenous absorptive capacity?⁴

An important contribution of this paper is that, in contrast to Cowan et al. (2007) and Baum et al. (2010), we account for differences in firms' absorptive capacity and how this affects their dynamics in the knowledge space. Absorptive capacity is endogenously defined by two factors: (i) a firm's distance both to a potential partner and to aggregate external knowledge, and (ii) its decision on the investment trade-off between inventive and absorptive R&D potentially compensating for a larger distance to a partner. This way, absorptive capacity combines elements of searching for, valuing, identifying and assimilating new knowledge (Zahra and George, 2002).

Furthermore, distinguishing between voluntary and involuntary spillovers allows us to examine our hypothesis with respect to different knowledge regimes. The intuition here is that at different times in the history of an industry, different extents of voluntary and involuntary spillovers will be observed due to varying levels of inter-firm cooperation. For instance, industries tend to cluster in the early stages when knowledge is more tacit and its diffusion require face-to-face interactions (von Hippel, 1989; Audretsch and Feldman, 1996). At such times, a higher proportion of inter-firm collaborations characterised by high levels of voluntary spillovers is likely to be observed. However, in later stages, the effects of localised spillovers have been reported to diminish significantly (Potter and Watts, 2011) partly due to congestion, obsolescence of local knowledge and, in particular, a high amount of codifiable intra-industry spillovers.

Our results do indeed replicate important empirical facts and generate some new insight. We observe networks with small world properties at all levels of spillovers that we examine. The effects of network structure on firm performance varies with changes in the knowledge regime. Aggregate profit in the networks increases with increasing involuntary spillovers but an inverted 'U'-shaped relationship is observed with increasing voluntary spillovers. Moreover, when involuntary spillovers are small, networks with high average path length - implying low accessibility and inefficient information flow within the network - are especially detrimental for innovation. High betweenness - that is, occupying some kind of brokerage positions - turns out to be a very profitable network strategy at low levels of involuntary spillovers. A particularly striking result is that firms which employ different network strategies do indeed differ in their absorptive capacities.

The rest of the paper is organized as follows. Section 2 presents the basic model. In Section 3 we address the parameter calibration issues of the present ABM. Section 4 illustrates the obtained results, while Section 5 contains some concluding remarks.

ration, especially in innovation, is that under some conditions there may be opportunities and incentives for free ridership, or for one party extracting more gain than others, or even expropriating their gain".

³With many firms, analytical solution of the dynamic model becomes intractable so we employ an agent-based simulation model (ABM). ABMs have gained an increasing interest in different fields of economic research having an advantage in (i) a more realistic representation of agents' behavior than in a standard representative agent model and (ii) possibility of an extensive and fast simulation analysis for different parameter settings due to the ongoing advances in computational performance.

⁴We do not imply a contrast between absorptive capacity and social capital; rather, we examine the networks generated when cooperation is motivated by knowledge gains rather than social capital.

2 The Model

In the model, a fixed population of firms (N) seeks to generate new knowledge over a certain number of periods within a defined knowledge space. Similar to Baum et al. (2010), a simple representation of firms in a two-dimensional metric space⁵ capturing cognitive distance is used. While firms' locations in the underlying space have no particular meaning, they "translate directly into a network of strategic alliances" (Baum et al., 2010, p. 2097), because the distances affect the learning ability and, hence, partnership formation.

In each period, innovations can be generated from new knowledge created within that period.⁶ Each firm maximizes its potential to innovate either alone or in cooperation with another firm. The decision to cooperate is influenced by absorptive capacity not only of the firm itself but also of its potential partner. The absorptive capacities, in their turn, depend on the extent to which the two firms' knowledge endowments both resemble and complement each other (cognitive distance). Bilateral partnerships among the firms yield an aggregate network. We are particularly interested in three issues:

- i. the kinds of aggregate network structures that emerge: here we examine whether the networks generated by our model display small world properties like many real life networks (Cowan and Jonard, 2004; Verspagen and Duysters, 2004).
- ii. the effects of different knowledge regimes on aggregate network structures and performance: here we analyze how the network structures respond to varying degrees of voluntary and involuntary spillovers.
- iii. the relationship between firms' network position and their innovation performance: the focus here is on individual firms and the manner in which the structural characteristics of the network relate to their performance.

Four important assumptions are made in the model. First, partnership formation is only a short-term profit-maximising decision. Second, each firm selects only one partner and conducts one R&D project in each period. Partnerships are reconsidered in every period so that previously formed alliances may be discontinued. Third, reciprocity in partnerships is only relevant in terms of shared knowledge; partners' trust and reliability are ignored. Last, firms are well informed about the knowledge base but are uncertain about the investment decisions of other firms.

2.1 R&D investments

For each firm i , we distinguish between investments in directly in R&D (rdi_i^t) for the creation of own knowledge (which is a share, ρ , of total research budget, RD_i), and investments for exploring the environment for new complementary knowledge (aci_i^t):

$$RD_i = rdi_i^t + aci_i^t = \rho_i^t RD_i + (1 - \rho_i^t) RD_i, \quad (1)$$

Allocation of these investments is influenced by the potential quantity and complexity of external knowledge, either within a partnership or beyond it, both of which, in the

⁵A two-dimensional representation is the smallest suitable form allowing for transitivity relations in the metric space and provides a clear graphical representation of network formation and evolution.

⁶Thus, although the new knowledge may be combined with already existing knowledge to innovate, not every new recombination of knowledge is considered to be an innovation.

context of cooperation, are proportional to cognitive distance. In making its own R&D investment decisions, each firm takes into account the investment decision of its potential partner. It does this by forming an expectation, considering the investment decision of the partner to be equal to the average from the last three investment allocations made by the partner in a cooperation setting.⁷ This introduces uncertainty into the model as the expectation does not exactly coincide with the actual investment decision of the potential partner, which is itself based on an expectation. Thus,

$$E^i(\rho_j^t) \neq \rho_j^t = f(E^j(\rho_i^t)), \quad \text{where} \quad E^i(\rho_j^t) = \left(\sum_{\iota=1}^3 \rho_j^{t-\iota} \right) / 3. \quad (2)$$

2.2 Knowledge generation

Firm i 's stock of knowledge in period t (k_i^t) increases as a result of its direct investment in R&D (rdi_i) and involuntary spillovers (ek^t) from other (both cooperating and non-cooperating) firms. In an alliance, firm i can also appropriate voluntary spillovers (δ_c) from its strategic partner j . The extent to which the firm benefits from the two types of spillovers depends on its absorptive capacities: $ac_{i,j}^t$ and $ac_{i,ek}^t$. Thus,

$$k_i^t = (rdi_i^t)^\xi + ac_{i,j}^t (\delta_c rdi_j^t) + ac_{i,ek}^t (ek_i^t), \quad (3)$$

where $\xi \in (0, 1)$ is a parameter which defines the rate of returns to inventive R&D.

External knowledge, ek_i , is set as the total inventive R&D investment of companies in the knowledge space which firm i can potentially understand (in total, let us say, equal to H_i), rescaled by the parameter of involuntary spillovers,⁸ $\delta_n \in (0, 1)$:

$$ek_i = \delta_n \sum_{i \in H_i} rdi_h. \quad (4)$$

The understandability restriction ensures that firm i can utilize the involuntary spillovers from firms located in the knowledge space and sets a certain 'radius' around the firm, within which external knowledge (also from the side of strategic cooperation) can be considered. Hence, firms having a more central position in the knowledge space have an access to potentially more spillovers than those being in the periphery.

In our model we focus on the situation, where $1 > \delta_c > \delta_n > 0$ with $\delta_c + \delta_n$ reflecting total spillovers available to a cooperating firm.

⁷The expectation formation plays an important role in our ABM and several alternatives have been analyzed. In the simplest case, the expectation was set equal to the decision made by the partner in the previous period, regardless of its decision to cooperate. Naturally, this approach was most imprecise. Alternatively, a generalization as an average over several periods was taken, but still providing a big mismatch between the expectation and the actual decision. Finally, an average over the last three investment decisions made by the partner within a cooperation is taken. This approach provides a good approximation of the actual investments made by the partner in all the scenarios considered: correlation between expectation and actual investments is always between 0.7 and 0.95 (see Appendix B).

⁸This fraction reflects the portion of knowledge not appropriated by companies and is determined by the appropriability conditions which include the patent system in a particular industry and the efficacy of secrecy or other forms of protection of firm j 's internal knowledge

2.3 Cognitive distance

Firms select partners from whom they are likely to benefit the most. Such partners typically possess complementary assets and other endowments that might not be easily accessible elsewhere, or they may be easily understood by the focal firm. Consequently, the cognitive distance of a firm i from a potential partner j (d_{ij}) is not necessarily equal to that from other external knowledge ek (d_{iek}). The former is modeled as the Euclidean distance between the stock of knowledge of the two partners (ν_i and ν_j), which are independently and randomly attributed to the firms in the interval $[0, 1]$:

$$d_{ij} = \sqrt{(\nu_{i1} - \nu_{j1})^2 + (\nu_{i2} - \nu_{j2})^2}. \quad (5)$$

Cognitive distance from external knowledge is represented as the average distance to the firms in the knowledge space firm i is able to understand:

$$d_{iek} = \left(\sum_{i \in H_i} d_{ih} \right) / H_i, \quad (6)$$

so that the maximum distance to the external knowledge does not exceed the maximum distance to a single potential partner in this space.

2.4 Absorptive capacity

As noted earlier, a firm develops its absorptive capacity by investing a share $(1 - \rho_i)$ of its total R&D budget for that purpose. Since a firm aims to maximise the knowledge it absorbs given its current level of absorptive capacity, one can think of absorptive capacity as the actual amount of knowledge absorbed by the firm divided by the total amount of knowledge available to absorb:

$$ac_{i,j} = \frac{\alpha\beta_1 d_{ij} + \alpha\beta_1 d_{ij} ac_i^\psi - \alpha\beta_2 d_{ij}^2}{\frac{1}{4\alpha\beta_2} \left[\alpha\beta_1 (1 + ac_i^\psi) \right]^2} \in [0, 1]. \quad (7)$$

The parameter $\psi \in (0, 1)$ reflects the decreasing marginal returns to absorptive R&D investments. The function (7) is derived and discussed in Egbertokun and Savin (2012). In short, the function reflects two main empirical findings. First, the cognitive distance (d_{ij}) between cooperation partners has an inverted ‘U’-shaped relation with the knowledge the partners obtain Lin et al. (2012); Gilsing et al. (2008); Nooteboom et al. (2007); Wuyts et al. (2005); Mowery et al. (1998). Second, investments in absorptive capacity allow firms to reach further in space in selecting cooperation partners (de Jong and Freel, 2010). This causes an increase in optimal distance between pairs of cooperating firms. Thus, an understandability–novelty trade-off exists such that effective learning by interaction is better accomplished by limiting cognitive overlap while securing cognitive proximity.

Note at this point that for $d_{ij} = 0$, the respective absorptive capacity in (7) equals zero as there is no new knowledge to absorb. In contrast, even with no investments in absorptive R&D, ac_{ij} may be positive for some minor cognitive distances. The latter assumption can be justified by the fact that firms working in similar fields have some level of mutual understanding even without explicit investments in absorbing R&D.

Similarly, $ac_{i,ek}$ has the same functional form as (7) with the only difference that d_{iek} replaces d_{ij} . Thus, for the same level of absorptive R&D investments, the absorptive capacities directed on each of the two sources of spillovers can be different.

Note also that the investment in absorptive capacity increases the distance (d_i) over which the firm can absorb external knowledge (the radius mentioned in Section 2.2).⁹

2.5 Profit generation

In our model innovation involves recombination of heterogeneous resources by a firm either working alone or in partnership with another firm. Therefore, the magnitude of a successful innovation is defined by the amount of knowledge firm i can appropriate (Π_i):

$$\Pi_i = \begin{cases} (k_i^t)^{\text{generated in cooperation}} / (1 + \kappa ac_{j,i}^t \delta_c r d_i^t) & \text{if } i \text{ has a partner,} \\ (k_i^t)^{\text{generated alone}} & \text{if } i \text{ has no partner.} \end{cases} \quad (8)$$

This variable is an output of the appropriated knowledge from a firm's continuous R&D effort and, as the case may be, partnerships. The appropriated knowledge is applied in the recombination process to generate an incremental innovation through which the firm maintains its competitiveness. Consequently, one can think of the firm's profit as being proportional to its knowledge input into the innovation process.¹⁰ In this sense, (8) is henceforth referred to as profit and used as a main indicator of firms' performance.

In a partnership, Π_i decreases proportionally with the amount of knowledge spillovers ($ac_{j,i}^t \delta_c r d_i^t$) that the partner can absorb (which is a constituent part of k_j that reduces the appropriability of k_i). This is in contrast to Cohen and Levinthal (1989) where Π_i is reduced proportional to the knowledge generated by the partner (k_j).¹¹ This 'cost of partnership' affects the choice of an R&D partner. Although this cost provides a disincentive to cooperate, one should remember that the cost is contingent upon the partner's absorptive capacity which itself is imperfect preserving the cooperation incentives.¹²

To avoid the problem of increasing Π_i for $ac_{j,i}^t \delta_c r d_i^t < 1$, we introduce a 'natural' leak-out that is fixed and equal to 1. Furthermore, since the cost of partnership affects Π_i multiplicatively, while absorbed knowledge from external sources comes into Π_i only additively, we downsize $ac_{j,i}^t \delta_c r d_i^t$ by factor κ to ensure that cooperation in our model brings more benefits than losses.¹³

To sum up, the function (8) is meant to introduce the trade-off between cooperative and non-cooperative strategies in our model: it provides a larger pool of knowledge spillovers for a cooperating firm i to benefit from, but also penalizes it by the spillovers the partner j can absorb. This function is later used as a main objective function of the firms reflecting the short-term profit-maximising objectives stated earlier. More information on the calibration of the function is provided in Section 3.

⁹Specifically, this radius is set to be not larger than $\frac{\beta_1}{\beta_2}(1 + (RD_i)^\psi)$ and can be derived from (7).

¹⁰Obviously, we limit ourselves here to successful innovation. Nevertheless, it can be argued that the learning effects from failed innovation efforts will be proportional to the knowledge input.

¹¹Recall that in Cohen and Levinthal (1989), $\frac{\partial \Pi_i}{\partial k_i} > 0$, $\frac{\partial \Pi_i}{\partial k_j} < 0$ and $\frac{\partial \Pi_i}{\partial k_i \partial k_j} < 0$, which is also fulfilled in (8) with the distinction that Π_i reduces proportional to the firm's i spillovers j can absorb.

¹²See Egbertokun and Savin (2012); Hammerschmidt (2009, p. 426); Cohen and Levinthal (1989, p. 575-6) for a more elaborate discussion of this.

¹³Simulations show that setting $\kappa = 1$ leads firms to prefer the standalone mode in most of the situations, while correlation between firm performance and the number of partnerships becomes negative.

2.6 Innovation and learning

There are two sources of dynamics in the knowledge space: learning and innovation. First, firms move closer together according to their learning capacity with respect to their partner (ac_{ij}), the extent to which they disclose own knowledge (δ_c) in a cooperation agreement and the extent to which the partner conducts inventive R&D in the period (ρ_j). Logically, the higher the three variables are, the faster the two firms learn from each other, reducing their cognitive distance.¹⁴ Technically this ‘convergence’ in knowledge space is implemented similar to Baum et al. (2010) with the distinction that the learning is potentially much faster¹⁵ and is endogenously driven by firms’ investment allocation:

$$\begin{aligned}\nu_{i1}^{t+1} &= (\delta_c ac_{ij}^t \rho_j^t) \nu_{j1}^t + (1 - \delta_c ac_{ij}^t \rho_j^t) \nu_{i1}^t \\ \nu_{i2}^{t+1} &= (\delta_c ac_{ij}^t \rho_j^t) \nu_{j2}^t + (1 - \delta_c ac_{ij}^t \rho_j^t) \nu_{i2}^t.\end{aligned}\tag{9}$$

In interpreting (9) we believe that the learning capacity (ac_{ij}^t) itself should not be set to some small value¹⁶ simply by arguing that firms do not learn from each other that quickly. In contrast, we believe that what actually matters is how much information the two firms disclose to each other in a cooperation agreement and how much original (inventive) R&D each of them conducts in that period.

The second driver of dynamics in cognitive distance is innovative activity. When a firm produces a radical innovation,¹⁷ it dislocates firms in its surrounding according to its innovative success ($\Theta \Pi_i$),¹⁸ where Θ is a binary outcome with one standing for a successful innovation. We set Θ equal to one for only one randomly drawn firm per period so that on the one hand, we do not have too many innovations and dislocations at each period, but also avoid the situation where all firms converge to one specific location in the knowledge space. Important also is that we do not distinguish between probabilities to successfully innovate alone or in a cooperation in order not to introduce even more heterogeneity in an already complex model. The difference in terms of innovation between the cooperative and non-cooperative strategies is that in a partnership, firms have the potential to generate innovations of larger magnitude¹⁹ due to voluntary spillovers.

We also restrict the effect of dislocation to the firms which are located close enough to the firm producing an innovation at time t , i.e. again the ‘radiuses’ of surrounding firms

¹⁴In this respect, there is a clear parallel with the ‘cost of partnership’ operationalized in (8) except that instead of R&D budget, the firms’ coordinates in the knowledge space are taken.

¹⁵A standard in the literature is setting the learning capacity (ac_{ij}^t) equal 0.01.

¹⁶Otherwise, why would firms put so much effort in increasing their absorptive capacity (Fabrizio, 2009), if it should always be equal to a value like 1%?

¹⁷Here we distinguish between two types of innovation. Whereas incremental innovation in (8) is considered to be a consistent indicator of firms’ performance in generating new knowledge, radical innovation (taking place infrequently and randomly) here is an important source of firms’ dynamics in the knowledge space. This distinction is useful for technical convenience. Nevertheless, it is logical because, while incremental innovations tend to encourage the status quo, radical innovations cause paradigmatic shifts with entire industries emerging or transforming (see Section 2 in Koberg et al., 2003, for a more detailed discussion). Thus, one can argue that the effects of incremental innovation reflect more in the innovator’s profit without significantly shifting other firms’ position in the knowledge space, whereas a radical innovation forces other firms to adjust their position within the knowledge space.

¹⁸Again, the magnitude of innovation is endogenized in this model and not drawn from any exogenous distribution as it is done in some recent studies (Baum et al., 2010).

¹⁹Here one can think of magnitude in terms of quality improvement or cost reduction that is significant enough to generate reactions from other firms in the knowledge space.

are used checking that the distance between the innovating (i) and a surrounding (h) firm is not too big for the latter one to comprehend the innovation. A technical aspect of this limitation allows one not to ‘shake’ the entire population of firms but only a certain number of them located in the specific region of the knowledge space. Furthermore, it also matters how large is the distance of the affected firms to the innovating ones (d_{ih}) – the smaller it is, the larger the dislocation:

$$\begin{aligned}\nu_{h1}^{t+1} &= \nu_{h1}^t + \epsilon_1 \ln(\Pi_{i,t})/d_{ih}^t \\ \nu_{h2}^{t+1} &= \nu_{h2}^t + \epsilon_2 \ln(\Pi_{i,t})/d_{ih}^t,\end{aligned}\tag{10}$$

with $\epsilon_1, \epsilon_2 \in [-\frac{1}{200}; \frac{1}{200}]$ rescaling the entire dislocation effect on each firm below $d_{ih}^{max}/2 = \sqrt{2}/4$. The dislocation effect in (10) holds irrespective of whether or not a firm h is involved in a partnership in the specific period. Thus, moving along the knowledge space according to their learning and innovativeness, firms may essentially form some clusters within which they exchange knowledge.

2.7 Optimal investment allocation and partnership formation

For certain levels of distances d_{ij} and d_{iek} that determine understandability and novelty, firm i is incentivized to invest in absorptive R&D to maximize the amount of external knowledge absorbed. The trade-off that the firm faces is how to optimally distribute its total R&D investment between the creation of own knowledge and the improvement of absorptive capacity. This necessitates a comparison of the marginal returns to each type of investment with respect to the profit gained:

$$\frac{\partial \Pi_i}{\partial aci_i} = \frac{\partial \Pi_i}{\partial rdi_i}\tag{11}$$

In Egbetokun and Savin (2012) the derivatives from (11) have been analyzed and considered for a representative firm in a cooperative and non-cooperative setting.²⁰

Ultimately, the decision to cooperate (or not) is a profit-maximizing one which depends on the potential profit generated when working alone in comparison with profit generated by cooperating with any of the possible partners:

$$\max \left(\Pi_i^{\text{generated alone}}; \Pi_i^{\text{with any of the possible partners}} \right).\tag{12}$$

Therefore, for each firm in each period the investment trade-off (ρ_i) is obtained for all potential partners, taking into account the expectation about those partners’ R&D investments (2) and not their actual investments. After that, the amount of knowledge k_i to be generated by each company either alone or in partnership is estimated. Based on this the most lucrative partner for each company is selected by maximizing profit in (8).

²⁰However, already in that setting those functions are highly complex and non-linear with multiple local minima depending on the particular set of parameter values applied. To solve such a non-trivial optimization problem and find an optimal ρ_i for each firm both in a cooperating and non-cooperating setting a heuristic optimization technique, Differential Evolution (Storn and Price, 1997) is applied. Among its advantages are comparative simplicity in tuning (Blueschke et al., 2013), a good approximation of a global optimum satisfying (11) for different sets of our model’s parameters and fast convergence. For more details see Appendix 2 in Egbetokun and Savin (2012).

Although the most lucrative partner for each firm is identified, partnership formation is a non-trivial task. The reason is that the incentives of a firm i to build a partnership with a firm j are asymmetric: while distance between the partners is the same, the investment decision is individual for each company. Hence, there is no ‘Nash stable network’.²¹ Hence, few alternatives on forming partnerships are considered:

- *Unilateral partnership formation*: in each period in a random order (to ensure that none of the firms has an advantage over others throughout the entire simulation process) firms sequentially identify their most fitting partner (based on the estimation of ρ_i and Π_i for each potential partner). Once the partner is found, partnership is formed (i.e. the chosen firm simply adjusts its ρ to the given partner) and the two firms are excluded from the search process in the respective period. If for any given firm the standalone mode is more lucrative, it generates new knowledge alone and is excluded from further search in this period. The main advantage of the method is its simplicity and lowest computational time required. However, reciprocity is not required; therefore, a firm i can exploit its partner j in a given period. Hence, this method is expected to result in the largest discrepancy in firms’ performance and can be considered as a certain benchmark to compare with.
- *Reciprocal partnership formation*: at first, both ρ_i and Π_i of each firm for each potential partner are estimated. After all firms preferring a standalone mode are excluded, again in a randomized order in each period, firm i ‘makes a proposition’ to firm j as its most lucrative cooperation partner, which will be successful if and only if it belongs to the ‘top’ 5% of the companies with whom firm j would cooperate. Then the two firms are excluded from further search. This approach is clearly more computationally intensive and sets the strictest limitation on reciprocity.
- *A ‘popularity contest’*: starts exactly as with reciprocal partnership by computing ρ_i and Π_i . Afterwards, for each firm a ‘rating’ of each firm’s popularity among other firms is calculated. This rating is measured as the number of times a particular firm is listed as the first, second or last ‘fittest’ partner for each firm in the population (including itself), while the weights are the inverses of the positions (i.e. weight of being first is equal one, of being tenth is equal one tenth):

$$Rating_i = \sum_{j=1}^N weight_i^{by\ firm\ j}, \text{ where } weight_i^{by\ firm\ j} = \frac{1}{Order\ of\ i\ by\ firm\ j}.$$

After that firms choose themselves a partner sequentially according to their rating: the most popular choose first. Although some reciprocity is present in the model, it is not ensured in every partnership. Computational cost of this approach is comparable to the one before, but it has some additional merits. First, the order in which partnership choice is made is not random. Second, the popularity ranking introduces some form of hierarchy (heterogeneity) among the firms. This presents a modest representation of competition: since firms ranked higher choose first, firms with the desire to form alliances may tend to adopt strategies which improve their ranking, thereby competing for alliances.

²¹‘A stable network is one in which for each agent (or pair of agents) there is a payoff maximizing decision about which link to form’ (Cowan et al., 2007, p. 1052)

It is worth noting that the procedures just described take place every period. This implies that the process of partner selection occurs in every period, and a partnership formed in one period may be terminated in the next one depending on its profitability.

3 Numerical Experiment

Given the complexity of the model at hand, we need to set some parameters described as fixed ones leaving only few to vary in the following extensive numerical simulation.²² The dynamics arising from this modeling should provide us with complex network information which we discuss in Section 4. The entire simulation runs over two hundred periods ($T = 200$) repeated ten times. In each case, the first hundred periods are removed from further consideration to avoid any effects arising from initial random allocation of parameters.

At the beginning of each simulation restart, a population of $N = 100$ firms²³ are randomly distributed in the knowledge space $[0, 1] \times [0, 1]$. They are also given a certain fixed R&D budget uniformly drawn from the interval $[7.5, 12.5]$.²⁴ To form expectations about other firms' investment decision, we randomly allocate values between 0 and 1 to all firms in the first three periods. Throughout the next remaining 197 periods, the firms should form alliances (or stay alone) according to one of the three matching alternatives described in Section 2.7 always solving two trade-offs: what is the preferred distance to the partner to cooperate with and how much to spend on absorbing voluntary and involuntary spillovers. After all alliances are set, firms generate knowledge in an alliance or staying alone and subsequently move in the knowledge space. While cooperating firms learn from each other and move towards each other in the knowledge space according to (9), roughly once in each period one of the firms innovates and dislocates the surrounding ones according to (10).

We set $\alpha = \sqrt{2}/50$, $\beta_1 = \sqrt{2}/40$, $\beta_2 = 1$ and $\psi = \xi = 0.5$, thus allowing i) $ac_{i,j}$ to have the inverted 'U'-shaped form in d_{ij} : first increasing and then decreasing;²⁵ ii) aci_i to have a positive but marginally decreasing impact on absorptive capacity; iii) rdi_i to have a positive but marginally decreasing impact on k_i and iv) setting the radius (within which firm can find a partner, absorb involuntary spillovers or be affected by another firm successful in generating an innovation) equal ≈ 0.15 on average (depending on the exact R&D budget the firms have). We also set κ as the rescaling parameter for the costs of partnership equaling 0.1.²⁶ We describe both the firms and the arising networks over ten

²²The entire code is written in MATLAB except of few parts on network statistics made in C++ by means of Microsoft Visual Studio boosting the efficiency of simulation. In this respect we want to thank Lev Muchnik for his instructive tutorial and codes on social network analysis at <http://www.levmuchnik.net/Content/Networks/ComplexNetworksPackage.html>. The code can be obtained from the authors on request.

²³This number is neither too large nor too small for simulation purposes. Besides, it is comparable to the population of firms applied in previous related studies

²⁴In this we attempt to introduce some moderate heterogeneity between firms in our model. Clearly, an even more ambitious idea would be to set the R&D budgets in the future periods contingent on the firm's performance in the previous period, but this we leave for further research.

²⁵Furthermore, as stated in Section 2.4, without any investments in absorptive capacity (aci_i) $ac_{i,j}$ remains positive for $d_{ij} \in (0, 0.0353]$ and reaches its maximum level at $d_{ij} = 0.0177$ (i.e. in the middle of the interval). Thus, at some very moderate level of cognitive distance, assimilation of external spillovers can be efficient without any explicit investments in absorptive capacity.

²⁶We tried different values of κ both above and below 0.1 and it turns out that this value produces more meaningful results in terms of number of partnerships per period (10-90% of firms in the population coop-

histories each lasting hundred periods (after discarding the first hundred periods in each case). In particular, we concentrate on:

- i. the network statistics, taking the whole population as one network.
- ii. firms' ego-network statistics and their correlation to the firms' performance.

In this model four parameters that determine absorptive capacity also drive the network structure and its effect on firm performance. First, cognitive distance (d_{ij}) between cooperating firms influences the learning and innovation potential of an alliance. This distance changes according to the firms' learning and innovation. As a result, previously discontinued alliances may be re-formed.²⁷ Second, R&D investments (RD) are the major source of absorptive capacity. Allocating this investment between invention and the development of absorptive capacity is an important strategic decision that every firm makes in response to the behaviour of their potential partners. Third, appropriability conditions within a partnership (δ_c) determine both the pool of knowledge of its partner that each firm can benefit from and the magnitude of spillovers the partner can absorb from the firm. The size of δ_c is an important factor both in partner selection and learning speed. Fourth, magnitude of involuntary spillovers (δ_n) generated by surrounding firms defines the pool of external knowledge.

While the first two factors are exogenously given only in the initial stage and endogenized afterwards, the latter two are held as exogenous. In particular, we consider $\delta_c \in [0.1; 0.9]$ and $\delta_n \in [0.03; 0.1]$.²⁸ This setup allows us to examine the evolution and performance effects of the other parameters - and, by extension, of the entire network - with respect to changes in the knowledge regime within the knowledge space.

4 Results

In order to understand results contained in this section, certain knowledge of network analysis is necessary. To help readers to follow the discussion we clarify the network measures (both the unweighted ones and their weighted generalizations) in Appendix A.

4.1 Network measures and characteristic networks

To give an idea of what networks emerge from our model, we provide some descriptive information on networks formed with the three different matching rules applied, using

erate). Giving a calculation example from our model according to (3) and (8), to give an idea about what forces are actually at work, consider $\Pi^t = \left((rd_i^t)^\xi + ac_{i,j}^t (\delta_c rd_j^t) + ac_{i,ek}^t (ek_i^t) \right) / (1 + \kappa ac_{j,i}^t \delta_c rd_i^t)$ equaling to 4.62 (for $d_{ij} = 0.0375$, $\rho_i = 0.5 = E^i(\rho_j)$, $\rho_j = 0.5$, $\delta_c = 0.9$, $RD_i = 9.19$, $ac_{ij} = 0.89$, $RD_j = 10.15$, $ac_{ji} = 0.88$, $ek = 1.07$, $d_{iek} = 0.11$ and $ac_{iek} = 0.07$). While the knowledge generated (before the cost of partnership reduces it in Π_i) equals 6.30, Π_i is rescaled by factor 1.36 (and not 4.63 for $\kappa = 1$). This makes the profit (knowledge generated and appropriated) still higher than investing only in inventive R&D ($\Pi = 3.03$) or absorbing only involuntary spillovers with the same investing decision ($\Pi = 2.22$). As it is also clear, setting $\kappa = 0.01$ would make the cost of partnership negligible. Smaller deviations, e.g., $\kappa \in [0.05; 0.15]$ do not change the further results dramatically.

²⁷As noted by Rosenkopf and Schilling (2007), alliance duration often ranges from one to five years. Thus, over a sufficiently long time horizon and in a sufficiently large population of firms, alliances will repeat often enough that firms will re-select partners with whom they had cooperated in the past.

²⁸Here we restrict the amount of involuntary spillovers from below to ensure the emergence of a network in the 'popularity contest' matching rule.

some intermediate values of δ_c and δ_n (in particular setting them equal to 0.5 and 0.05, respectively). These results are given both on a single simulation run (first and last panel in Figure 1) as well as for ten restarts.

The network using unilateral partnership formation (presented to the left in Figure 1) has the largest density and contains all except one firm in its largest component. The density of the network generated in the popularity contest scenario is much smaller, indicating relatively few alliances taking place in each period. Even more, in this scenario the total number of unconnected components, several of which contain more than one firm, is above 20. However, the network appears to be highly centralised (as seen in the network representation on the bottom right of Figure 1). This feature, to some extent, is an effect of the fact that network formation is based on reputation such that much fewer firms form alliances and repeatedly connect with each other.²⁹ An interesting measure in the context of this study is the *fraction of reciprocal links* (i.e. the share of double-connected pairs - alliances formed on the ‘initiative’ from both sides). This is one of the few characteristics where we account for link direction and examine differences among the three matching rules. Clearly as a result of the barter-like setup, mutual linkages, or what we can term double coincidence of wants (Cowan and Jonard, 2004), are significantly higher in the network generated from the reciprocal partnership rule. The other two networks are not particularly different in this respect; they both have relatively fewer mutual linkages.

To see how dense the resulting networks are (in other words, to what extent partners of partners are connected to each other), we measure the *clustering coefficient*. Since our networks are cumulated over 100 periods with many alliances taking place several times, a weighted generalization of this coefficient is more suitable. From the values we see that the local structure of the network with reciprocal partnership is the most dense. Another important network measure is the *mean path length*³⁰ which proxies the efficiency of information flow within a network - the smaller the value, the more efficient a network is in terms of information diffusion. The network with unilateral partnership has the shortest path, while the popularity contest network has the longest. The network with reciprocal partnership lies between these two extremes but has a mean path length much closer to the former one. Combining these two measures, we can assess the *small world* property of each network structure. The most common quantitative measure of this is the *small world ratio* obtained from dividing clustering coefficient by mean path length. Typically, values greater than one have been used in previous studies to indicate that a network is small worldly (Davis et al., 2003). Judging from the weighted version of this measure, the networks generated from our model, with the exception of the one with popularity contest, seem to represent small worlds.

By definition, a network with small world properties will be much more clustered than a corresponding random network but the average path length between its nodes will be comparable to that of the random network (Watts and Strogatz, 1998). Thus, to make sure that the resulting networks truly represent small worlds, one has to compare the relevant measures with an equivalent random network (benchmark). Specifically, the benchmark is one with not only the same number of nodes and links (which would have

²⁹This feature reflects in most of the other measures that we examine subsequently. Thus, the discussions in the rest of the paper relies more on the results from the other two matching rules.

³⁰This should not be confused with distance. Mean path length is the average number of nodes separating two distinct nodes (e.g., i and j) in the network, while the distance refers to the (Euclidian) distance between two nodes in the knowledge space: d_{ij} .

been sufficient for an unweighted one-period network), but also the same distribution of weights. For that we generate random networks having the same number of nodes and links as those observed in our ABM. We then randomly assign weights to the links in the benchmark from the distribution of weights of our observed networks. This procedure was replicated 100 times³¹ and average values for clustering coefficient, mean path length and small world ratio were obtained. After that, we divide measures derived from the ABM by the values from benchmark networks, denoting the resulting network characteristics as *rescaled*.³²

	Unilateral partnership		Reciprocal matching		Popularity contest		
Single example	# of components	2	6		39		
	Average degree	26.20	21.76		15.42		
	Density	0.262	0.209		0.100		
	Fraction of reciprocal links	54.06	69.77		55.83		
	Weighted clustering coefficient	1.01	1.56		0.63		
	Rescaled clustering coefficient	2.16	3.28		2.32		
	Weighted mean path length	0.55	0.75		1.59		
	Rescaled mean path length	1.16	1.66		1.17		
	Weighted Small World ratio	1.85	2.08		0.40		
	Rescaled Small World ratio	1.86	2.04		2.00		
Results over 10 restarts	# of components	2	(0.7)	4	(1.05)	24.67	(14.01)
	Average degree	23.04	(2.84)	19.92	(1.27)	10.99	(2.25)
	Density	0.242	(0.018)	0.199	(0.019)	0.101	(0.021)
	Fraction of reciprocal links (in %)	54.80	(2.93)	65.92	(2.42)	48.43	(4.65)
	Weighted clustering coefficient	1.06	(0.10)	1.35	(0.12)	0.79	(0.22)
	Rescaled clustering coefficient	2.50	(0.33)	3.28	(0.28)	2.39	(1.29)
	Weighted mean path length	0.64	(0.09)	0.71	(0.05)	1.88	(0.33)
	Rescaled mean path length	1.21	(0.09)	1.40	(0.11)	1.40	(0.36)
	Weighted Small World ratio	1.69	(0.28)	1.89	(0.17)	0.43	(0.12)
	Rescaled Small World ratio	2.06	(0.23)	2.36	(0.24)	1.92	(1.54)

Note: In the lower panel, average values (standard deviations in parentheses) are reported.

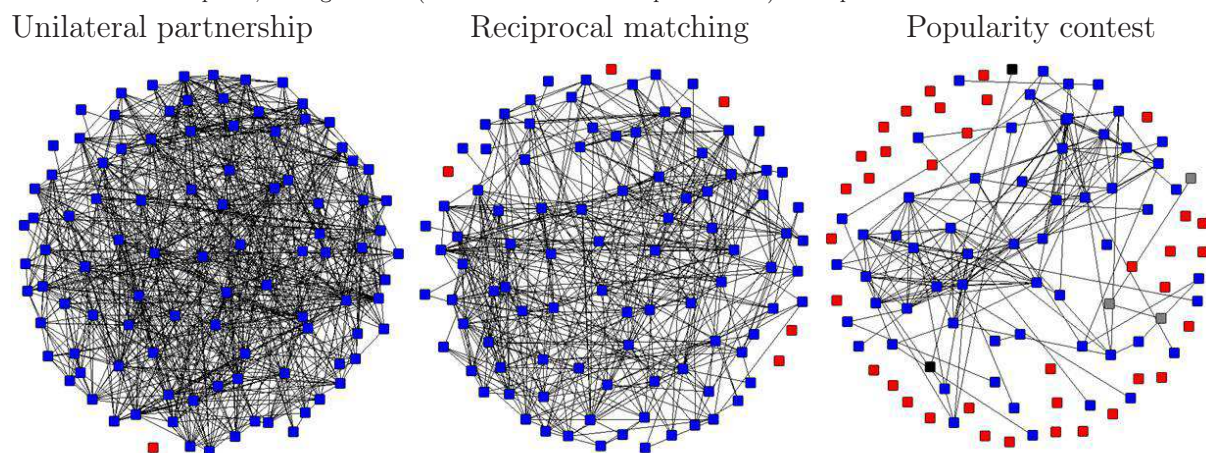


Figure 1: Characteristic networks with different alliance formation rules

³¹As before, this figure is in keeping with earlier studies.

³²In doing so and given the types of networks we obtain in our model, we concentrated our comparison on the largest connected components only. This is because in the popularity contest scenario, there are potentially several disconnected components with a relatively moderate total number of links. This gives rise to very sparsely connected random networks and thus very low clustering coefficients. Whereas in the first two matching rules (unilateral and reciprocal) this approach leads to only marginal differences in the rescaled values, the difference for the popularity contest approach is dramatic. The implication of this approach is that the network with popularity contest is not necessarily small worldly as a whole but its largest connected component is.

The characteristic networks generated in our ABM have rescaled small world ratios about twice as high as the one from a corresponding random network. This is primarily because of the much denser local structures generated in our model and not the shorter paths (whereas random networks have systematically shorter paths than constructed networks, their clustering coefficients are two to three times smaller).³³ In general, this implies that even in the absence of network-based structural and strategic motives, networks that emerge from bilateral partnership based on knowledge considerations with endogenous absorptive capacity demonstrate small world properties.

The 3-D graphs on the right in Figure 2, also obtained from a single simulation run, display the ‘matrices of cooperation’ (also referred to as weighted adjacency matrices) for the different matching rules. This refers to the number of partnerships aggregated over the 100 periods, distinguishing whether a particular alliance was ‘initiated’ by partner i : in such a case the partnership is attributed to i ’s row - left axis. We see a much larger number of alliances in the unilateral partnership formation resulting also in a larger density. In contrast, in the reciprocal partnership scenario, the alliances are more symmetrically dispersed in the matrix resulting in a higher reciprocal rate of partnerships. To illustrate their difference, we compare these two networks in terms of the amount of profits generated by firms over the time interval under consideration. It is seen that although on average firms in the unilateral setting generate more (474 versus 448), this comes at the cost of a larger disproportion between the firms. The coefficient of variation³⁴ in the unilateral setting is 5.95 (versus 5.63 in the reciprocal partnership setting).

The popularity contest scenario results in the smallest mean and coefficient of variation: 377 and 5.3, respectively. As earlier pointed out, in this scenario there is some level of competition for alliances among these firms which detrimentally affects aggregate performance (though particular firms perform well). Here only a moderate number of firms form an alliance more than ten times, and many of those firms cooperate only with each other. Those firms performing well, as expected, are the ones allying most (correlation between the aggregated profits and the total number of partnerships $\approx 82\%$). The dramatic difference in the number of partnerships cannot be explained by the R&D budget allocation (correlation is merely $\approx 32\%$), but rather the network location of those firms (correlation with weighted betweenness centrality is $\approx 74\%$).

4.2 Different knowledge regimes

To provide a more systematic insight about the network characteristics and the contingent effects of firms’ position (both in the network and the knowledge space), we analyze results by varying the magnitude of voluntary and involuntary spillovers. In particular, we fix $\delta_n = 0.05$ varying the voluntary spillovers and then similarly alter involuntary ones by fixing $\delta_c = 0.2$. The results are produced for ten restarts³⁵ and reported in medians together with 5-95% quantiles to account for the variance in those results. The logic is as follows. Alliance formation in our model is solely for knowledge sharing, and partner selection is driven not by social capital but mainly by absorptive capacity. Firms therefore tend to select partners from whom they expect to gain the largest amount of knowledge

³³Note, however, that the result for the popularity contest approach is not that stable, and in certain restarts the small world ratio of the generated network can be close to that of the benchmark.

³⁴Simply the standard deviation divided by the mean.

³⁵A single restart of the current ABM for a given parameter setting requires from 130 to 250s using Matlab 7.11 and Pentium IV 3.3 GHz (depending on the matching rule applied).

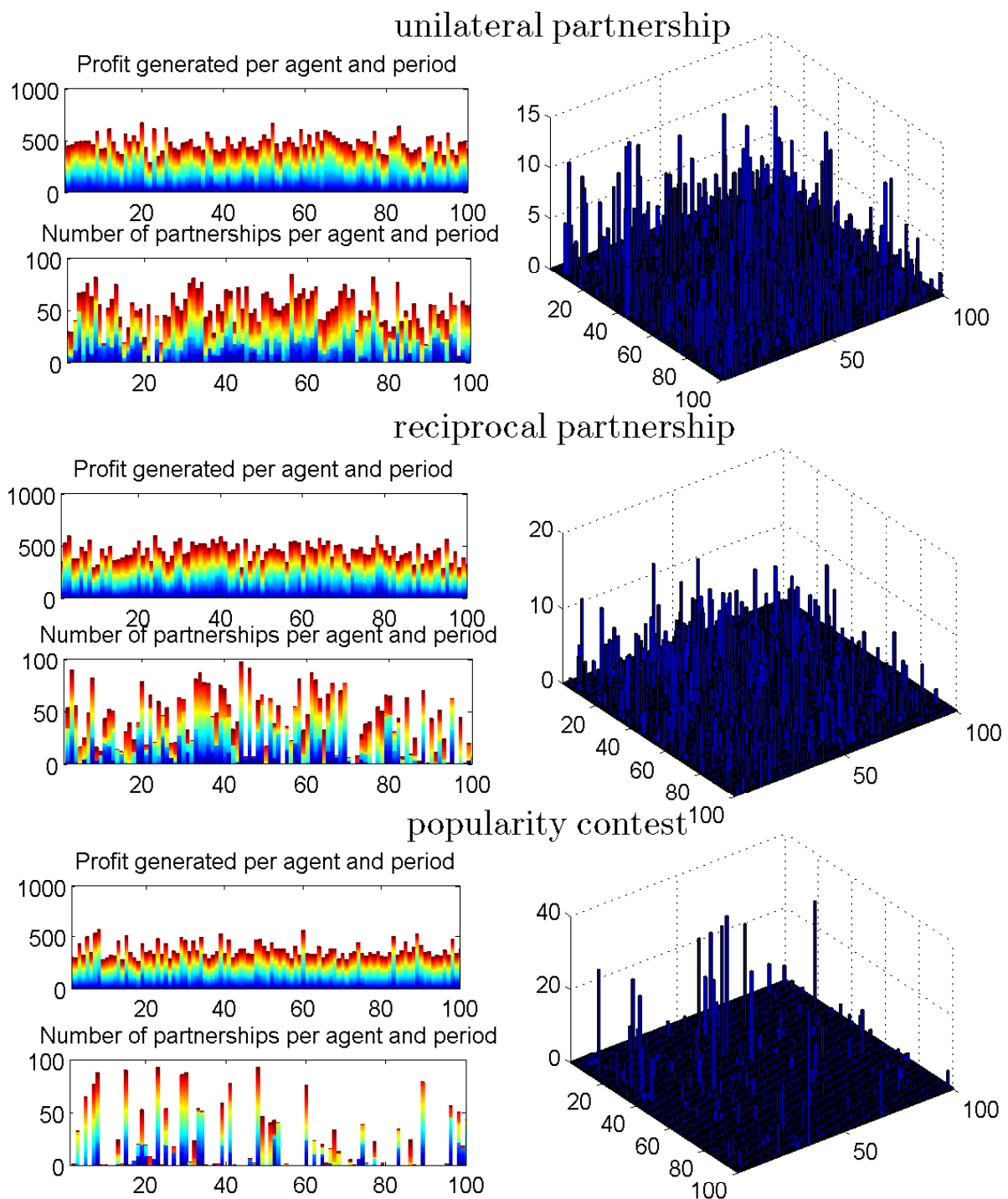


Figure 2: Descriptive information on the network with different formation rules

Notes for Figure 2: On the left upper plots firms are given on the X-axis, while aggregated profits over 100 periods are on the Y-axis. Colors in the stacked bars indicate attribution of the profits to a particular period. On the lower left plots of the respective matching rule again firms are on the X-axis, while the total number of alliances firm have participated in are on the Y-axes. Similarly, the color on the stacked bars represent a particular period a partnership was taking place in. On the right plots the weighted adjacency matrices are given: the X-axis (left side of the 3-D plot) indicates a firm successfully 'offering' an alliance, while Y-axis - the company accepting it. Z-axis illustrates the total number of alliances between the firms.

at the lowest cost (Bala and Goyal, 2000). Thus, changes in the quantity of knowledge available through partnership (voluntary spillovers) should affect partnership formation. Moreover, the knowledge gained from an alliance is combined with externally available knowledge (involuntary spillovers) as inputs into the innovation process. Consequently, changes in the quantity of involuntary spillovers should also affect alliance formation. Taken together, changes in those spillovers reflect in aggregate network structure and performance as well as the contingent effects of firms' network position on their innovativeness. Such changes are supposed to occur at different times in the history of an industry. Typically, when an industry is young, knowledge is more tacit and requires cooperation to gain access. Thus, a higher intensity of voluntary spillovers can be observed. In contrast, in a mature industrial setting, knowledge is more codified and thus firms do not necessarily need to cooperate to gain access to external knowledge (higher involuntary spillovers).

4.2.1 Network structures in different knowledge regimes

Figures 3-5 show the effects of involuntary spillovers (δ_n) on the aggregate network structure and performance. The effects of voluntary spillovers are shown in Figures 6-8. In all cases, we report first the dynamics in the small world ratios (subplot (a)). In subplot (b) we illustrate the dynamics in the *length of cooperation* (or duration of alliances). Here we count all the cooperations between any two firms actually taking place, measure how long they were lasting without discontinuation - irrespective of whether an alliance was formed on the 'initiative' of one or the other firm - and take the average. The same plot shows the trend in the total number of alliances in each period. Subplot (c) shows the trend in aggregated profits generated within the network. Recall that the absorptive capacity within the context of cooperation (ac_{ij}) is different from that which is directed on external knowledge (ac_{iek}), due to different distances between the firm and these knowledge sources. How these two capacities respond to changes in the quantities of knowledge is shown in subplot (d). In subplot (e) the trend in *% of reestablished cooperations* is shown. This is defined as the share of cooperations during a given period which did not exist in the immediate preceding period, but existed during one of the three periods before that, i.e. i and j form an alliance in period t , but not in period $t - 1$, while they also had an alliance at least once between $t - 2$ and $t - 4$.³⁶ Finally, in all of Figures 3-8, subplot (f) shows the relative benefits (i.e. amount of knowledge generated from inventive R&D as well as voluntary and involuntary spillovers) and costs of cooperation. The costs are calculated not as the denominator in equation (8), but as the difference between $(k_i^t)^{\text{generated in cooperation}}$ and Π_i , i.e. by how much knowledge generated in cooperation has been reduced due to outgoing spillovers to the partner.

First we discuss how different δ_n affect aggregate network characteristics and performance (Figures 3-5). In general, the networks retain their small world properties for different amount of involuntary spillovers. Clearly, the networks with unilateral and reciprocal matching are more small worldly, mostly because changes in δ_n do not affect the average mean path length and clustering in these networks (subplots (a) and (b), Figures 15-17 in Appendix B).

³⁶The choice of exactly three periods here is meant to be a trade-off: to consider on the one hand a potentially larger time horizon (since only one period is too short to account for), and on the other hand to avoid too much of double counting (since the larger the time horizon, the higher the chance that within it a cooperation might have been discontinued and re-formed several times).

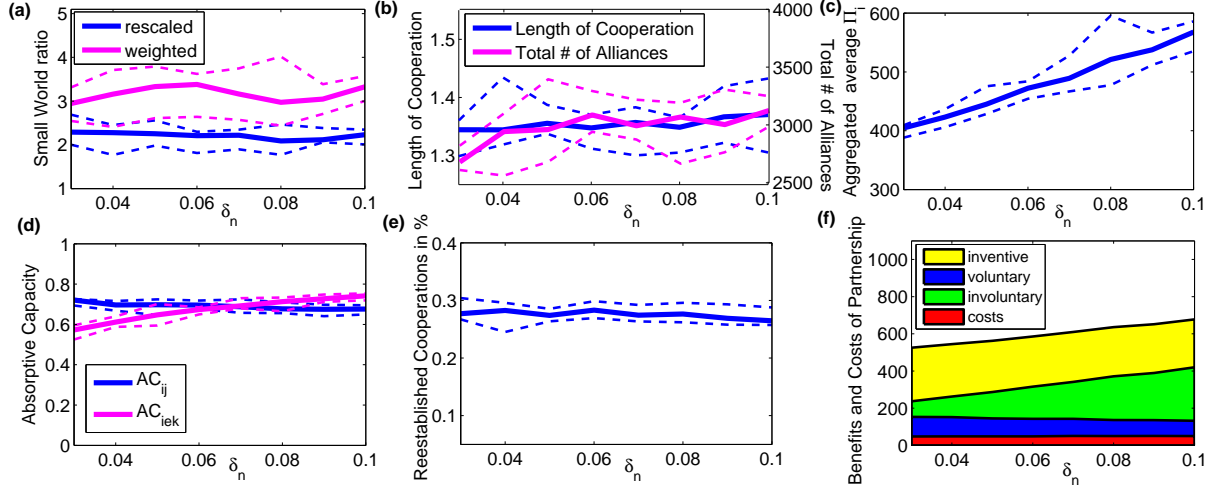


Figure 3: Network characteristics for different δ_n and unilateral partnership

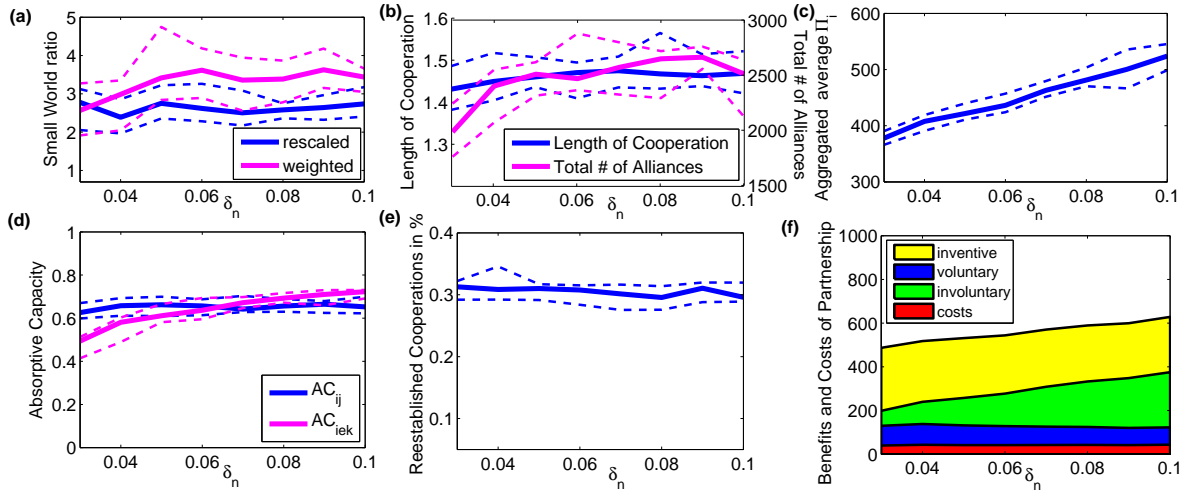


Figure 4: Network characteristics for different δ_n and reciprocal partnership

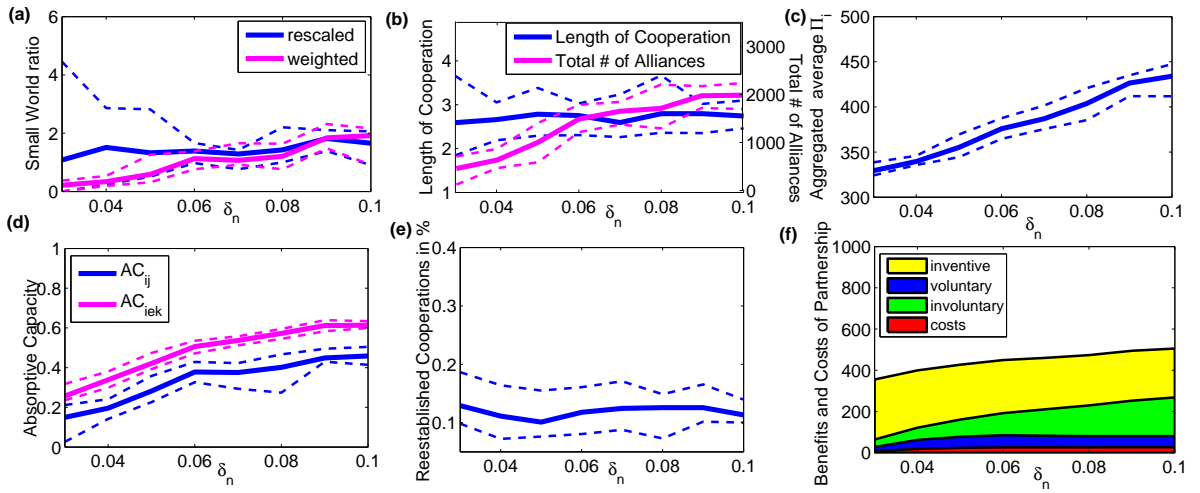


Figure 5: Network characteristics for different δ_n and popularity contest

Specifically, an increase in the magnitude of involuntary spillovers from firms located

close enough (in particular, within the radius discussed in Section 2) has no substantial impact on the rescaled small world ratio (being ≈ 2). In comparison, the network with popularity contest is less small worldly. Only $\delta_n \approx 0.1$ (i.e. 10% of knowledge from surrounding firms ‘spilling over’ at no cost to a given firm) allows to reach a small world ratio comparable to the other two networks. For smaller δ_n the rescaled ratio sometimes also reaches this value, but this result is not robust.

In general, alliance durations are relatively short and not very responsive to changes in δ_n . The values range from ≈ 1.35 periods in the unilateral matching to ≈ 1.45 periods in the reciprocal matching and ≈ 2.5 periods in the popularity contest. These figures are generally consistent with empirically observed alliance duration averages of between 1 and 5 years (Rosenkopf and Schilling, 2007). The average duration of cooperation maps directly onto the total number of alliances. The generally short durations imply increasing number of partnerships as δ_n increases. Shorter durations in the unilateral matching context correspond to higher alliance rates (and, by extension, a highly dense and clustered network) while comparatively longer durations lead to relatively lower alliance rates in the popularity contest scenario. A related measure in this respect is reestablished cooperations (subplot (e)). Now, reversely, the popularity contest network has the lowest share of about 10%. Since partners tend to stay longer together in this context, the possibility for alliances between them to re-occur is comparatively lower than in the other two networks. Moreover, there are fewer alliances in the popularity contest network. In the other two matching rules, repeated partnerships are close to 30%, with the reciprocal partnership having a marginally higher share.

It is worth noting that in the popularity contest network alliance formation increases consistently across the entire range of δ_n and more than doubles. The marked increase may be explained by the fact that for low δ_n in this scenario cooperations are ‘initiated’ by the most ‘popular’ firms in the population which benefit from their central position most and invest funds in absorbing both voluntary and involuntary spillovers from more distinct (in the knowledge space) partners, while majority of other agents prefer the standalone scenario by absorbing only involuntary spillovers. This is confirmed by the look on the average distance between partners in cooperation (subplot (d) in Figures 15-17 in Appendix B): for $\delta_n = 0.03$ this distance is twice as high in the popularity contest network than in the other two. However, as involuntary spillovers rise, firms located in the ‘periphery’ of the knowledge space get an incentive to invest more in absorbing new knowledge and engaging in R&D cooperation, and the average distance falls.³⁷

Increasing alliance rates, especially at low levels of δ_n , can also be explained by the fact that when involuntary spillovers are small, firms rely more on cooperation partners as sources of additional knowledge for innovation. But as δ_n rises, firms start to pay more attention to involuntary spillovers while maintaining their access to knowledge from alliances. Consequently, the number of alliances in any given period tends to level out at higher levels of δ_n . This may also be the reason for the slightly lower clustering coefficients and mean path lengths observed for higher δ_n in Figures 15-17 in Appendix B.

This trend is even better observed by looking at the dynamics in learning capacities (subplot (d)). The absorptive capacity directed towards external knowledge (ac_{iek}) is seen to rise consistently in all the matching scenarios while the absorptive capacity

³⁷In the other scenarios, this distance increases in δ_n as there are more spillovers to absorb and firms more equally engage in cooperation. In the unilateral matching this distance is slightly higher as firms ‘offering an alliance’ may *ex parte* impose a partnership if they find it profitable not requiring any reciprocity.

within alliances (ac_{ij}) slightly falls in the unilateral and reciprocal matching networks and increases only in the popularity contest approach as more and more firms start to co-operate.³⁸ These results imply that when involuntary spillovers increase, firms' capacity to appropriate them also increase. This is quite logical since, in the scenarios observed here, the level of voluntary spillovers is fixed; firms may, therefore, shift their learning attention towards involuntary spillovers which are consistently increasing.

Nevertheless, aggregate profits increase (subplot (c)) mostly due to increasing amounts of involuntary spillovers assimilated. Profits due to inventive R&D reduce as firms tend to invest less in invention (subplot (e) of Figures 15-17 in Appendix B). The contribution of voluntary spillovers assimilated in the R&D profit (darkest area in subplot (f)) as well as the costs related to R&D cooperation remain relatively the same (slightly fall in the first two scenarios and increase in the last one). This is to be expected since δ_c is fixed and firms do not necessarily become better at appropriating spillovers from cooperation.

Now we consider how changes in voluntary spillovers (δ_c) affect aggregate network structure and performance (Figures 6-8). We find that higher voluntary spillovers - which increase the speed of learning and convergence in the knowledge space according to (9) - cause the networks to become somewhat less small worldly: the weighted small world ratios reduce by nearly half.³⁹ However, the rescaled measures which compare our networks to the random benchmarks do not change much as voluntary spillovers increase (subplot(a)) though in the popularity contest network, this result is not robust. The small world properties of these networks appear to be more sensitive to changes in voluntary spillovers than to involuntary ones.

When voluntary spillovers are small, alliance durations are generally much longer (also in comparison to the regime of involuntary spillovers) but fall rapidly as the spillovers increase. This is primarily due to increased pace of learning which makes continuous cooperation with the same partner less profitable over time. Thus, alliances are more often discontinued and firms either find other partners or innovate on their own. This reflects in the reducing total number of alliances since, in reality, it takes time for firms to adjust their investments appropriately and to find other suitable partners. The share of reestablished cooperations remains rather stable: it only slightly falls in the unilateral matching, while a marginal increase can be observed in the popularity contest (subplot (e)). This could be because, as earlier demonstrated with the characteristic networks (Figure 2), cooperation intensity is lower in the popularity contest scenario, and many cooperating firms repeatedly ally only with each other. Thus, there is an increased likelihood that previously discontinued alliances are re-formed in this scenario.

Another pronounced effect of the learning rate is on the dynamics in absorptive capacity (subplot (d)). Learning from involuntary spillovers increases when δ_c is high and remains stable above the learning that occurs from voluntary spillovers. Combined with what was observed earlier in the case of δ_n , this suggests that firms tend to put more effort into absorbing spillovers when they are high. More of this learning takes place from involuntary spillovers because the costs are much lower than for voluntary ones.

³⁸Note here that estimating average absorptive capacity in cooperation AC_{ij} we ignore periods when firms do not form an alliance in order not to downsize the actual absorptive capacity and highlight its level in cooperation periods.

³⁹Figures 18-20 show also that weighted clustering reduce and mean path length increase marginally.

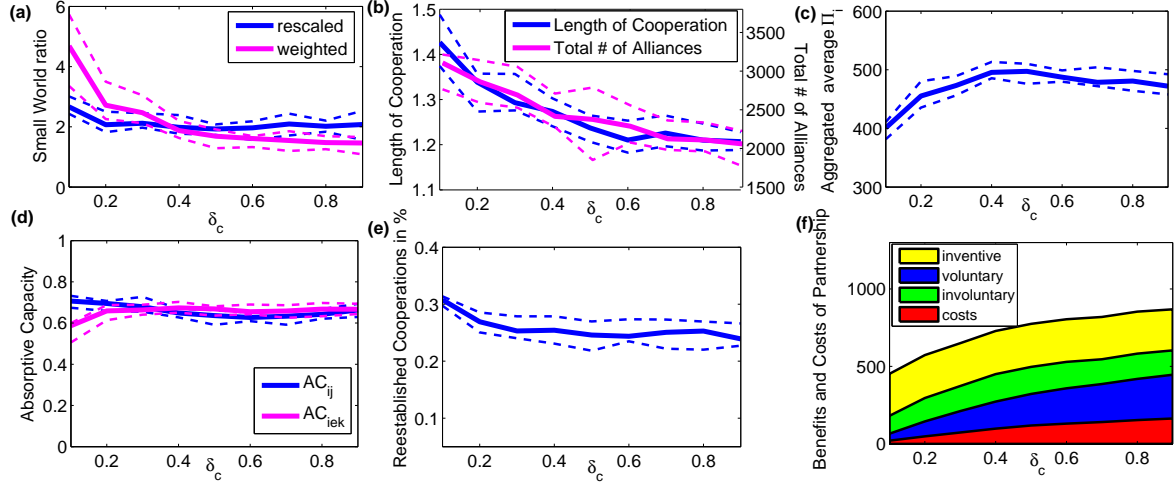


Figure 6: Network characteristics for different δ_c and unilateral partnership

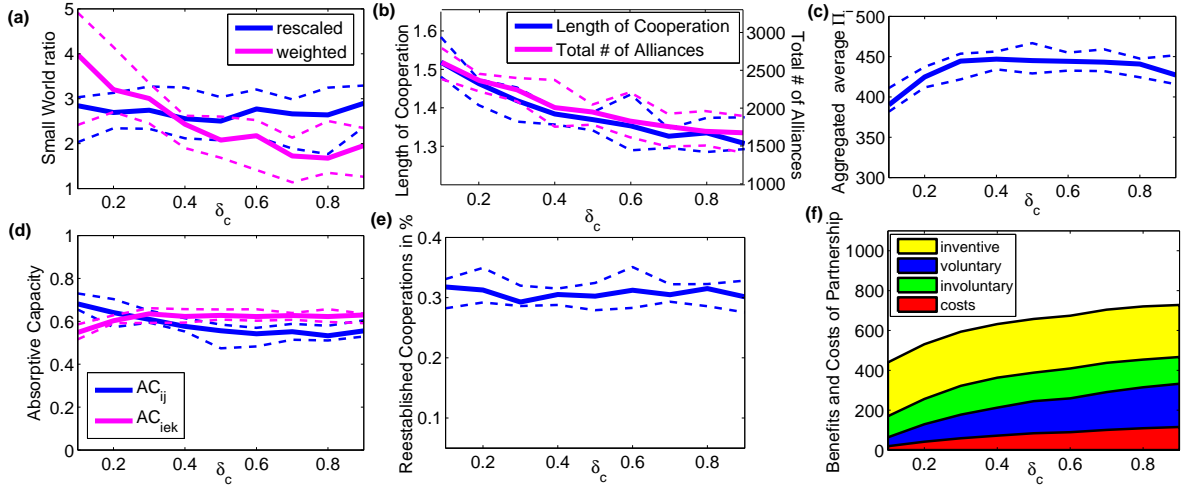


Figure 7: Network characteristics for different δ_c and reciprocal partnership

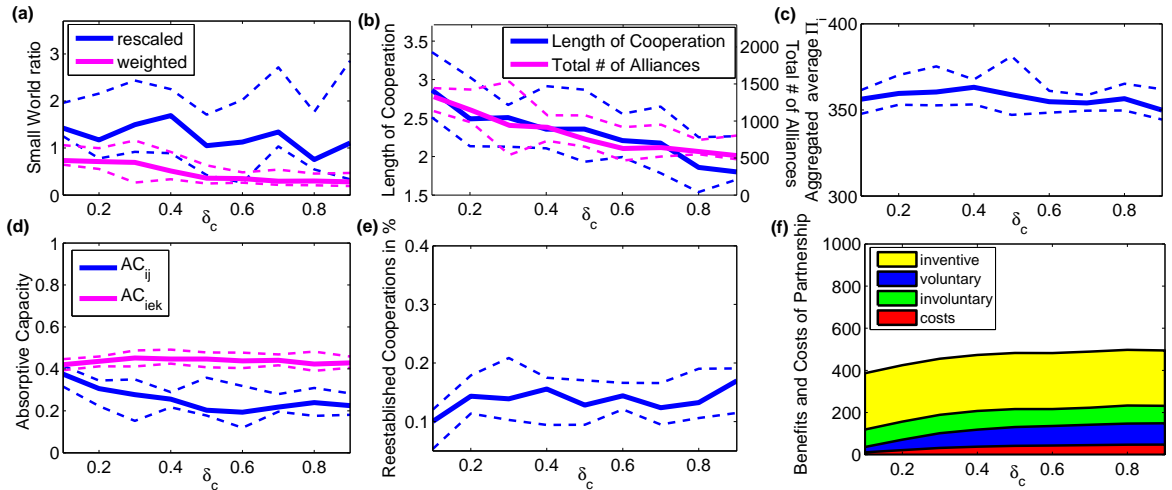


Figure 8: Network characteristics for different δ_c and popularity contest

Moreover, ac_{ij} slightly falls for some intermediate levels of δ_c and rises afterwards. This can be explained by changes in cognitive distance. As shown in subplot (d) of Fig-

ures 18-20 in Appendix B, average cognitive distance between partners first falls in δ_c since learning implies that firms move closer in the knowledge space. As δ_c reaches its middle range ($\delta_c \approx > 0.4$), average distance increases. This may be a result of increasing absorptive R&D (subplot (e) of Figures 18-20 in Appendix B), where firms reach further in the knowledge space to find cooperation partners with novel knowledge. In fact, firms first reduce investments in absorbing external knowledge (with smaller distance less investments are required) but then increase them back to roughly the same level. This causes ac_{ij} to rise. The combined dynamics here further illustrates the ambiguous relationship between cognitive distance and absorptive capacity that we analyzed in Egbetokun and Savin (2012).

It is worth noting that aggregate profit reaches its maximum at an intermediate level of δ_c in all matching scenarios (subplot (c)). This happens because firms' learning capacities allow them to benefit from the combination of shorter distances to partners and increasing investments in inventive R&D. However, at high levels of δ_c , aggregate profit drops in spite of the benefits from inventive R&D and involuntary spillovers as well as increasing assimilated voluntary spillovers observed in subplot (f). The reason for this is that the costs of cooperation rise consistently and become more dominant as cooperation becomes more intense. The inverted 'U'-shaped dynamics draws attention to the potential pitfalls of cooperation as emphasised in the empirical literature. Intense cooperation, whether in terms of repeatedness or persistence, limits the potential for recombinant novelty, thereby reducing innovative profits. Again, networks with popularity contest consistently demonstrate the worst aggregate performance which is due to lowest number of alliances (which, in turn, is a result of competition).

4.2.2 Relationship between firms' performance and network position

A widely held belief in the literature on alliances and firm networks is that the diffusion of knowledge in networks characterised by short path lengths is more efficient. Also, it is thought to be beneficial for firms to occupy influential positions - such as having high betweenness centrality which allows them to act as knowledge brokers - in networks. These results are normally explained in terms of social capital. Our model, in which networking is entirely knowledge-driven and any kind of social capital is excluded from consideration, yet shows results which are consistent with the empirical regularities. The value in this is that knowledge and technological fit, rather than just social capital, contribute to the observed performance effects of inter-firm cooperation. An important extension derived from the results here is how the relationship between network structure and innovativeness varies in response to changes in the characteristics of the knowledge space. In this section we discuss the relationship between an individual firm's performance⁴⁰ and the structure of the network. Figures 9-11 contain the results for varying levels of involuntary spillovers and Figures 12-14 for varying levels of voluntary spillovers.

In all cases, we report first the correlation between profits and betweenness centrality (subplot (a)). In subplot (b) we illustrate the correlation between absorptive capacity to a partner and betweenness centrality. The same plot shows the correlation of absorptive capacity to external knowledge and betweenness. Subplot (c) shows the correlation between profits and number of partnerships. In subplot (d) the correlation between profits and mean path length is shown. The correlations between the two different absorptive capacities and mean path length are given in subplot (e). Finally, in all of Figures 12-14,

⁴⁰Recall that performance refers to the amount of R&D profit that the firm generates in each period.

subplot (f) shows the correlation between profits and absorptive capacity.

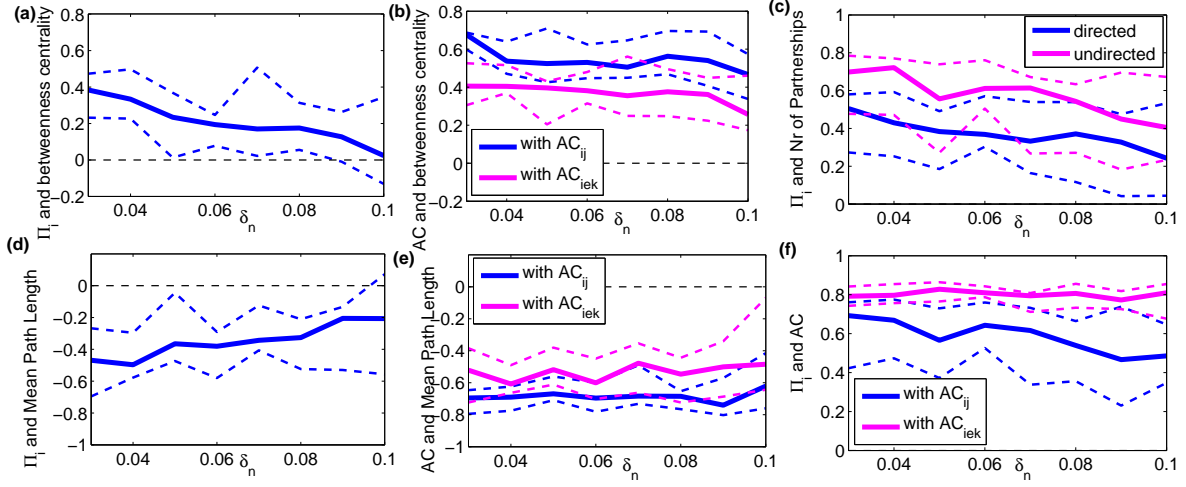


Figure 9: Correlations with firm performance for different δ_n and unilateral partnership

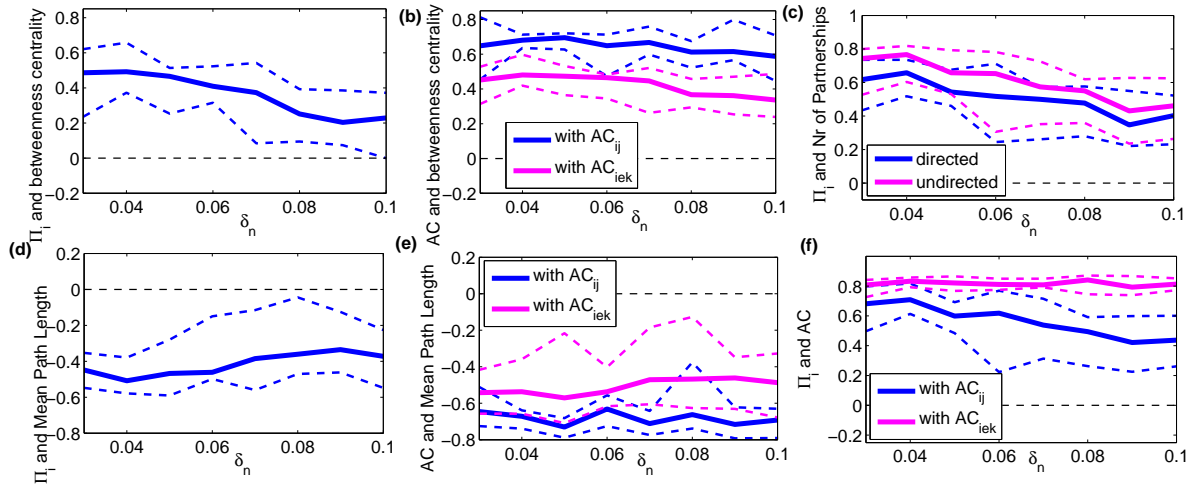


Figure 10: Correlations with firm performance for different δ_n and reciprocal partnership

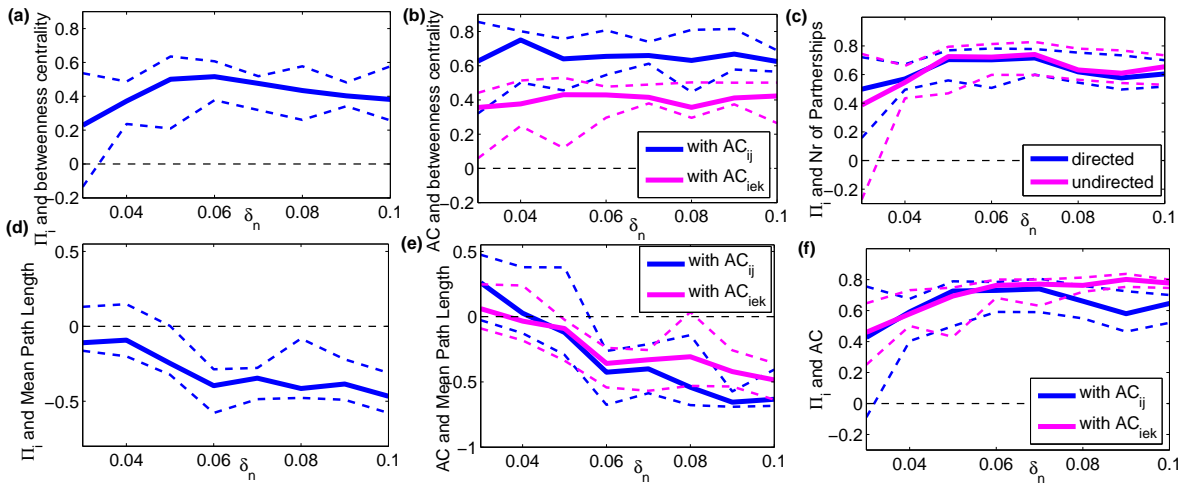


Figure 11: Correlations with firm performance for different δ_n and popularity contest

First, we consider changes in the level of involuntary spillovers (Figures 9-11). In subplot (a), betweenness is positively correlated with performance, suggesting that occupying a brokerage position pays off for the firm. At least in the unilateral and reciprocal matching networks, the highest correlations coincide with very low values of δ_n . The correlation then reduces greatly but remains positive as δ_n increases. The extreme values of δ_n correspond to different stages of an industry's life cycle. Typically, in the early stages, involuntary spillovers are low. In this context, being in a brokerage position improves access to tacit knowledge. In the later stages when knowledge is mostly codified and freely available, although being in a brokerage position is good, it becomes less relevant.

This is quite logical because, in such scenarios, networks are less clustered and brokers tend to become redundant to gain access to spillovers which are freely available. The only real constraint that each firm faces then is its capacity to absorb and not necessarily the absence of a broker in its ego-network. The correlation of profit with mean path length (subplot (d)) tells a consistent story. High values imply late arrival of knowledge and potentially lower innovation performance. However, when freely available knowledge becomes more abundant, the severity of this effect reduces significantly. Simply put, a combination of high betweenness and short path length becomes less critical for performance in mature industries wherein involuntary spillovers are generally high.

The total number of alliances is positively correlated with performance at all levels of δ_n but the strength of the correlation reduces as δ_n increases (subplot (c)). In particular, the correlation of a firm's performance with 'directed' partnerships (that is, when it initiates the partnership) is consistently lower than the 'undirected' partnerships (that is, when it either initiates or accepts a partnership). This is consistent with the empirical finding that too many partnerships can be problematic (Uzzi, 1997; Ahuja and Lampert, 2001), mostly for social capital reasons. In contrast, our result here is driven by changes in the underlying knowledge regime. When intra-industry spillovers are high, it is less efficient to maintain a large portfolio of alliances. As we have noted earlier, this situation is characteristic of the later stages of an industry when firms might be more dispersed and localised spillovers are less useful (Potter and Watts, 2011).

Also, as expected, absorptive capacity is positively correlated with performance (subplot (f)). The reducing correlation of ac_{ij} further emphasizes the view that at higher levels of involuntary spillovers, learning from cooperation becomes less important. Particularly interesting is to observe the correlations of absorptive capacity with betweenness (subplot (b)) and mean path length (subplot (e)). The correlations somewhat reflect the relationship between profits and these network measures. It seems that firms having favourable network positions (high betweenness and short paths) are motivated to build up absorptive capacities. This implies that firms adjust their learning (particularly from partners) depending on their position.

Now we turn to the effect of changes in voluntary spillovers (Figures 12-14). First we observe from the correlation in subplot (a) that betweenness is highly positively correlated with performance. The correlation does not change much with variations in δ_c . This suggests that brokerage positions are consistently favourable in a regime characterised by increasing voluntary spillovers. In such regimes, tacitness is high and cooperation is considered to be essential (von Hippel, 1989). Occupying brokerage positions thus confers some controlling power on firms. Again, it is crucial to note that this result arises not from social capital but out of knowledge-driven alliance formation. In this sense, a high betweenness value can be interpreted as being located in a clustered part of the knowledge space and having influence in the knowledge diffusion process.

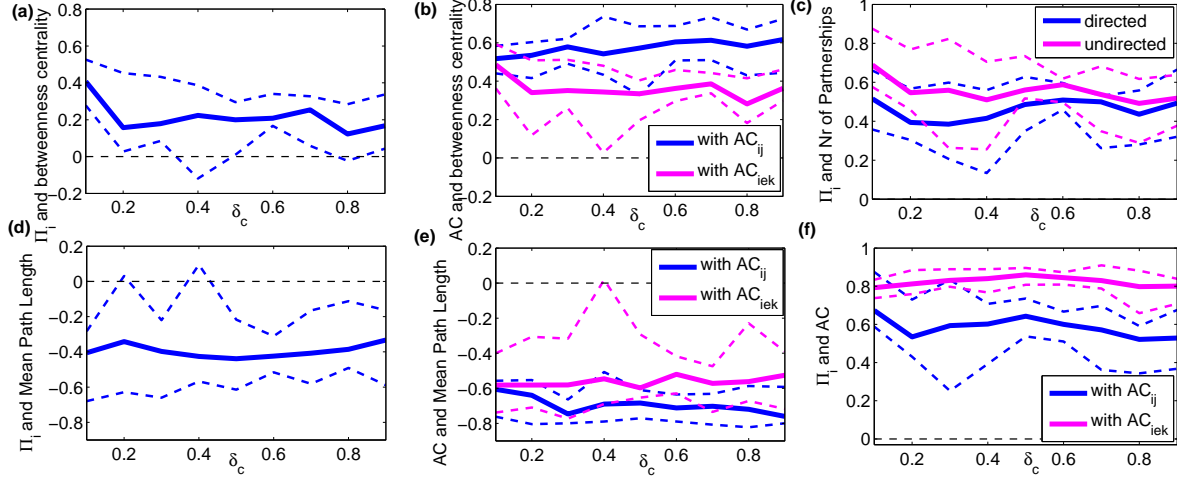


Figure 12: Correlations with firm performance for different δ_c and unilateral partnership

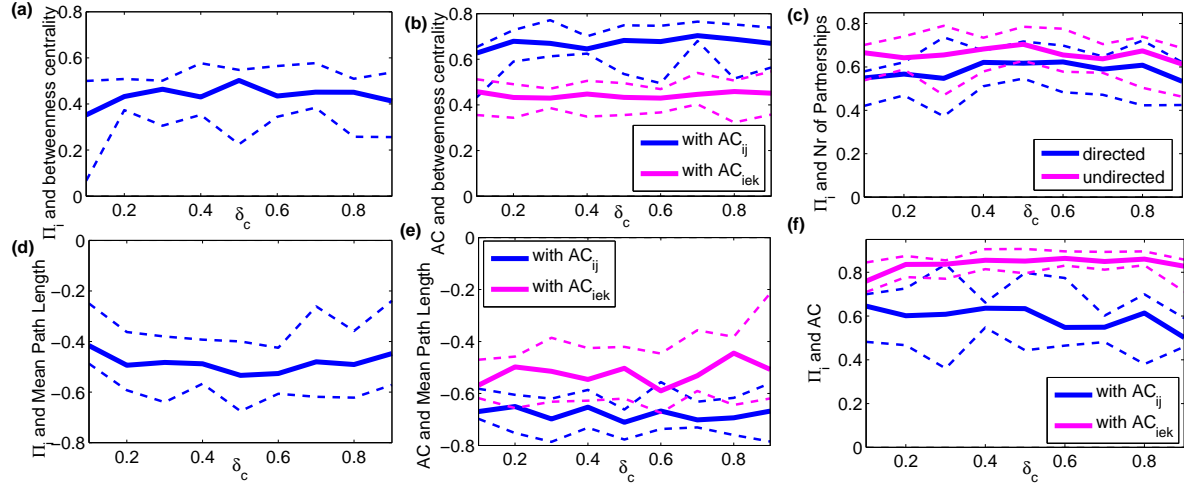


Figure 13: Correlations with firm performance for different δ_c and reciprocal partnership

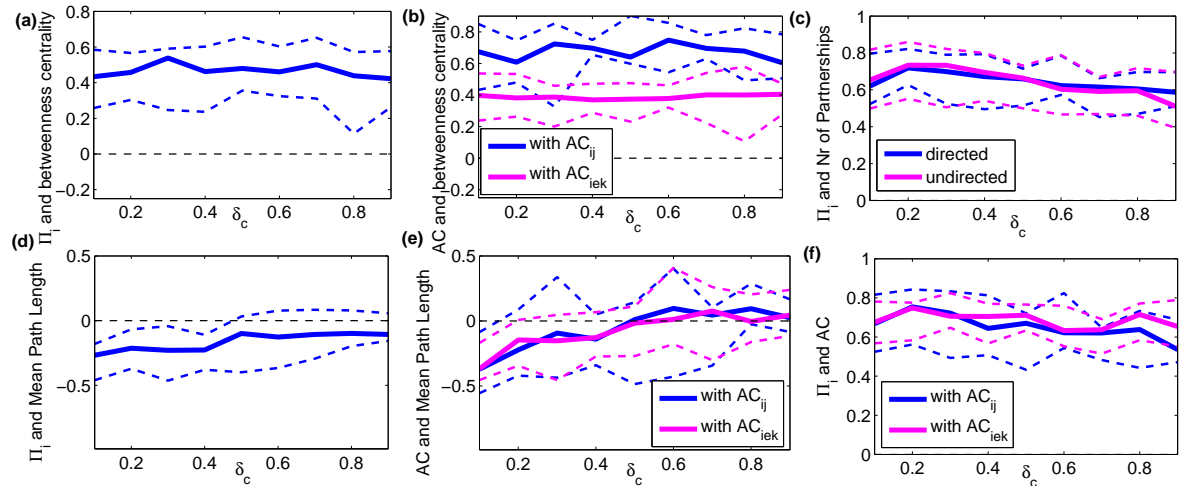


Figure 14: Correlations with firm performance for different δ_c and popularity contest

Mean path length is negatively correlated with performance (subplot (b)) meaning that low accessibility impairs innovativeness. In contrast to involuntary spillovers, this

relationship here does not vary much for different voluntary spillovers. In sum, a combination of high betweenness and short path length are consistently important for innovation performance in a highly tacit knowledge regime.

The performance effects of the number of partnerships (subplot (c)) fluctuates as δ_c rises. It is highest at some intermediate values of δ_c , falling otherwise. This seems to reflect an empirically observed problem associated with alliances. As Ahuja (2000) argued, at high levels of embeddedness, the marginal costs of every additional linkage will outweigh the marginal benefits. Absorptive capacity is consistently positively associated with profits. This correlation does not change much with changes in the amount of voluntary spillovers (subplot (f)). As observed with involuntary spillovers, the correlations of absorptive capacity with betweenness and mean path length somewhat reflect the relationship between profits and the network measures. Taken together with the earlier observation, these results indicate that firms display heterogeneity in building up absorptive capacity depending on the network positions they occupy.

4.3 Robustness checks

To analyse robustness of the results discussed above, a number of alternative settings are examined.⁴¹ First, in the simulation we have set the marginal returns to both inventive and absorptive R&D as equal.⁴² It is appealing, however, to try out scenarios where this does not hold (that is, $\psi \neq \xi$). To this end, we set either ψ or ξ equal to 0.75 leaving all other parameters unchanged and repeat the ABM simulation comparing results with the *baseline scenario* (the one described before).

What we find is that for $\xi = 0.75$ investments in inventive R&D become naturally more lucrative (ρ_s rise close to 90%), firms' absorptive capacities in all the scenarios reach on average lower values as in the baseline setting (between 0.2 and 0.6), which makes partnerships less efficient and their number drops by almost one half. The latter results in lower weighted small world ratios (due to lower clustering coefficients). However, rescaled small world ratios remain robust and are consistently above one at least in the unilateral and reciprocal partnership formation rules. Similarly, patterns of correlations identified in the baseline scenario between R&D profits and network positions (betweenness and mean path length) and also the positions and firms' absorptive capacities remain stable. 40%-60%), which results in somewhat higher average absorptive capacity values (particularly, in the popularity contest matching) - between 0.7 and 0.8 - and larger numbers of R&D alliances in all the scenarios considered (about 25% higher than in the baseline setting). Small worlds ratios of the emerging networks remain robust: while rescaled values consistently exceed one (even in the popularity contest matching), the weighted values are slightly higher (due to higher clustering). The patterns of correlations both between firms' network positions and profits, and network positions and learning capacities remain stable.

Second, we have assumed that firms have perfect knowledge about cognitive distance between them and others. In other words, firm i knows how far it is from j in the knowledge space and j also has the same information. To see, how crucial this assumption is, we introduce some uncertainty in this knowledge by adding some uniformly distributed error term ϵ reaching in its absolute maximum 50% of the distance between two firms,

⁴¹Detailed results are available upon request, but are not included in the paper for the sake of brevity.

⁴²In doing this we were aiming to obtain more general results not giving any preference to one of the investment directions.

$\epsilon \in [-\frac{1}{2}d_i; \frac{1}{2}d_i]$.⁴³ This estimation error is added to the distances both to voluntary and involuntary spillovers (i.e. to d_{iek} as well) during the matching scenarios and, hence, affects the partner choice and the investment decision. However, the R&D profits are then estimated with actual distances.

What we observe is that the main findings (on the small world properties, interdependencies between firms' network positions and profits/learning capacities) remain remarkably robust (not only qualitatively, but quantitatively). Among the most noticeable changes are:

- clearly lower quality of expectation about other firm's investment decision (correlation between $E^i(\rho_j)$ and ρ_j is about 50% only). This could have been expected since the uncertainty in the distances' evaluation affects the investment decision and, hence, the latter becomes less predictable;
- some lower average absorptive capacities of firms (0.4-0.6 for unilateral and reciprocal matching, 0.2-0.4 for popularity contest). Again, failing to estimate the distance exactly naturally leads to under-/over-investments in absorptive R&D and, hence, lower absorptive capacities (recall the inverted 'U'-shaped function in (7));
- some lower aggregate R&D profits of firms (by 10-20% maximum compared to the baseline scenario) which is primarily due to lower absorptive capacities.

It is worth to mention that increasing ϵ further (up to 100% of the distance between two firms), the emerging networks loose their small world properties, while firms' profits plummet further down. Thus, the ability to approximate the distance with a sufficient precision is found to be a very important competence firms must have to be efficient.

In brief, one can conclude that the main findings remain robust for different settings, although naturally it becomes impossible to try out all the different parameter combinations given the complexity of the model and the number of parameters included. A possible further step along this line would be to estimate some of the model parameters as described, for instance, in Gilli and Winker (2003); Winker and Gilli (2004). However, due to lack of suitable and readily available data, this is left for further research.

5 Conclusion

As an important determinant of learning, absorptive capacity plays a key role in firm-level innovativeness. Its role in the formation of R&D partnerships, and the resulting networks, is however, not well understood. This paper starts with the observation that earlier work on alliances has heavily focused on social capital explanations and that recent works which attempt to overcome this limitation seem to understate the role and the complexity of the absorptive capacity phenomena neglecting an important source of heterogeneity between firms resulting from it. We develop and simulate an agent-based model in which social capital is absent and alliances are formed based on knowledge fit depending on endogenous absorptive capacity. Three different matching scenarios are tested, one of which - the popularity contest - presents a simplified representation

⁴³In this way, the higher the distance between the two firms, the larger the potential error in estimating the cognitive distance between them.

of competition.⁴⁴ The remarkable result from the modeling exercise is that well-known empirical characteristics of networks are replicated by this ABM. What this tells is that disregarding the knowledge dimension in trying to explain the emergence, evolution and performance effects of networks gives, at best, a partial picture of reality.

The networks generated in the model display small world properties which respond to different extents to changes in the underlying knowledge regimes. The effects of these networks on performance vary depending on whether the knowledge space is characterised by intense cooperation and high voluntary spillovers (regime of high tacitness) or by relative dispersion and high involuntary spillovers (regime of high codification). In particular, in a regime of high tacitness, it seems to be more profitable for firms to occupy some kind of brokerage (high betweenness) and easily accessible (short path length) positions. This effect is less pronounced in a regime of high codification. Thus, at different stages of an industry's history, firms require different network strategies to achieve and maintain competitiveness through innovation.

A particularly important result relates to the role of absorptive capacity in network evolution. We observed a consistently strong and positive correlation between firms' absorptive capacity and their network centrality. This implies that being in a favourable network position relies on a higher level of absorptive capacity than being on a periphery. The consistently negative relationship between absorptive capacity and mean path length tells a consistent story. Efficient knowledge diffusion within a network requires that firms build up sufficient levels of absorptive capacity. To maximise their benefits, therefore, firms tend to adjust their absorptive capacity depending on their network positions. This heterogeneous behaviour is pronounced at extreme spillover levels.

Echoing recent studies (Cowan et al., 2007; Baum et al., 2010), our model further advances the possibility that empirically observed properties of inter-firm networks may be due to the characteristics of the knowledge space rather than purely social capital. Beyond this, however, we identify a time-varying characteristic of the knowledge space which helps to explain the network properties - that is, variations in the amount of knowledge spillovers. Network structures observed in mature industries characterised by high amounts of involuntary spillovers affect firm-level performance differently from the structures observed in early-stage industries characterised by high amounts of voluntary spillovers. By extension, network-based policy mechanisms (such as clustering initiatives) need to take into account the stage of an industry's development.

Our study may serve as a basis for a large number of extensions. Among those, one may set firms' R&D budgets dependent on their past profits, instead of time invariant and randomly allocated. Besides, we hope that the results of this modeling exercise will guide a fresh wave of empirical investigations. In particular, analysis of strategic alliances in industries where networking is pervasive (such as biotechnology, pharmaceuticals and information and communication technologies) may benefit from the results stated above.

⁴⁴What we generally find in this way is that competition reducing the number of alliances detrimentally affects firms' innovativeness. On the complexity of modeling competition in this type of models see, e.g., Baum et al. (2010, p. 2108).

Appendices

A Formal Definitions on Network Analysis

The definitions of the networks and its characteristics used in this study are consistent with the latest studies in this research area (see, e.g. Baum et al., 2010). However, some minor differences are possible. This section is meant to clarify them.

A.1 Networks

The simulated population of firms N and the L links (or ‘partnerships’, or ‘alliances’ throughout the paper) over 100 periods represent the resulting *network*. The two firms (*nodes*) are connected if there exists a *link* l_{ij} in the network L . The more links to distinct partners firm i has, the larger its *neighbourhood* (firms to whom i is directly connected) $Ne_i^L = \{ij \in L, i \neq j\}$, which is sometimes denoted as *degree* (the number of links to distinct partners held by i : l_i). The *average degree* of a network, therefore, is simply an average over all nodes’ degrees. Another related measure in this context is *density* measured as the sum of all links presented in the network L divided by the number of all possible ties (repeated alliances do not count here), i.e.

$$Density^L = \frac{\sum l_{ij} \in L}{N(N-1)/2}$$

with $N(N-1)/2$ being the total possible number of (undirected) links in the network.

A.2 Unweighted measures

Considering network characteristics we are most interested in the following three:

- The *clustering coefficient* measured as an average over neighbourhood clustering of each firm in the network, where the neighbourhood clustering of firm i is the proportion of neighbours who are neighbours of each other, i.e. are directly connected:

$$c_i = \frac{\sum l_{jh} \in L : j, h \in Ne_i^L}{l_i(l_i - 1)/2}.$$

- The *mean path length* is the average of all pairwise shortest distances between two nodes in a given network (computed by means of the Dijkstra’s algorithm). The more the distinct nodes are located on the shortest path, the larger the resistance of the path. To cope with infinite distances (if the population of firms is split in two distinct networks), we equalize them to the maximum distance within the network containing the node and adding one more unit to the distance, i.e. making the distance largest available within the given network.
- The *betweenness centrality* of firm i in the network L is the proportion of the shortest paths between any two other nodes in the network which pass through i ($p_{h,i,j}$) to

total number of shortest paths between these two nodes ($p_{h,j}$):

$$b_i = \frac{\sum_{h,j \neq i} p_{h,i,j} \in L}{p_{h,j}}.$$

A.3 Weighted measures

To take into account the number of times each partnership was over the last 100 periods, we construct a cumulative matrix of firms' past alliances W (an example of such matrix with the distinction between direction of links is illustrated in Figure 2). Hence, each element of the matrix has a weight $0 \leq w_{i,j} \leq 100$, with $\forall i \neq j$, $\sum w_{i,j}$ capturing the *strength* of the link, i.e. its weighted degree.

Using the matrix W all three network characteristics described in Section A.2 can be 'weighted'. For the *weighted clustering coefficient* there is a large variety of ways of doing this (a brief but comprehensive review is provided by Saramäki et al. (2007)). In this study we implemented the version described in Onnela et al. (2003), which is similar to the one used in Baum et al. (2010) (e.g., by taking into account weights of all links of triangles in which firm i is involved). In particular, weighted clustering coefficient of each node is defined as the geometric average of subgraph link weights:

$$c_i^w = \frac{1}{l_i(l_i - 1)} \sum_{j,h} (\hat{w}_{ij} \hat{w}_{ih} \hat{w}_{jh})^{1/3},$$

where \hat{w}_{ij} are node weights normalised by the maximum weight in the network L : $\hat{w}_{ij} = w_{ij}/\max(w)$. Thus, $c_i^w \in [0, 1]$ due to the normalisation and if $\hat{w}_{ij} \in [0, 1]$ an unweighted clustering coefficient can be recovered. Furthermore, contribution of each triangle to c_i^w is proportional to the weight of each link in the triangle.

As for *weighted mean path length*, the Dijkstra's algorithm finds the least resistance paths with the distinction that the each link's resistance equals the inverse weight, $w'_{i,j} = 1/w_{i,j}$, indicating the lowest resistance by the most frequently activated partnership.

Finally, the *weighted betweenness centrality* again uses each link's resistance set equal to the inverse weight.

B Further Results

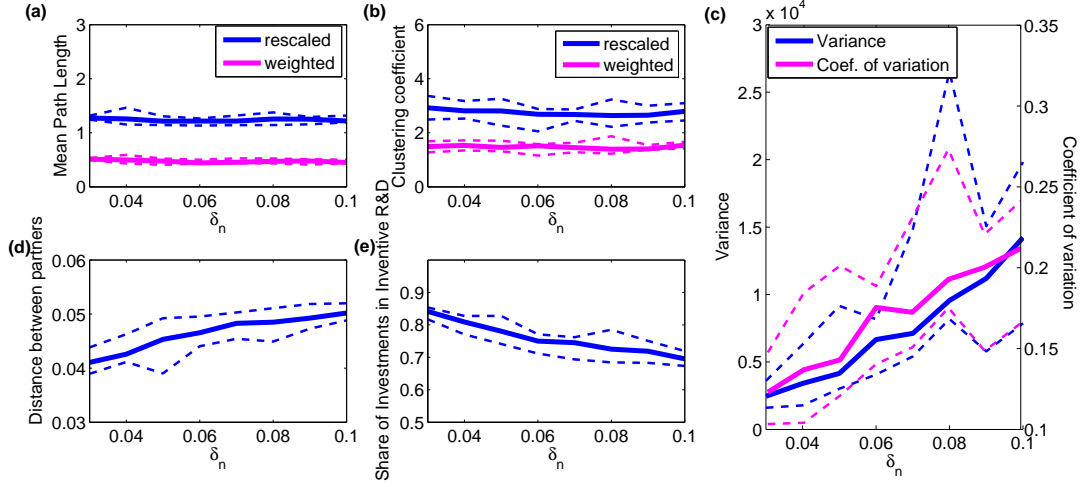


Figure 15: Network characteristics for different δ_n and unilateral partnership

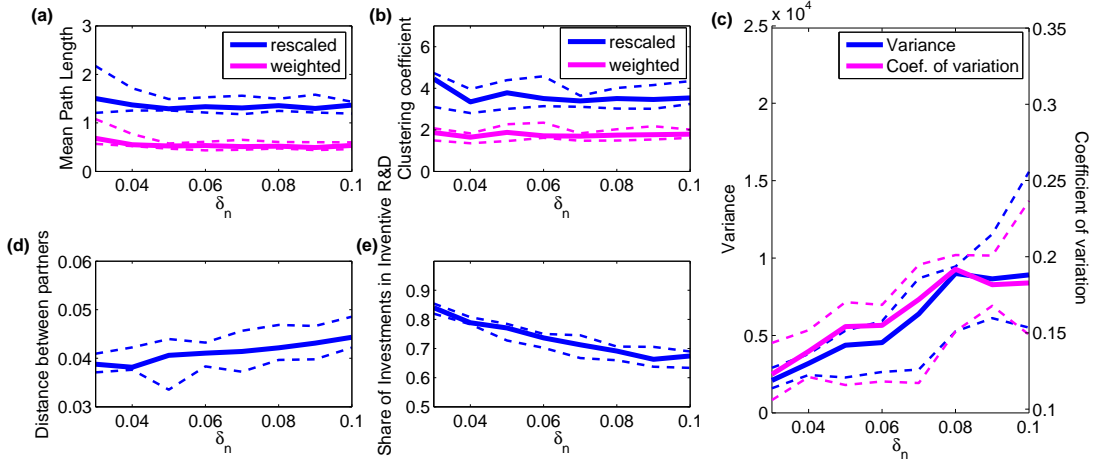


Figure 16: Network characteristics for different δ_n and reciprocal partnership

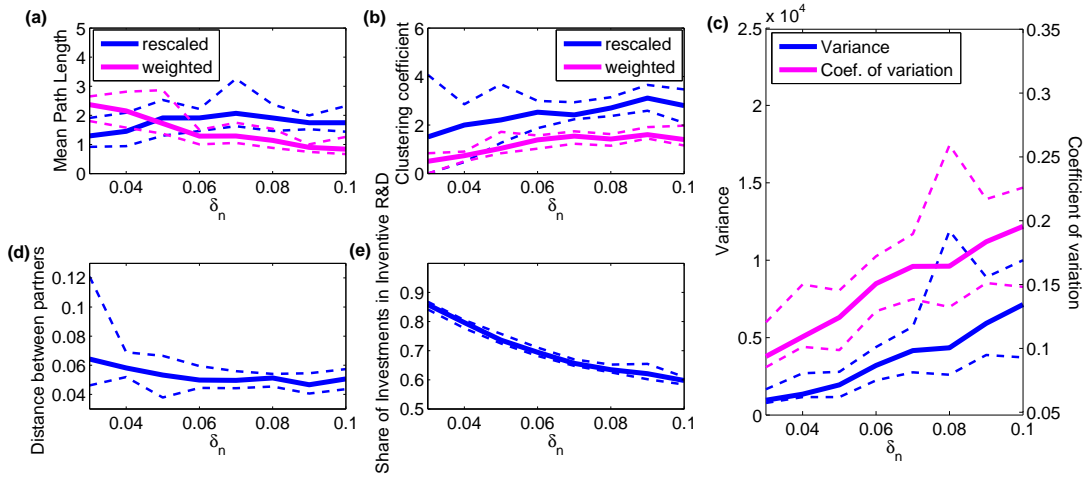


Figure 17: Network characteristics for different δ_n and popularity contest

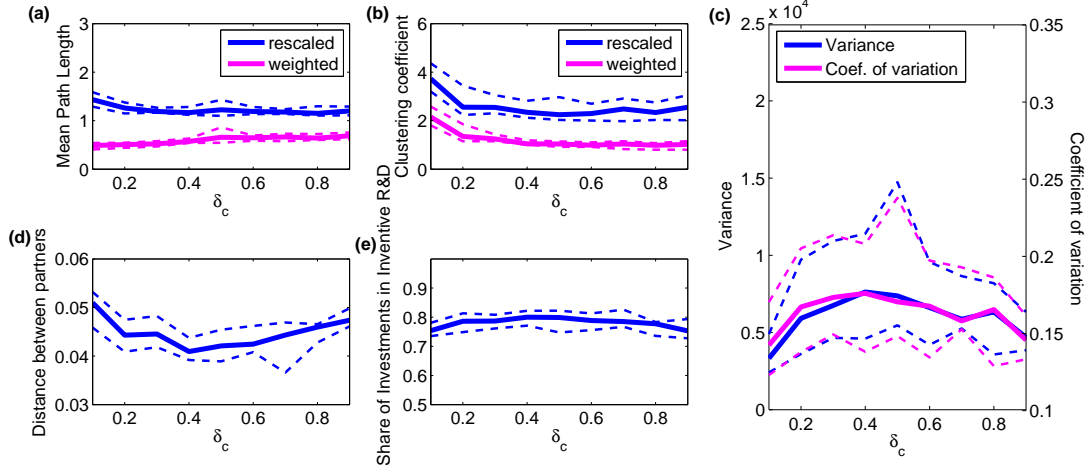


Figure 18: Further network characteristics for different δ_c and unilateral partnership

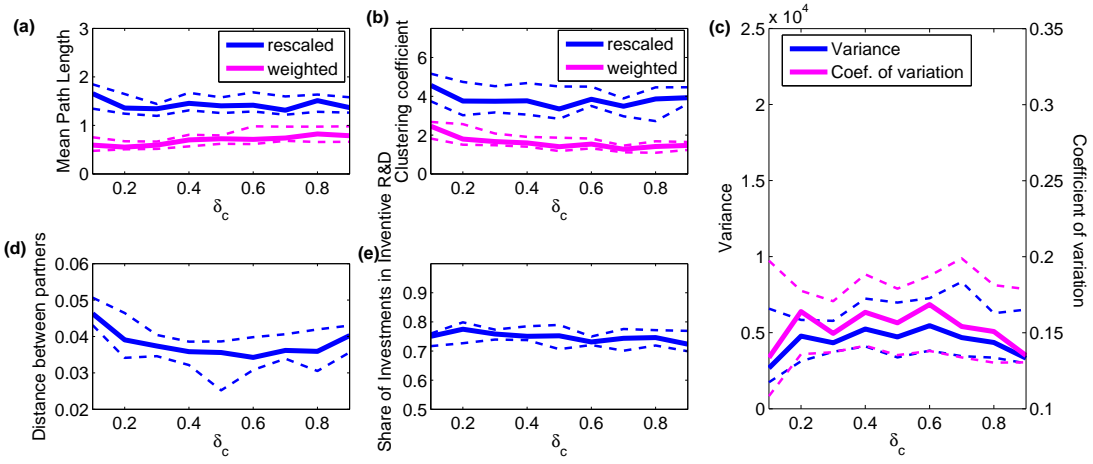


Figure 19: Further network characteristics for different δ_c and reciprocal partnership

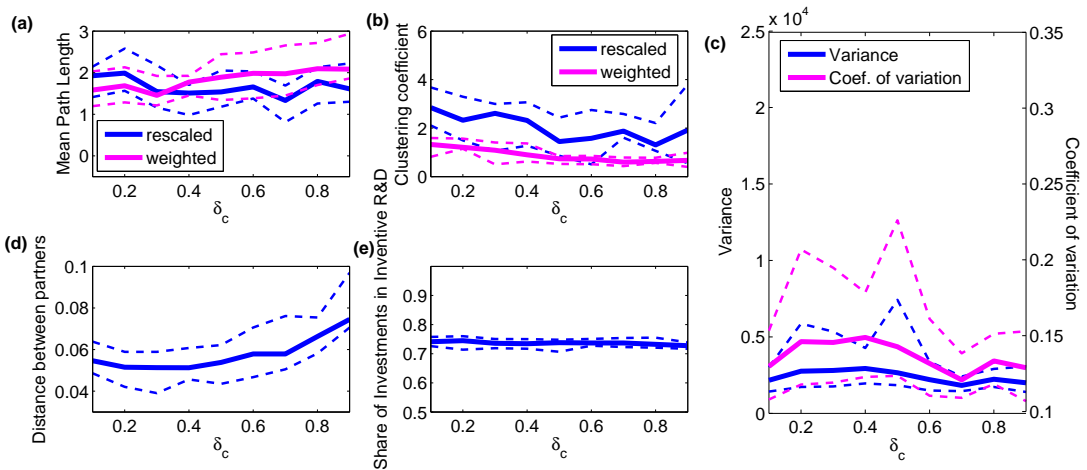


Figure 20: Further network characteristics for different δ_c and popularity contest

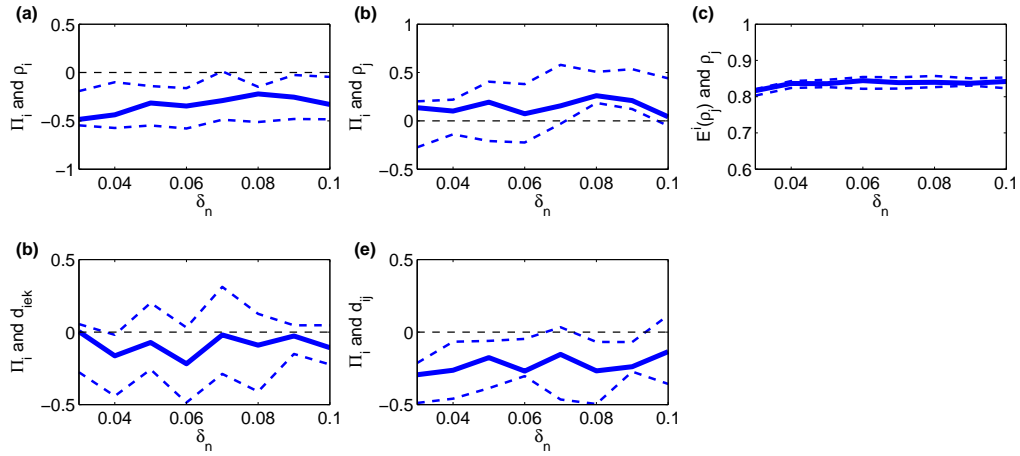


Figure 21: Further correlations with firm performance for δ_n and unilateral partnership

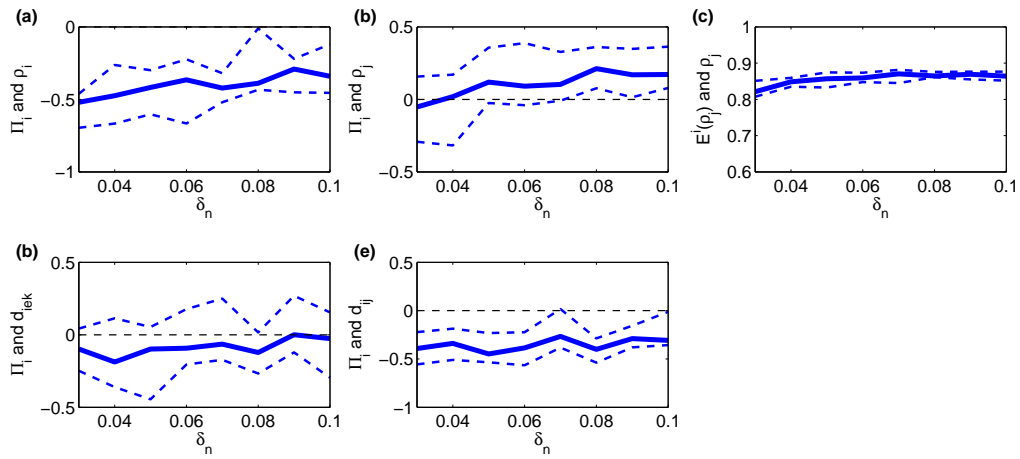


Figure 22: Further correlations with firm performance for δ_n and reciprocal partnership

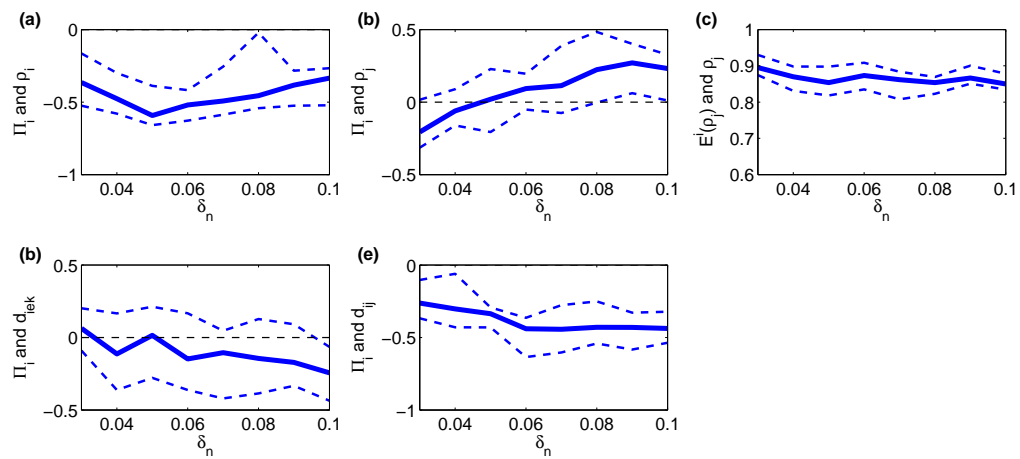


Figure 23: Further correlations with firm performance for δ_n and popularity contest

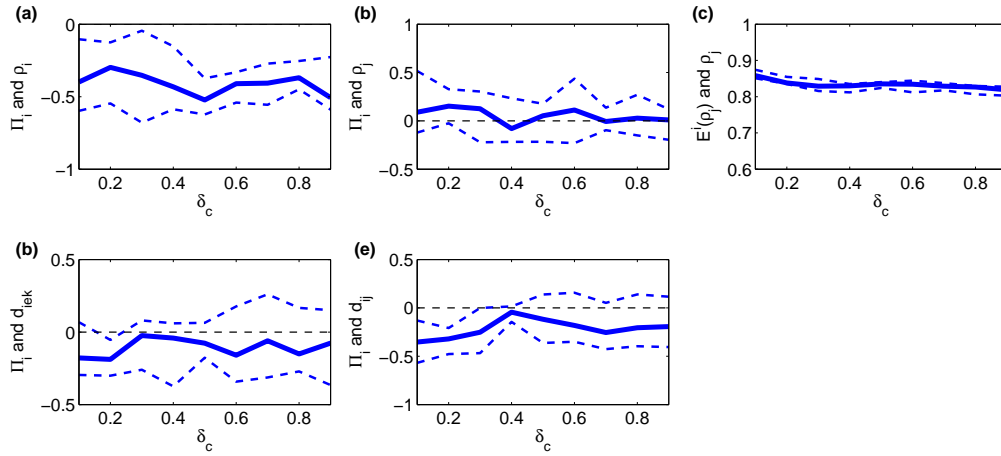


Figure 24: Further correlations with firm performance for δ_c and unilateral partnership

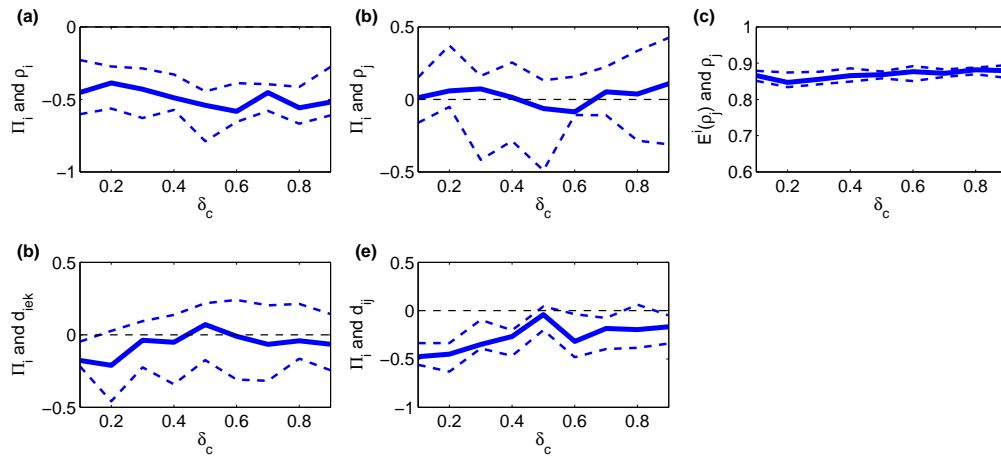


Figure 25: Further correlations with firm performance for δ_c and reciprocal partnership

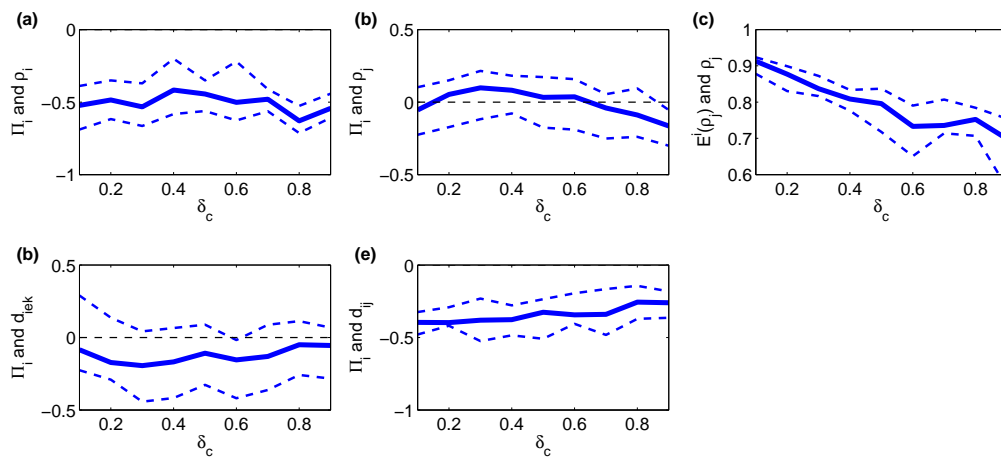


Figure 26: Further correlations with firm performance for δ_c and popularity contest

References

- Ahuja, G. (2000). Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly*, 45(3):425–455.
- Ahuja, G. and Lampert, M. C. (2001). Entrepreneurship in the large corporation: a longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal*, 22(6-7):521–543.
- Antonelli, C., Krafft, J., and Quatraro, F. (2010). Recombinant knowledge and growth: The case of ICTs. *Structural Change and Economic Dynamics*, 21(1):50–69.
- Audretsch, D. and Feldman, M. (1996). Innovative clusters and the industry life cycle. *Review of Industrial Organisation*, 11:253–273.
- Bala, V. and Goyal, S. (2000). A noncooperative model of network formation. *Econometrica*, 68(5):1181–1229.
- Baum, J., Shipilov, A. V., and Rowley, T. J. (2003). Where do small worlds come from? *Industrial and Corporate Change*, 12(4):697–725.
- Baum, J. A. C., Cowan, R., and Jonard, N. (2010). Network-independent partner selection and the evolution of innovation networks. *Management Science*, 56(11):2094–2110.
- Blueschke, D., Blueschke-Nikolaeva, V., and Savin, I. (2013). New insights into optimal control of nonlinear dynamic econometric models: Application of a heuristic approach. *Journal of Economic Dynamics and Control*, 37(4):821 – 837.
- Burt, R. (2004). Structural holes and good ideas. *American Journal of Sociology*, 110(2):349–399.
- Cantner, U. and Meder, A. (2007). Technological proximity and the choice of cooperation partner. *Journal of Economic Interaction and Coordination*, 2:45–65.
- Cohen, W. and Levinthal, D. (1989). Innovation and learning: The two faces of R&D. *The Economic Journal*, 99(397):569–596.
- Coleman, J. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94:95–120.
- Cowan, R. (2005). Network models of innovation and knowledge diffusion. In Breschi, S. and Malerba, F., editors, *Clusters, Networks and Innovation*. Oxford University Press.
- Cowan, R. and Jonard, N. (2004). Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control*, 28(8):1557–1575.
- Cowan, R., Jonard, N., and Zimmermann, J. (2007). Bilateral collaboration and the emergence of innovation networks. *Management Science*, 53(7):1051–1067.
- Davis, G. F., Yoo, M., and Baker, W. E. (2003). The small world of the American corporate elite, 1982-2001. *Strategic Organization*, 1(3):301–326.
- de Jong, J. and Freel, M. (2010). Absorptive capacity and the reach of collaboration in high technology small firms. *Research Policy*, 39(1):47–54.

- Dosi, G. (1988). The nature of the innovative process. In Dosi, G., Freeman, C., Nelson, R., Silverberg, G., and Soete, L., editors, *Technical Change and Economic Theory*, pages 221–238. Pinter London.
- Egbetokun, A. and Savin, I. (2012). Absorptive capacity and innovation: When is it better to cooperate? Technical Report 2012-056, Jena Economic Research Papers.
- Fabrizio, K. (2009). Absorptive capacity and the search for innovation. *Research Policy*, 38(2):255–267.
- Gilli, M. and Winker, P. (2003). A global optimization heuristic for estimating agent based model. *Computational Statistics & Data Analysis*, 42(3):299–312.
- Gilsing, V., Nooteboom, B., Vanhaverbeke, W., Duysters, G., and van den Oord, A. (2008). Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. *Research Policy*, 37(10):1717 – 1731.
- Gulati, R. (1995). Social structure and alliance formation patterns: A longitudinal analysis. *Administrative Science Quarterly*, 40(4):619–652.
- Hammerschmidt, A. (2009). No pain, no gain: An R&D model with endogenous absorptive capacity. *Journal of Institutional and Theoretical Economics*, 165(3):418–437.
- Koberg, C. S., Detienne, D. R., and Heppard, K. A. (2003). An empirical test of environmental, organizational, and process factors affecting incremental and radical innovation. *The Journal of High Technology Management Research*, 14(1):21–45.
- Lin, C., Wu, Y.-J., Chang, C., Wang, W., and Lee, C.-Y. (2012). The alliance innovation performance of R&D alliances: the absorptive capacity perspective. *Technovation*, 32(5):282 – 292.
- Moran, P. (2005). Structural vs. relational embeddedness: Social capital and managerial performance. *Strategic Management Journal*, 26(12):1129–1151.
- Mowery, D. C., Oxley, J. E., and Silverman, B. S. (1998). Technological overlap and interfirm cooperation: implications for the resource-based view of the firm. *Research Policy*, 27(5):507–523.
- Nooteboom, B. (1999). Innovation and inter-firm linkages: new implications for policy. *Research Policy*, 28(8):793 – 805.
- Nooteboom, B., Haverbeke, W. V., Duysters, G., Gilsing, V., and van den Oord, A. (2007). Optimal cognitive distance and absorptive capacity. *Research Policy*, 36(7):1016–1034.
- Onnela, J.-P., Chakraborti, A., Kaski, K., Kertesz, J., and Kanto, A. (2003). Dynamics of market correlations: Taxonomy and portfolio analysis. *Phys Rev E*, 68(<http://www.ncbi.nlm.nih.gov/pubmed/17358454>).
- Pittaway, L., Robertson, M., Munir, K., Denyer, D., and Neely, A. (2005). Networking and innovation: a systematic review of the evidence. *International Journal of Management Reviews*, 5(3-4):137–168.

- Potter, A. and Watts, H. D. (2011). Evolutionary agglomeration theory: increasing returns, diminishing returns, and the industry life cycle. *Journal of Economic Geography*, 11:417–455.
- Powell, W. (1998). Learning from collaboration: knowledge and networks in the biotechnology and pharmaceutical industries. *California Management Review*, 40(3):228–240.
- Powell, W., Koput, K., Smith-Doerr, L., and Owen-Smith, J. (1999). Network position and firm performance: Organizational returns to collaboration in the biotechnology industry. *Research in the Sociology of Organizations*, 16(1):129–159.
- Rosenkopf, L. and Schilling, M. (2007). Comparing alliance network structure across industries: observations and explanations. *Strategic Entrepreneurship Journal*, 1(3-4):191–209.
- Rowley, T., Behrens, D., and Krackhardt, D. (2000). Redundant governance structures: an analysis of structural and relational embeddedness in the steel and semiconductor industries. *Strategic Management Journal*, 21(3):369–386.
- Saramäki, J., Kivelä, M., Onnela, J., Kaski, K., and Kertesz, J. (2007). Generalizations of the clustering coefficient to weighted complex networks. *Phys Rev E Stat Nonlin Soft Matter Phys*, 75(<http://www.ncbi.nlm.nih.gov/pubmed/17358454>).
- Schilling, M. and Phelps, C. (2007). Interfirm collaboration networks: The impact of large-scale network structure on firm innovation. *Management Science*, 53(7):1113–1126.
- Storn, R. and Price, K. (1997). Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359.
- Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly*, 42(1):35–67.
- Verspagen, B. and Duysters, G. (2004). The small worlds of strategic technology alliances. *Technovation*, 24(7):563 – 571.
- von Hippel, E. (1989). Cooperation between rivals: informal know-how trading. In Carlsson, B., editor, *Industrial Dynamics, Technological, Organizational and Structural Changes in Industries and Firms*, pages 157–176. Kluwer Academic Publishing, Boston.
- Watts, D. J. and Strogatz, S. (1998). Collective dynamics of small-world networks. *Nature*, 393(6684):440–442.
- Winker, P. and Gilli, M. (2004). Applications of optimization heuristics to estimation and modelling problems. *Computational Statistics & Data Analysis*, 47(2):211–223.
- Wuyts, S., Colombo, M., Dutta, S., and Noteboom, B. (2005). Empirical tests of optimal cognitive distance. *Journal of Economic Behaviour & Organization*, 58(2):277–302.
- Zahra, S. and George, G. (2002). Absorptive capacity: A review, reconceptualization, and extension. *Academy of Management Review*, 27(2):185–203.