

Report on "Modeling the formation of R&D alliances: An agent-based model with empirical validation" (*MS2482*).

January 4, 2018

The paper provides interesting insights into the formation of R&D alliances by using an approach which combines agent-based model and network theory. Even if this paper goes in the right direction in order to find new analytical tools able to better identify, monitor and address the dynamics of R&D networks, it is very weak from many points of view. In particular: i) an accurate analysis of the empirical and theoretical literature on R&D alliances and R&D networks is missing; ii) a clear description of the original contribution of this paper with respect to existing literature, and specifically with respect to Tomasello et al. 2014, 2017, is missing; iii) many model assumptions are not clearly justify or linked with empirical evidence derived from other studies; iv) the analysis on the empirical dataset, which could be a winning point, is quite poor. Finally, the authors do not present a clear robustness analysis on the simulated results. In conclusion, I wish that the authors respond to the following comments:

Point 1

The existing literature on R&D alliances and R&D network is quite extensive. Specifically, agent-based literature has illustrated different mechanisms able to generate R&D networks ranging from game theoretical models (see Goyal and Joshi 2003) to more sophisticated network models (see Koenig et al. 2011; Vitali et al. 2013). Moreover, some studies have generated R&D networks by using a fitness mechanism based on agents' performance (see Tedeschi et al. 2014). General speaking, all the above mentioned models are able to reproduce many stylized facts observed in innovation networks.

Other studies, have focused on important empirical phenomena on technological innovation and industrial dynamics (Dosi et al. 1997; Klepper and Simons 1997; Czarnitzki and Fier; Czarnitzki et al. 2007; dAspremont and Jacquemin 1988;

Dawid 2006)

I think that references to the literature are not exhaustive, covering not all the most important works in this field.

Point 2

Closely related to the previous point, some studies specifically deal with R&D fitness networks. These studies find results in line with the present paper (see Tomasello et al 2014-2017). In this respect a question arises: what is the original contribution of this paper?

Point 3

In the article there are sentences not appropriately justified. For instance, just to name a few: "To address this questions, we can build on a number of empirical studies about R&D alliances". "Previous modeling attempts in this field have...". I think you have to justify these sentences by mentioned some of these studies.

Point 4

About the empirical findings, I have some doubts:

i) the dataset, presented in the section 2.1, should be better investigated. The data have a daily frequency (if I correctly understand the expression "with a granularity of 1 day"), however all the statistical analysis refers (if I understood correctly) to the aggregation of the whole sample from the 1984 to 2009. I think authors should better explain this point and, also produce some descriptive statistics on the daily dataset.

ii) it is not clear how the authors build pictures 1 and 2. Specifically, do graphs just represent one shot of the configuration of the R&D network or, again are the aggregation of the whole sample from the 1984 to 2009.

iii) Authors say: "Isolate nodes,...are simply excluded". I am fine with this simplification, however authors should clearly indicate the number of companies that remove the daily average and the standard deviation.

iv) Authors say: "...in R&D networks (see Tomasello et al., 2016), activity is a stable attribute..." This sentence must be clearly proven. By aggregating 26 years of data, the authors are in fact eliminating all the evolutionary analysis and the dynamics. This is a strong simplification and I am not sure authors can consider the whole R&D networks stable during 26 years. I would very much prefer to see all the analysis by using just the data of one single year (or smaller time windows).

v) Authors say: "It should be noted that an identification of the functional

form of the distribution (e.g., power-law, exponential, log-normal and so on) is outside of the scope of this study”. This is an important missing point. The ability to reproduce just the shape of a curve is not enough. The model must be able to reproduce the statistical distribution of Fig. 2 and also some moments of this distribution.

Point 5

About the modeling approach, I do not understand how many time each agents can play for simulation and how many independent simulations authors run.

Point 5

Authors identify the optimal region of the three parameters space (a_s, a_l, R) using the maximum likelihood estimation. In this regard, I have two comments: i) the effect of the exponent $1/N$ must be investigated ii) Authors find that R can vary between 1 to 20 and so they choose to use $R = 20$. However, they should investigate how the shape and the distribution of figure 5 change by varying R between 1 and 20. iii) A clear comparison between Figures 2 and 5 must be presented, something like relative errors between the two or statistical tests proving they belong to the same statistical distribution.

Point 6

In the Conclusion, authors present some interpretation of the results that are not presented in the results' analysis. I suggest authors to cut conclusion and to move the results in the appropriated section.

References:

- Goyal S, Joshi S (2003) Networks of collaboration in oligopoly. *Game Ec1on Behav* 43(1):57-85.
- Koenig MD, Battiston S, Napoletano M, Schweitzer F (2011) Recombinant knowledge and the evolution of innovation networks. *J Econ Beh Organ* 79:145-164.
- Vitali S, Tedeschi G, Gallegati M (2013) The impact of classes of innovators on technology, financial fragility and economic growth. *Ind Corp Chang* 22(3)
- Tedeschi, G., Vitali, S. & Gallegati, M. *J Evol Econ* (2014) 24: 817.
- Czarnitzki, D. and A. Fier (2003), Publicly Funded R&D Collaborations and Patent Outcome in Germany. ZEW, Zentrum fur Europaische Wirtschaftsforschung/Center for European Economic Research, Discussion Papers, Number 03-24.
- Czarnitzki, D., B. Ebersberger and A. Fier (2007), The relationship between

R&D collaboration, subsidies and R&D performance: Empirical evidence from Finland and Germany, *Journal of Applied Econometrics*, 22(7), 1347-1366.

-d'Aspremont, C. and A. Jacquemin (1988), Cooperative and noncooperative R&D in duopoly with spillovers, *The American Economic Review*, 78(5), 1133-1137.

-Dawid, H. (2006), Agent-based models of innovation and technological change, *Hand-book of Computational Economics*, 2, 1235-1272.