R&Dsimulab: a micro-policy simulator for an ex-ante assessment of the effect of public R&D policies

Pierpaolo Angelini, IRES-CGIL
Federico Cecconi, CNR-ISTC
Giovanni Cerulli, CNR-IRCrES *
Maria Augusta Miceli, Sapienza University
Bianca Maria Potì, CNR-IRCrES

Abstract. This paper presents "R&Dsimulab", a micro-policy simulator for an ex-ante assessment of public R&D policy effect when R&D and non-R&D performing companies are located within a network. We set out by illustrating the behavioural structure and the computational logic of this model, to then provide a simulative experiment where the total level of R&D activated by a fixed amount of public support is studies as function of companies' network topology.

More specifically, the suggested experiment shows that a large "hubness" of the network is – on average – more likely accompanied with a decreasing median of the aggregated total R&D performance of the system. Since the aggregated firm idiosyncratic R&D (i.e., the part of total R&D independent of spillovers) is slightly increasing, we conclude that positive cross-firm spillover effects - in the presence of a given amount of support - have a sizeable impact within less centralized network, where fewer hubs emerge. This may question the common belief which suggests that larger R&D externality effects should be more likely to arise when few "central" champions receive a support.

Keywords: R&D policy; networks; complexity; simulation

April 29, 2015

* Corresponding author. E-mail: giovanni.cerulli@ircres.cnr.it

1. Introduction

R&Dsimulab is a micro-policy simulator for an ex-ante assessment of public Research & Development (R&D) policy effect on companies' R&D activity. It is an agent-based computational model based on the interaction between a public agency, entitled to manage a direct (or grant-based) R&D policy, and a given set of companies eligible for receiving a monetary support to increase their actual level of R&D activity.

On the part of policymakers, such model can be used to build and compare ex-ante evaluation scenarios related to alternative policies aimed at fostering the R&D activity of companies undergoing a given public R&D support.

R&Dsimulab can be run either using a pre-defined set of parameters, thus exploring outcomes' sensitivity to parameters' changes, or by a calibration based on empirical evidence.

R&Dsimulab assumes that agents (the public agency and the companies) maximize an objective function under reasonable constraints, and assumes that companies doing R&D are placed within a network of firms where possible positive (or negative) externality effects can arise.

To our knowledge, no previous models of this type have been proposed in the literature so far. Therefore, R&Dsimulab constitutes a first attempt to build a policy simulator for an ex-ante assessment of R&D policy effects, whose scientific and policy-oriented scope can be worth exploring.

2. Modelling approach and methodology

R&Dsimulab is an agent-based simulative model. The agents constituting the model are: a public agency, which provides public funds to support private R&D companies, and a set of eligible-for-fund private companies. Companies are then assumed to be located within a network which can generate possible R&D externality effects.

Both types of agents take decisions by maximizing an objective function under reasonable constraints. The model runs under a series of assumptions which are illustrated below.

2.1 Companies' behaviour: the optimal R&D expenditure

Companies choose the level of R&D expenditures (R) which maximizes their profits. Thus, the optimal R is the one equalizing the (expected) marginal rate of return and the (expected) marginal capital cost of doing R&D. The optimal level of R&D is in turn a function of the R&D support (S) that a firm may potentially receive from the public agency.

We assume that each company owns an optimal level of the subsidy, thus making the R&D optimal equation as a convex function of the public support (a *parabola*, for instance). Finally, we also assume that R&D spillovers among firms may take place, due to companies' relationships within their R&D network, where the R&D flows from one company to another according to the strength of the relationship between firms. Therefore, each company R&D includes both an idiosyncratic component (R_{idio}) and an "additional" component due to the presence of R&D externalities. The sum provides the total R&D outlay of a company (R_{total}).

Firms run in a competitive market where they sell a single research output at a parametric/normalized price p_{rd} =1. Each single company owns a profit function convex in R_{idio}, where the production function is linear and costs are increasing (Howe and McFetridge, 1976; David, Hall and Toole, 2000; David and Hall, 2000). The total company R&D revenues are:

$$\phi R_{idio}^2$$

while the total company R&D costs:

$$h(S)R_{idio} + a$$

where $h(S)$ is a proper function of the subsidy S received, and a is a scale factor. The company profit function is thus defined as:

$$\Pi(Q,S,R_{idio}) = \phi R_{idio}^2 - \left[h(S)R_{idio} + a\right]$$

In order to maximize this function over R_{idio}, we calculate the first order condition obtaining:

$$\frac{\partial \Pi(Q,S,R_{idio})}{\partial R_{idio}} = 2\phi R_{idio} - h(S) = 0$$

If we hold that:

$$h(S) = aS - bS^2 - (Z - M)$$

implying that each company owns a desired level of S, given the parameters a, b, Z and M. Given this, we finally get:

$$R_{idio}(S) = \frac{1}{2\phi}\left\{aS - bS^2 - (Z - M)\right\} = p_{rd}S - kS^2 - (F - D)$$

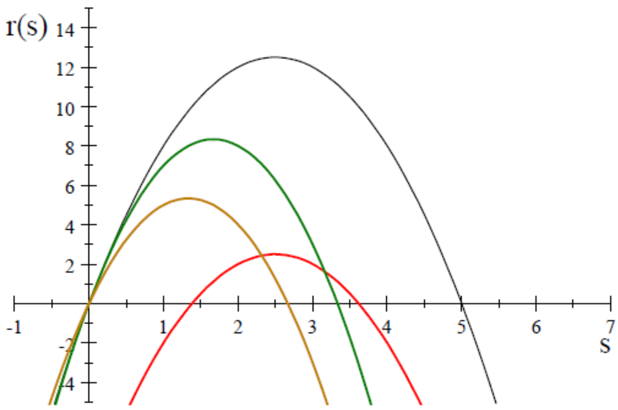
or, given that p_{rd} =1:

$$R_{idio}(S_i) = S_i - k_i S_i^2 - F_i + D_i$$

which represents the optimal "idiosyncratic" R&D expenditure of firm *i* as function of the level of the subsidy *S*. The parameters *F* and *D* in the previous formula have the following interpretation: *F* represents the fixed costs of performing R&D, implying that the higher *F*, the lower the level of the optimal R&D the company would be willing to provide (other things being equal). A large *F* could be however compensated by a large *D*, being *D* defined as the "degree centrality" of each company within the network. This mean that high fixed costs can be compensated by a more central role of the company in the network, an assumption that seems reasonable. Indeed, *ceteris paribus*, the more the firm is central in the network, the more its R&D should be high; this accounts for the stylized fact that more central companies in R&D networks are also those with a larger size of R&D expenditure.

Figure 1 illustrates the shape of company optimal R&D as a function of the subsidy *S* received from the public agency. Although such a shape entails that companies own an optimal R&D subsidy, they however cannot choose this level which is decided on the part of the public agency. It is worth to observe that such a shape basically suggests that subsidies produce not only benefits for the firm, but also costs that, beyond a certain threshold, can overcome benefits thus yielding a convex form of the *R*-*S* relationship.

Figure 1. Shape of the optimal R&D function as a function of the subsidy *S*.



Finally, since the network degree-centrality of the company strictly depends on network topology, it is more correct to indicate D as a function of the network M in which the company operates; we can thus re-write the last formula as:

$$R_{idio}(S_i) = S_i - k_i S_i^2 - F_i + D_i(M)$$

which makes it explicit that D depends on M, the network matrix.

2.2 Agency behaviour: optimal subsidy provision

Given a constant total amount of subsidy equal to $\bar{S}$, the direct objective of the public agency is that finding the optimal allocation of such amount by maximizing the total level of R&D (i.e., the sum of all companies' idiosyncratic R&D spending) (Cerulli, 2012; Laincz 2009; Jou and Lee, 2001).

To this end, we first assume that the agency knows the company ability to perform R&D and its centrality within the network, but it has no knowledge of firms' R&D network relationships. As objective, the agency wants to determine two things: (i) which companies are worth to support and which are not (i.e., selection-process); (ii) which share of $\bar{S}$ has the agency to provide to each firm selected for support. Thus, the agency comes up with two optimal solutions: (i) the N_1 (out of N) selected companies; (ii) the optimal allocation of the subsidy $\bar{S}$ within the N_1 selected companies.

The agency optimization problem is the following:

$$\begin{cases} \max_{\{S_1, S_2, \ldots, S_N\}} W(S_1, S_2, \ldots, S_N) = \sum_{i=1}^{N} R_{idio}(S_i) = \sum_{i=1}^{N} \left[S_i - k_i S_i^2 - F_i + D_i(M) \right] \\ s.t. \\ \sum_{i=1}^{N} S_i \leq \bar{S} \\ R_{idio}(S_i) \geq R_{idio}(0) \quad \forall i = 1, \ldots, N \\ S_i \geq 0 \end{cases}$$

Such a problem assumes that:

- the total sum of subsidies is fixed and equal to $\bar{S}$ (*budget constraint*);

- the expected firm R&D conditional on the subsidy is higher when a company is supported than when it is not supported;
- subsidies are positive numbers.

Observe that, in doing its choice, the public agency does not take into account network externalities, but only the arguments of the idiosyncratic R&D.

Once the Agency has chosen the N_1 units to support along with their level of support, all the N companies perform their actual R&D expenditure, given the support received, that is, R_{total}. The global policy effect is thus given by:

$$\hat{R}_{total} = \sum_{i=1}^{N} R_{total,i}$$

which is a function of network topological parameters, given any other idiosyncratic factor considered in the model.

2.3 Total R&D outlay by introducing network externality

In order to calculate the total R&D expenditure, once the level of the idiosyncratic R&D is known, it is first necessary to introduce the network in which company operate.

We assume the network to be represented by a weighting matrix $M=[m_{ij}]$, where m_{ij} represents the generic element of this matrix. Observe that M can contain either positive or negative values, thus accounting for positive or negative externalities respectively. At this step, however, we only consider positive externalities (i.e., $m_{ij} \geq 0$).

We define the total R&D outcome of company i as:

$$R_{total,i} = R_{idio,i} + \sum_{j=1}^{N} m_{ij} R_{idio,j}$$

where $\sum_{j=1}^{N} m_{ij} = 1$, and N is the total number of companies forming the network. This implies that, for the generic company i, its total R&D outlay is equal to its idiosyncratic R&D *plus* a weighted average of the idiosyncratic R&D of other companies. It means that a company can increase its total R&D performance either if the size of its relational weights is large, or if the number of its relational weights is large.

2.4 Simulating the effect of the policy under different network parameters

The model is characterized by the parameter space $\mathbf{\Theta}$ which defines the entire set of parameters under which the model can be simulated. This set also contains the network parameters as distinct parameters governing simulation results. If we indicate by θ one parameter characterizing a specific network topology, we can write:

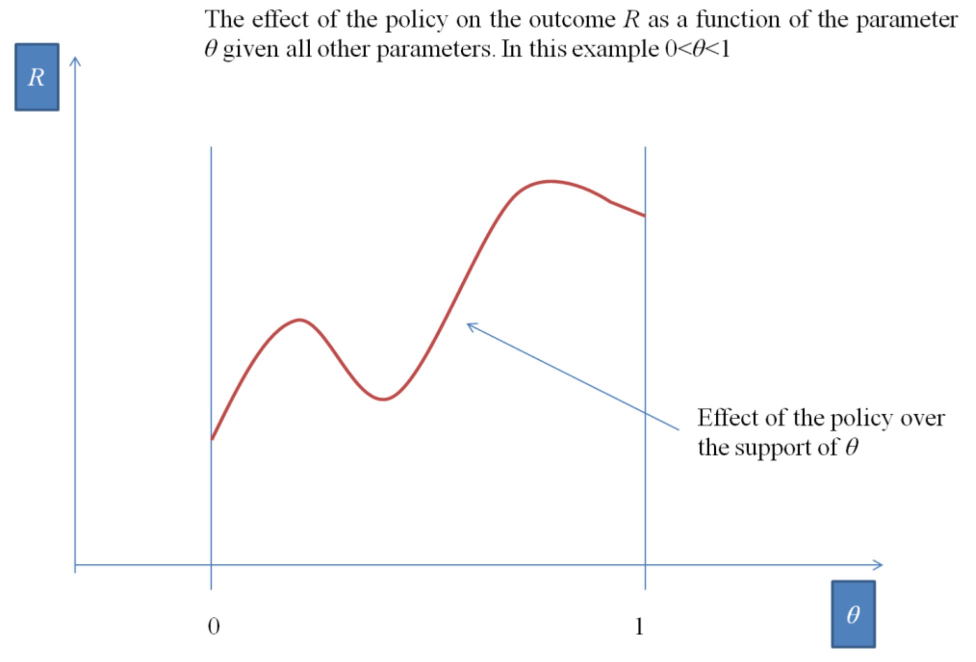
$$\mathbf{M} = \mathbf{M}(\theta).$$

Given this definition, we may be interested in studying the pattern of the function:

$$\hat{R}_{total} = \hat{R}_{total}(\theta \mid \mathbf{\Theta}_{-\theta})$$

to see how the policy effect changes under different level of θ, given the value of other parameters. The meaning of θ depends on the specific network topology considered. Figure 2 presents an illustrative example of the typical result we obtain in simulating our model this way.

Figure 2. Total R&D policy effect as function of the network topology parameter θ .



3. The logic of R&Dsimulab's simulations

Companies are located within an R&D network, and different network topologies can produce different policy effects. The network impacts on R_{total} in two ways: (i) the more a company is central in the network, the more a lower barrier to do R&D is assumed (thus reducing the fixed costs of doing R&D); (ii) different network topologies could provide different R&D performance. Therefore, running simulations under different policy scenarios provides guidance to detect the emerging properties of the R&D pattern under specific model's parameterizations.

R&Dsimulab uses Monte Carlo methods to provide reliable conclusions about simulation results. Are specific configurations of the network more likely to produce larger R&D effect than other types of settings? In order to answer questions like this, we can run a number of R&Dsimulab simulation exercises.

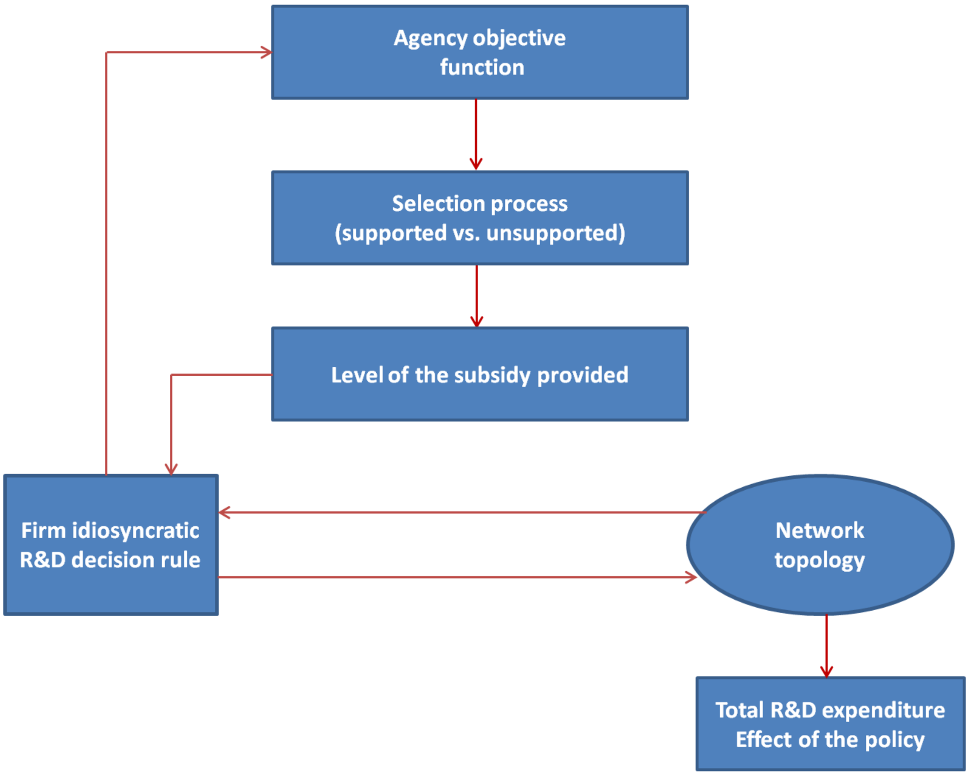
For example, one could be interested in identifying whether, ceteris paribus, a quasi-random network is or is not more conducive to higher levels of R&D than, for instance, networks characterized by the emergence of specific nodes playing as hubs. It may thus be interesting to assess whether the policy effect on R_{total} will show an increasing or decreasing pattern as a function of the network's "hubness".

Other experiments could also include the assessment of policy effect when other significant network parameters are changed or when one considers different network topologies, such as "scale-free" or "small-world" networks. Moreover, if an empirical measure of the actual network is available so that an empirical calibration of the model's parameters can be done, one may also provide an assessment of the impact of the R&D support policy on a real study context; this way, it is possible to use R&Dsimulab as a tool for an ex-ante evaluation of the considered R&D policy.

Figure 3 illustrates the path-diagram of the model simulation. The point of departure is the generation of a weighted network of firms (i.e., the matrix M) which determines company degree-centrality in the network and thus the value of $D_i(M)$. Given the value taken by other firm idiosyncratic parameters, the agency can operate by selecting the companies to support, and then by optimally allocating the subsidy $\overline{S}$ among them. Companies then provide their idiosyncratic R&D outlay which depends on the subsidy received (S_i), and other idiosyncratic parameters. Subsequently, companies perform their total R&D expenditure by transmitting and by receiving R&D spillovers according to the structure of matrix M, which defines the network topology. Finally, the sum of total R&D expenditure over the entire population of companies is calculated, which is the main outcome of interest.

Table 1 illustrates three network topologies we can consider within the model: random, scale-free, and small-world. This table also reports the main parameters characterizing these network topologies, and their meaning.

Figure 3. Path-diagram of the R&Dsimulab model simulation.



Finally, **Figure 4** sets out the model programming flow, which explains the computational steps through which the model is simulated. This figure shows that we have three simulation layers, based on: (i) type of experiment, (ii) given level of the network parameter(s), and (iii) run of the single simulation. To understand how this works, consider the random network case: the type of experiment is "random network", the parameter to fix is p (the "edge probability"), and a single run generates one single network based on the fixed p, as well as a draw from the distributions of all the idiosyncratic parameters needed to solve numerically the model.

As for the distributions from which firm idiosyncratic parameters are drawn, Table 2 displays meaning and type of distribution (or just the value) of each parameter.

Table 1. Description of network topologies and parameters.

Network topologies	**Parameters**	**Meaning**
Random	p	Edge probability
Scale-free	m_0	Initial number of nodes
	m	Number of edges added by each new node
	$m \leq m_0$	Condition
Small-world	k	Number of neighbours in the initial ring lattice
	p_w	Rewiring probability
	$N >> k >> ln(N) >> 1$	Condition

Figure 4. Model programming flow.

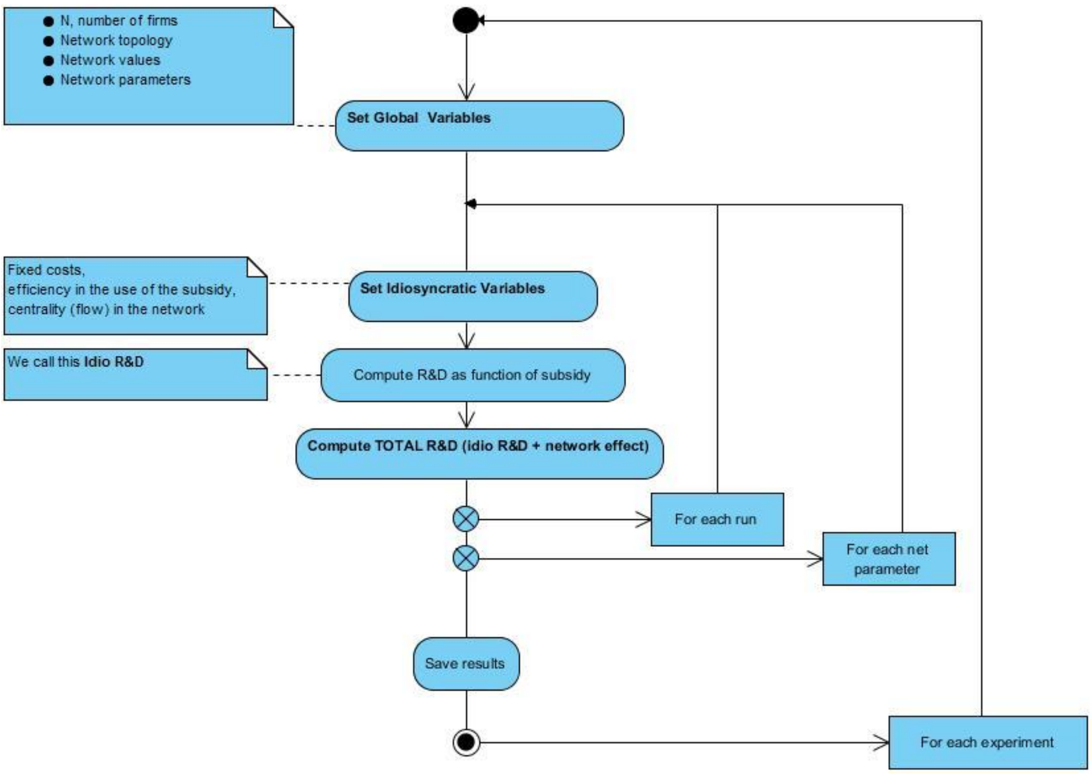


Table 2. Distribution function of model parameters.

Parameters	**Meaning**	**Distribution**
p_{rd}	Unit price of R&D	Fixed to the value 1
k_i	Degree of the R&D concavity as function of S	Uniform[0; 1]
F_i	R&D fixed costs	Uniform[0; 1]
D_i (M)	Firm degree-centrality in the network	Normalized to vary between [0; 1]
m_{ij}	Generic weight of the network matrix M	Uniform[0,1] and the normalized to get $\sum_j m_{ij} = 1$

4. Assessing the effect of R&D public support on total R&D by increasing network "hubness"

In this application of R&Dsimulab, we are interested in identifying whether, ceteris paribus, a random network is or not more conducive to higher levels of R&D than, for instance, networks characterized by the emergence of specific nodes playing as hubs. We will look at whether the policy effect on R_{total} shows an increasing or decreasing pattern as a function of an increasing "hubness" of the network.

We run a simulation by setting 100 companies, 5 runs, 334 distinct values of the hubness parameter, thus obtaining a simulated dataset of 167,000 observations. For each value of the hubness parameter we first calculate the median of the log of R_{total}, and then a regression of this variable on the parameter of hubness. The result is reported in Table 3, while Figure 5 shows the fit graphically.

Table 3. Linear regression of the median of the log of R_{total} on the hubness parameter.

Source	SS	df	MS		
Model	101.693861	1	101.693861	Number of obs =	166999
Residual	2457.61489	166997	.014716521	F(1,166997) =	6910.2
Total	2559.30875	166998	.015325386	Prob > F =	0.0000
				R-squared =	0.0397
				Adj R-squared =	0.0397
				Root MSE =	.12131

med_y	Coef.	Std. Err.	t	P>|t|	Beta
hubness	-.0002559	3.08e-06	-83.13	0.000	-.1993361
_cons	-.1750011	.000595	-294.10	0.000	.

It is immediate to see that the larger the hubness of the network, the lower the optimal level of R&D activated by an amount of subsidy $\overline{S}$ = 10. The beta coefficient is around -20 and it is highly significant. Figure 5 confirms this decreasing pattern, although it also emphasizes that the variability around the fit is rather huge, as confirmed by the low R-squared of this regression.

If we look at R_{idio}, the result is in the opposite sense, with a statistically significant (but moderate) increase of R_{idio} as soon as the hubness increases (see table Table 4). Figure 6 confirms this finding.

These findings suggest that, in the presence of positive R&D spillovers, and for a fixed amount of support, the median of the total population R&D tends to decrease as soon as few nodes become highly important in the network. This was not a trivial regularity to find out without such analysis. On average, and just by considering the median as representative moment of the entire distribution of the log of R_{total}, one can conclude that the policy R&D return to an investment equal to $\bar{S}$ is higher in R&D networks characterized by lower hubness. Observe, also, that this result is independent of the specific idiosyncratic parameters chosen, as what we performed in this experiment is a Monte Carlo simulation which explores by subsequent draws the entire support of the distribution of these parameters.

The fact that the R_{idio} is a moderately increasing with larger hubness may be explained by the fact that fixed costs of R&D decrease with a larger centrality-degree – i.e. $D(M)$ – which is an argument of the R_{idio} function.

Finally, Figure 7 illustrates the trend of the median of $\log(R_{total})$, and that of the median of $\log(R_{idio})$ as function of the hubness within the same graph. As expected, in the presence of positive R&D spillovers, the overall level of the median of $\log(R_{total})$ is higher than that of $\log(R_{idio})$.

Figure 5. Pattern of the median of log of R_{total} as a function of the network hubness parameter. A fixed amount of support is assumed.

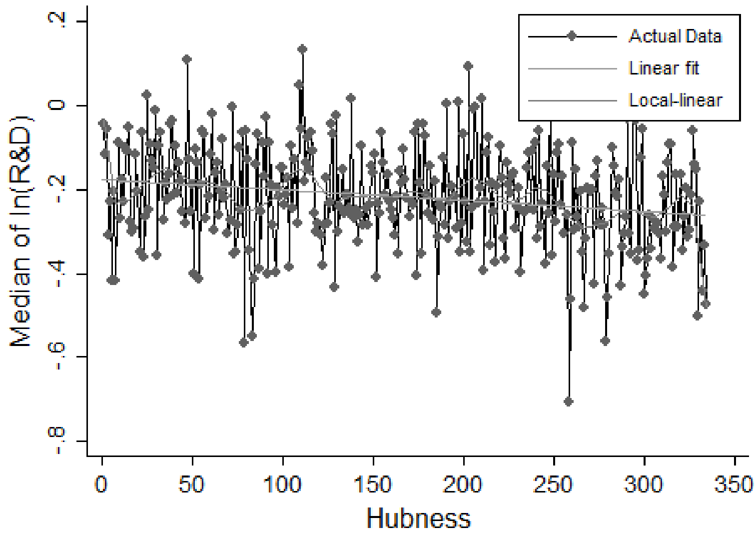


Table 4. Linear regression of the median of the log of R_{idio} on the hubness parameter.

Source	SS	df	MS
Model	.489141345	1	.489141345
Residual	7947.57963	166997	.047591152
Total	7948.06877	166998	.047593796

Number of obs	=	166999
F(1,166997)	=	10.28
Prob > F	=	0.0013
R-squared	=	0.0001
Adj R-squared	=	0.0001
Root MSE	=	.21815

med_y_ridio	Coef.	Std. Err.	t	P>|t|	Beta
parametro	.0000178	5.54e-06	3.21	0.001	.0078449
_cons	-1.588483	.0010701	-1484.46	0.000	.

Figure 6. Pattern of the median of log of R_{idio} as a function of the network hubness parameter. A fixed amount of support is assumed.

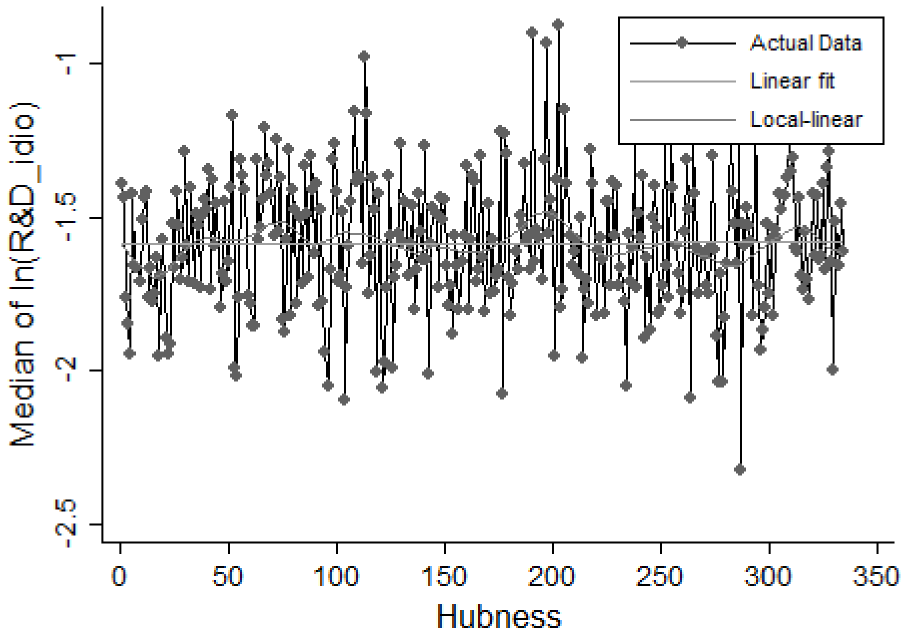
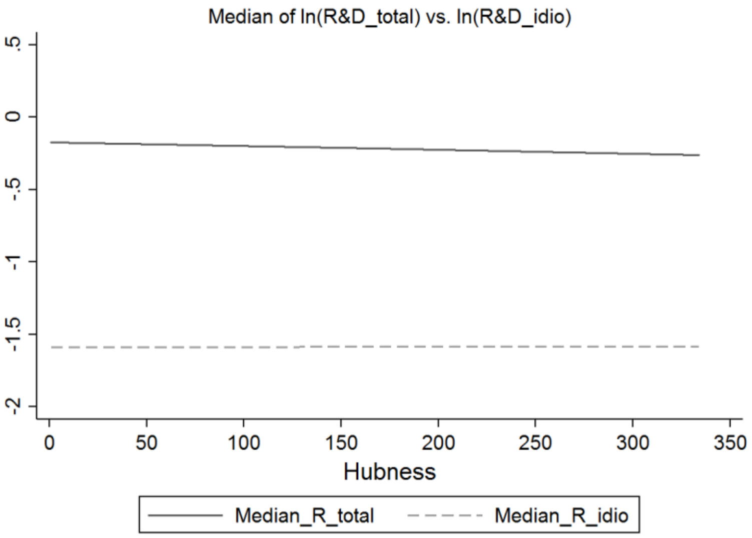


Figure 7. Trend of the median of log(R_{total}), and of the median of log(R_{idio}) as function of increasing network hubness.



Another important finding concerns the pattern of the standard deviation of the log(S) as function of the parameter of hubness. We expect a decreasing standard deviation of the distribution of the R&D support as long as the level of the hubness increases. This should be so, as an increase in the hubness should reduce the number of supported units to few companies. Table 5 shows in fact that this is the case: the beta regression coefficient is negatively significant with a value equal to -0.073, which is however not too large in size.

Table 5. Linear regression of the standard deviation of the log of *S* on the hubness parameter.

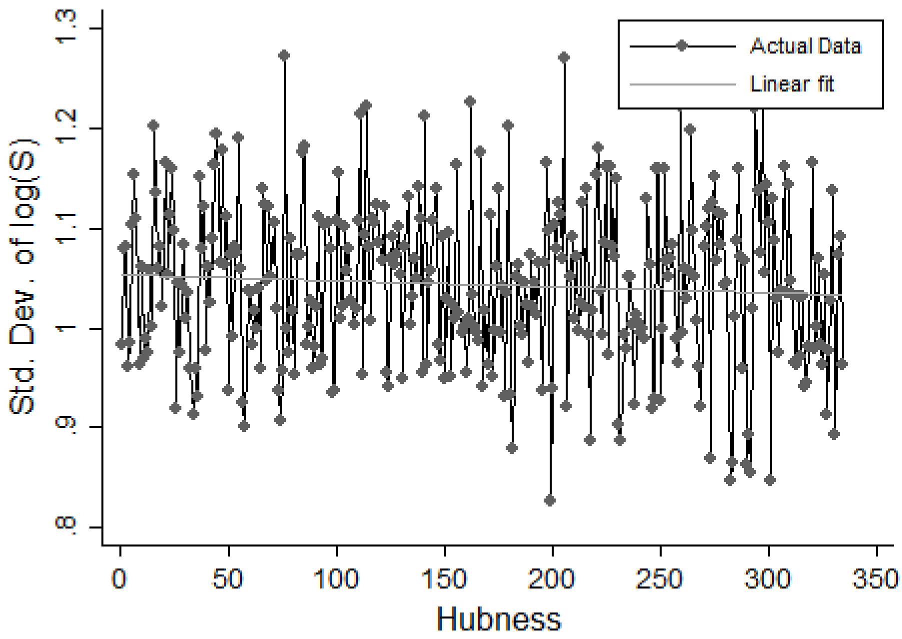
```
------------------------------------------------------------------------------
      Source |       SS       df       MS              Number of obs =  166999
-------------+------------------------------           F(  1,166997) =  906.01
       Model |  6.06357944     1  6.06357944           Prob > F      =  0.0000
    Residual |  1117.64884166997  .006692628           R-squared     =  0.0054
-------------+------------------------------           Adj R-squared =  0.0054
       Total |  1123.71242166998  .006728897           Root MSE      =  .08181

------------------------------------------------------------------------------
     sd_ln_s |      Coef.   Std. Err.      t    P>|t|                     Beta
-------------+----------------------------------------------------------------
   parametro |  -.0000625   2.08e-06   -30.10   0.000                -.0734576
       _cons |    1.05357   .0004013  2625.52   0.000                        .
------------------------------------------------------------------------------
```

Figure 8 confirms this finding although it also sets out a large variability around the linear fit.

Finally, the degree-centrality index increases – as expected – with the hubness, with a significant beta regression coefficient of 0.073.

Figure 8. Linear pattern of the standard deviation of the log of *S* on the hubness parameter.



5. Conclusions

This paper has presented R&Dsimulab, a micro-policy simulator for an ex-ante assessment of public R&D policy effect on companies' R&D activity. We have shown the behavioural structure and the computational logic of this model, and then illustrated a simulative example where the total level of R&D activated by a fixed amount of public support becomes function of companies' network topology. More specifically, the experiment here presented shows that a large "hubness" of the network is – on average – accompanied with a decreasing median of the aggregated total R&D. Since the aggregated firm idiosyncratic R&D (the part of total R&D independent of spillovers) is slightly increasing, we conclude that positive cross-firm spillover effects in the presence of a given amount of support, have a larger impact in less centralized network where fewer hubs emerge. This may question the common wisdom which suggests that larger R&D externality effects are more likely to arise when few "central" champions receive a support.

References

David P A and Hall B H (2000) Heart of Darkness: Modeling Public-Private Funding Interactions inside the R&D Black Box. Research Policy, 29, pp. 1165-1183.

David P A, Hall B H and Toole A A (2000) Is Public R&D a Complement or Substitute for Private R&D? A Review of the Econometric Evidence. Research Policy, 29, pp. 497-529.

Laincz C A (2009) R&D subsidies in a model of growth with dynamic market structure. Journal of Evolutionary Economics, 19, 5, pp. 643-673.

Jou J and Lee T (2001) R&D investment decision and optimal subsidy. R&D Management, 31, pp. 137-148.

Cerulli G (2012) Are R&D Subsidies Provided Optimally? Evidence from a Simulated Agency-Firm Stochastic Dynamic Game, Journal of Artificial Societies and Social Simulation, 15, 1, 7.

Howe J D and McFetridge D G (1976) The Determinants of R&D Expenditures. Canadian Journal of Economics, 9, pp. 57-71.