Collaboration Networks and Innovation Results in Spain

Pablo Galaso
Instituto de Economía – Universidad de la República

Jaromir Kovarik
Universidad del País Vasco

Montevideo, July 2014

Motivation

- Private companies, universities and research centres tend to share information and resources in R&D cooperative projects.

- Collaboration strategies may determine the success of individual agents and territories (Allen, 1983; Saxenian, 1994; Brusco, 1999).

- Previous research used social network analysis to successfully measure the structure of collaboration and estimate its influence on agents' results (e.g. Singh, 2005; Schilling and Phelps, 2007; Uzzi, 2008).

Therefore:

- Companies may take care of their collaborative activities (the number of links they trace and maintain, the partners they choose, the type and extent of knowledge they share...)

- Policymakers should also contemplate collaboration networks in fields such as innovation activities.

Reseach Questions

- Networks can include actors from different territories (different cities, regions or even countries)
- In such case, collaboration patterns may differ based on geographical differences: cultural aspects, institutional issues, face to face vs telephone/email interactions…

How relevant are these geographical differences?

Do they shape the impact of collaboration on companies results?

Do regional networks influence companies results in a different way than national or international networks do?

Research Goals

The aim of the paper is to analize the cooperation networks of innovative companies in Spain:

- Describing their structural properties, evolution and geographical differences
- Estimating their influence on innovation results
- Identifying –if any– the differences of this influence that are motivated by geographical aspects

Do do so:

- We will estimate the impact of Spanish national network on its members' R&D outputs
- Then we will estimate the same impact using regional networks in Spain

Collaboration Networks

Applying social network analysis, several studies have reached to identify network properties that can influence R&D results:

1. Connectivity and closeness: improve information access, making it easier and more reliable, as more links imply more sources of knowledge and fewer intermediaries (Fritsch and Kauffeld-Monz, 2008; Burt, 2000; Schilling and Phelps, 2007).

2. Clustering: reciprocal ties facilitate the diffusion of complex and tacit knowledge (Monge et al., 2008; Fleming et al., 2007), creates a system of self-regulation that reduces opportunistic actions increasing trust (Ahuja, 2000; Schilling and Phelps, 2007; Cowan and Jonard, 2008) and alters individual incentives, moving separate preferences towards general targets shared by the group (Uzzi and Spiro, 2005)

3. Decentralization: separates non-redundant sources of information (Burt, 2000), increasing the diffusion of new ideas (Stone, 2003; Schilling and Phelps, 2007; Monge et al., 2008).

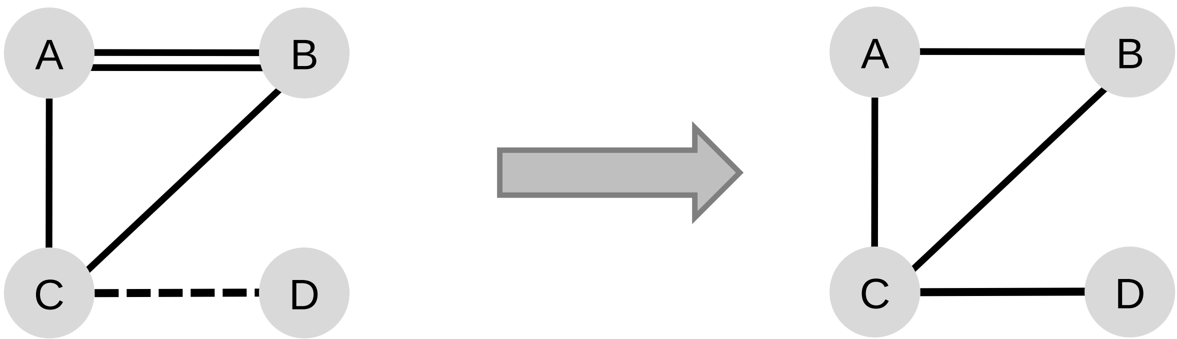
4. Small world: increases clustering and closeness advantages (Uzzi and Spiro, 2005; Schilling and Phelps, 2007; Uzzi, 2008).

Data

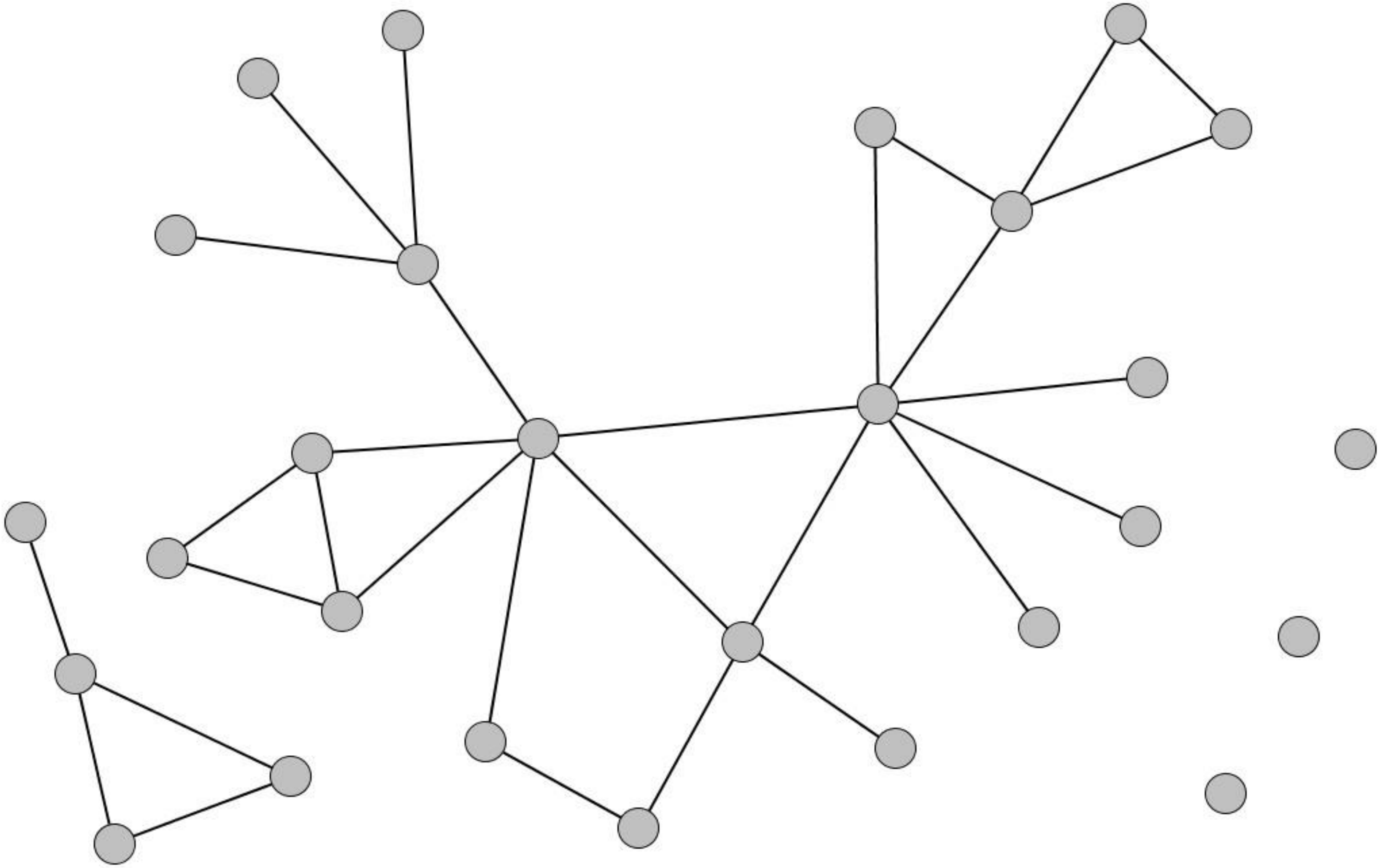
- Our source is the Spanish Patent Office (OEPM)
- In particular, we use all European patents presented in the Spanish Office from 1978 to 2008
- We construct a detailed database to identify, from each patent:
  1. Date of the application
  2. Names of the companies which have applied for the patent
  3. Names of the inventors who have worked on it
  4. The locations of both the companies and the inventors of the patent (the postal code)

Using Patent Data to Analyze Collaboration

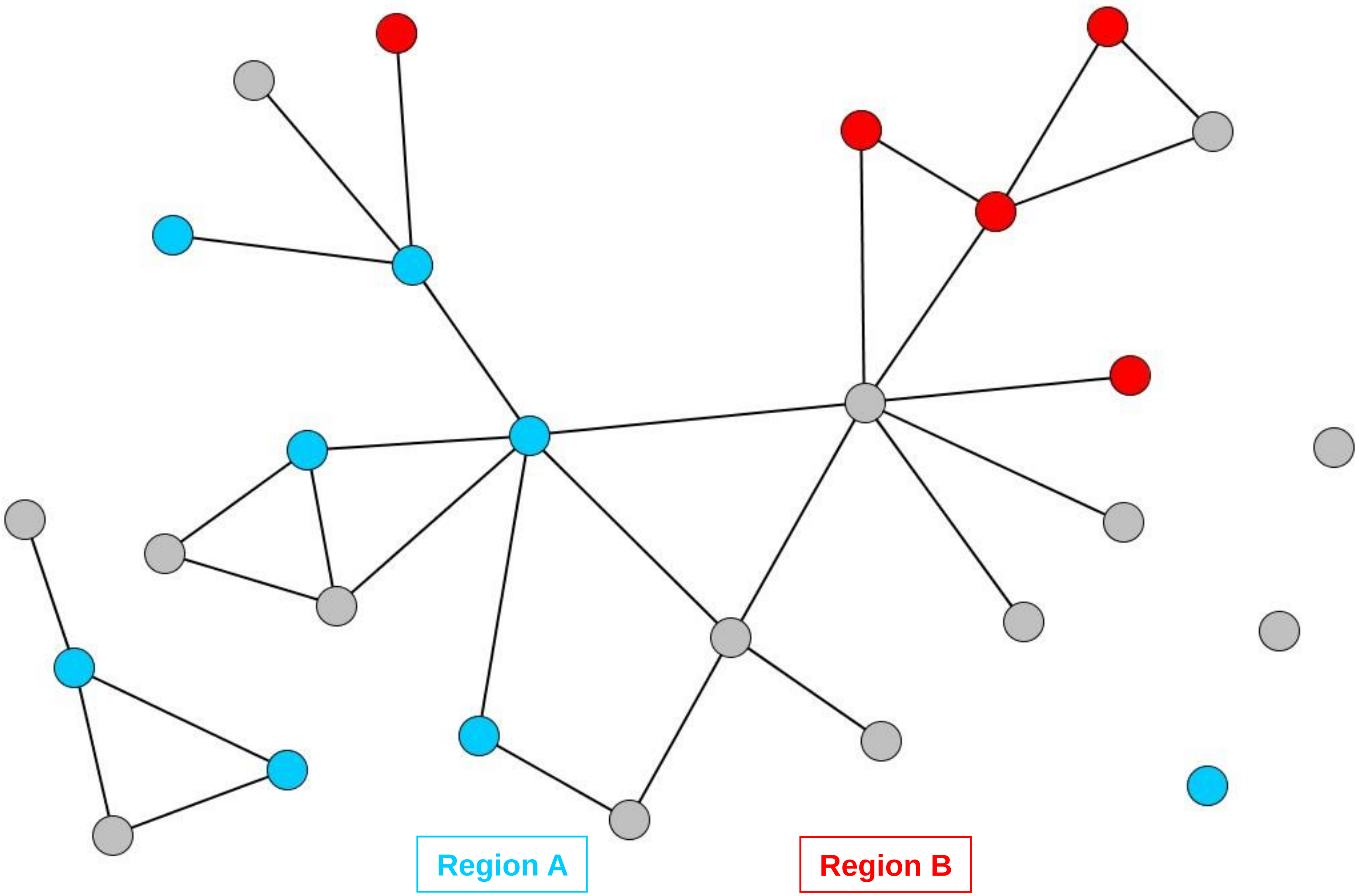
Patents	Owners	Inventors
1	A, B	X
2	A, B, C	Y
3	C	Z
4	D	Z



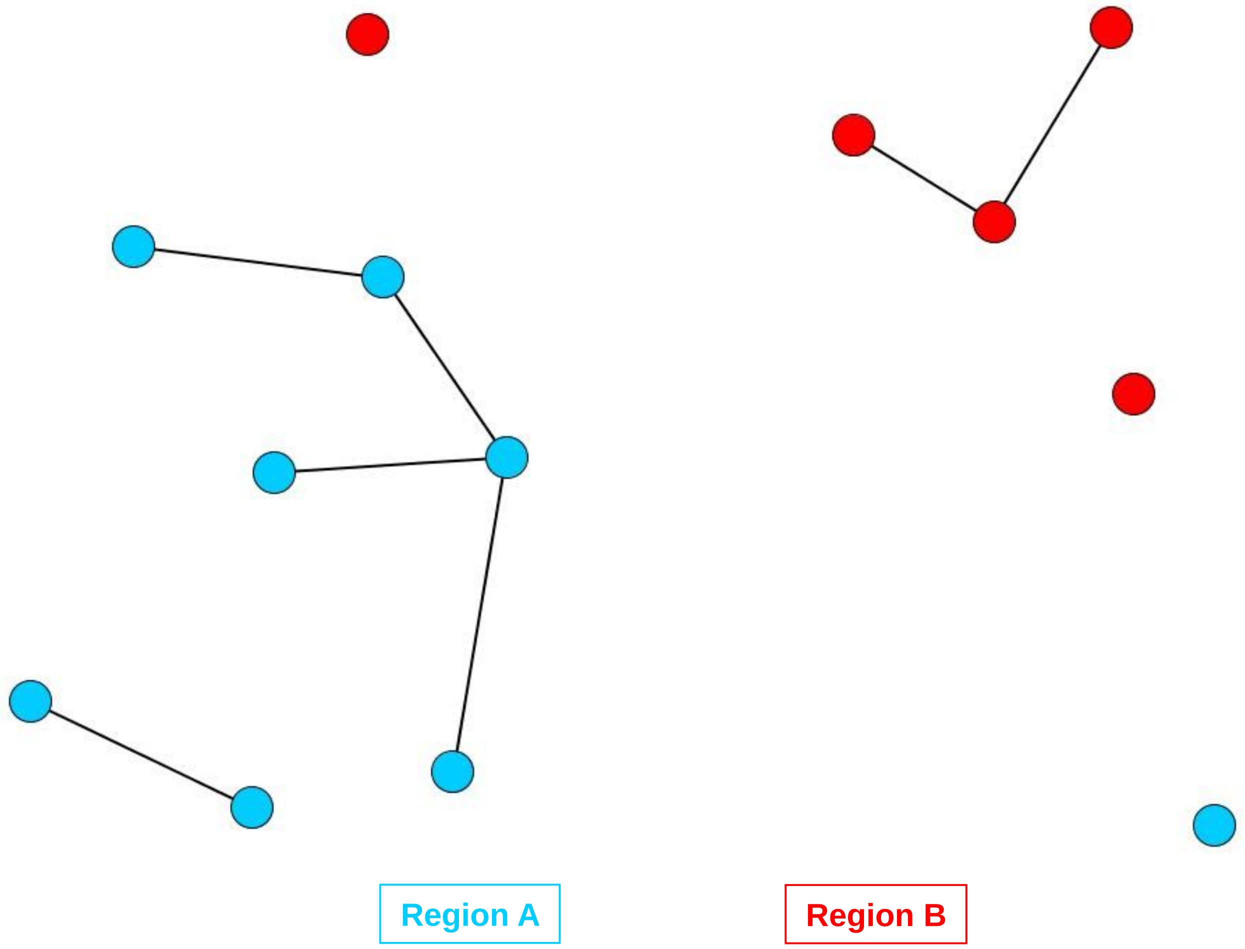
Geographical Perspective: National Network



Geographical Perspective: Regional Network



Geographical Perspective: Regional Network II



Network Variables

Variable	Description / Interpretation
Betweenness Centrality	Number of shortest paths from all nodes to all others that pass through a given node. Measures a firm's access to information. But maintaining links implies also costs.
Density	Share of actual links over the number of total possible links in a given network. Can increase the information diffusion but also the homogeneity.
In Giant	Indicates whether a node is part of the giant component. Measures the possibility of a firm to have direct or indirect connections to the largest group of innovative companies.
Giant Share	Share of total nodes that belong to the giant component of the network May reflect information spillovers.
Giant Size	Number of nodes included in the giant component. May reflect information spillovers.
Centralization	Calculates the similarity of a given network to a star-shaped network with the same number of nodes. Might lead to homogeneity of the information diffused.
Clustering	Degree to which the network contains groups of nodes highly connected. It can accelerates the circulation of trustworthy information and foster collaboration.
Reach	Measures how far all the nodes are from each other. Represents the level to which a network is expanded or tighter. It can help to have an easier access to diverse information.
Small World	Multiplication of the average clustering and the average reach. Makes the existence of both characteristics more valuable.

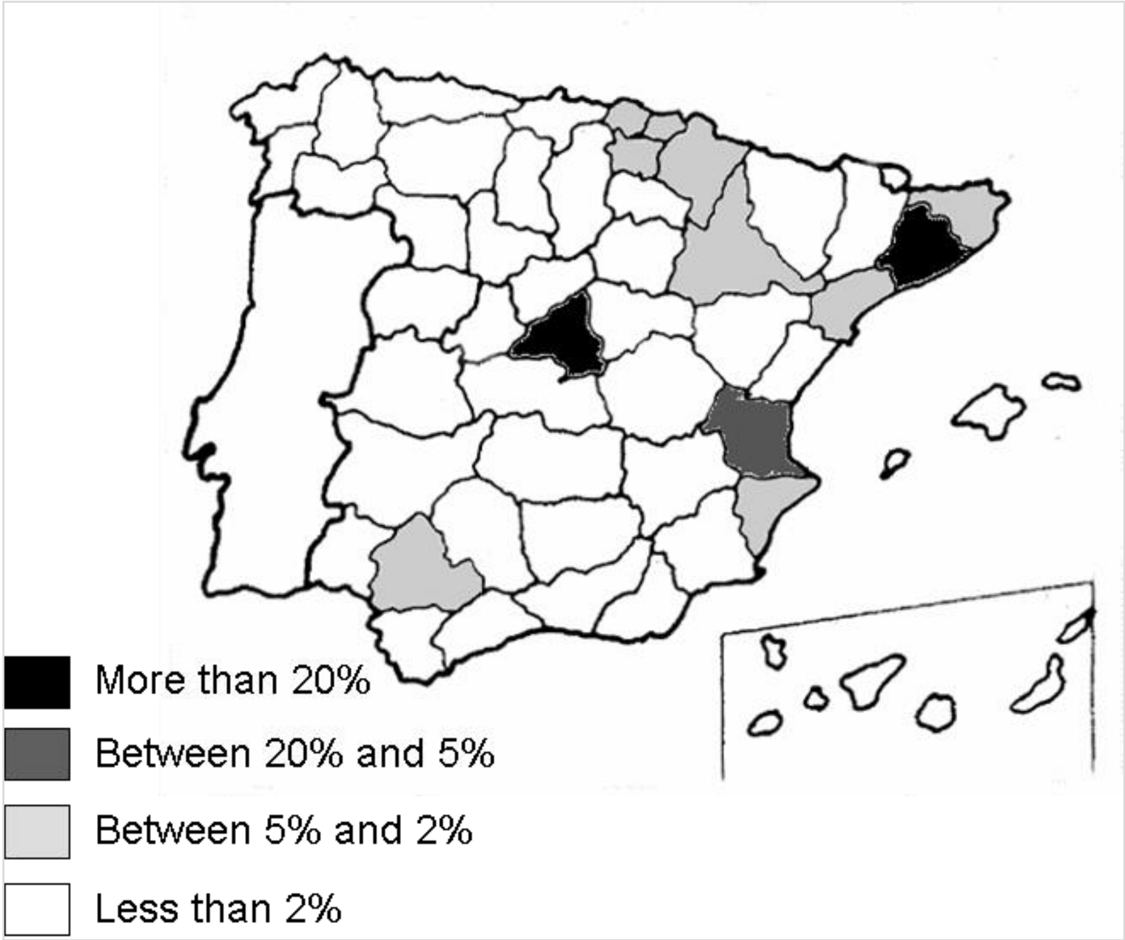
Models and Variables

Independent Variables	Individual and global network properties: Betweenness, Density, In Giant, Giant Share, Giant Size, Centralization, Clustering, Reach, Small World
Dependent Variable	Number of Patents in the next period
Control Variables	Patents, Degree, City, Period
Model Specification	Negative Binomial • Appropriate to model count data • Allows for overdispersion of the variance in the dependent variable (Hausman et al. 1984) • Used by previous literature (Schilling and Phelps, 2007; Fleming et al., 2007; Whittington et al., 2009)

$$Next\ Patents_{it} = f(Betweenness_t, Density_t, Centralisation_t, Clustering_t, Reach_t, Small\ World_t, Giant\ Share_t, Giant\ Size_t, In\ Giant_t, Patents_t, Degree_t, City, Period_t)$$

Territorial Distribution of Patent Production in Spain (1978-2008)

(% share of total patents registered)

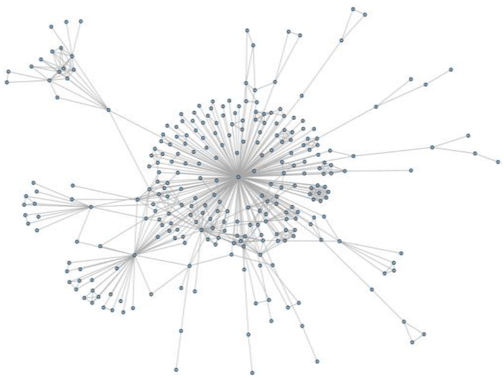


Spanish Innovation Network (Giant Component)

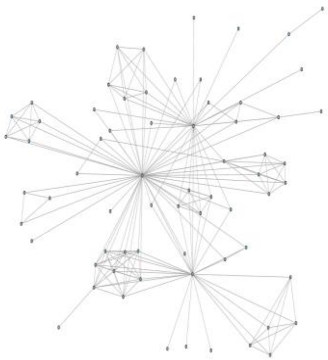
Spanish Main Regional Networks (Giant Components)

Barcelona

Madrid



Valencia



Barcelona

- *Connectivity*: equally distributed
- *Decentralisation*: less centralised
- *Clustering*: high levels
- *Closness*: expanded
- *Small world*: medium levels

Madrid

- *Connectivity*: concentrated
- *Decentralisation*: star-shaped
- *Clustering*: lowest levels
- *Closness*: short distances
- *Small world*: lowest levels

Valencia

- *Connectivity*: concentrated
- *Decentralisation*: star-shaped
- *Clustering*: highest levels
- *Closness*: shortest distances
- *Small world*: highest levels

Basic Network Properties

		Spain	Barcelona	Madrid	Valencia
Network size	Nodes	8,215	2,459	1,614	604
	Links	5,475	1,558	1,114	458
Density (%)		0.02	0.05	0.09	0.25
Degree	Av.	1.33	1.27	1.38	1.51
	St. Dev.	4.44	2.33	5.25	3.24
Giant component	Size	852	208	278	67
	% of total	*10.37*	*8.46*	*17.22*	*11.09*
Second largest	Size	17	16	10	11
	% of total	*0.21*	*0.65*	*0.62*	*1.82*
Isolates	Number	4,139	1,203	811	300
	% of total	*50.38*	*48.92*	*50.25*	*49.67*
Diameter		11	12	9	5

Estimations

	Model 1		Model 2		Model 3		Model 4	
	Regional	National	Regional	National	Regional	National	Regional	National
	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
Independent Variables								
Betweenness	-0,458 ***	-0,001 ***	-0,005 ***	-0,001 ***	-0,004 ***	-0,001 ***	-0,004 ***	-0,001 ***
Density	-3,943	-51,225	82,483	-4441,040 *	-24,362	-346,414	10,132	-4410,416 *
Centralization	-7,551 **	0,705	-0,824	-176,684 *	-22,283	-103,253 *	-8,672	-176,882 *
Clustering	-	-	2,150	11,840 *	-	-	2,646	11,375 *
Reach	-	-	-117,275	5038,596 *	-	-	-162,649	5017,437 *
Small World	-	-	-	-	-	-	188,504	ommited
Giant Share	-	-	-	-	0,007	-0,016	15,661	ommited
Giant Size	-	-	-	-	15,814	117,087	-0,007	ommited
In Giant	-	-	-	-	0,499 ***	0,588 ***	0,498 ***	0,588 ***
Control Variables								
Patents	0,044 ***	0,056 ***	0,047 ***	0,056 ***	0,040 ***	0,052 ***	0,041 ***	0,052 ***
City1	0,377 *	0,063	0,449 *	0,063	0,215	0,020	0,323	0,020
City2	0,078	-0,108	0,047	-0,108	-0,022	-0,085	-0,001	-0,085
Period 2	-0,469	-0,200	-0,555	ommited	-0,183	ommited	-1,035	ommited
Period 3	-0,459	-0,242 *	-0,677	ommited	-0,096	ommited	-1,211	ommited
Period 4	-0,192	ommited	-0,385	ommited	0,088	ommited	-0,961	ommited
Period 5	0,081	ommited	-0,247	ommited	0,031	ommited	-1,017	ommited
Period 6	0	ommited	0	ommited	0	ommited	0	ommited
Constant	1,284 **	1,038 ***	1,241 *	0,415	0,904	0,679	1,802	0,480

* p < 0,10
** p < 0,05
*** p < 0,01

Estimations Sumary

N. observations

385
(active nodes in, at least, two consecutive periods)

Regional			National		
Variable	**Coef.**	**p**	**Variable**	**Coef.**	**p**
Betweenness	(-)	***	Betweenness	(-)	***
In giant	(+)	***	In giant	(+)	***
Centralization	*(-)*	**	*Centralization*	*(-)*	*
			Density	(-)	*
			Clustering	(+)	*
			Reach	(+)	*

Conclusions

Main contributions

- Study of cooperation relations among innovative companies in Spain using patent data
- Ellaboration and analysis of Spanish innovation networks (1978-2008)
- Estimation of the impact of collaboration networks on innovators' outputs
- Identification of territorial differences on this impact

Reseach and Policy Implications

- Geographical aspects might be considered when studying collaboration networks
- Companies may expect different outputs when collaborating with partners from the same region than when they do it with firms from other regions
- Firms may follow different strategies: national vs regional collaboration
- Governments may apply different policies for national and regional innovation systems