The determinants of network formation: a dynamic perspective

Work in progress ; all comments are wellcome ☺

Lorenzo Cassi, CES, Université Paris 1
&
Anne Plunket, ADIS, Université Paris 11

Eurolio, Toulouse, 11 juin 2010

Localized Knowledge Spillovers

Geography vs Networks

- R&D, knowledge spillovers and economic growth
  - Firm and industry productivity is related to R&D spending and knowledge spillovers
- Knowledge spillovers are localized (Jaffe, Trajtenberg & Henderson, 1993) (using citations)
  - Emphasize spatial proximity
  - Industrial clusters : innovation is more concentrated than production (Audretsch & Feldman, 1996 ; Porter, 1990)
- LKS = black box : channels of knowledge diffusion ?
  - Pure knowledge spillovers versus interpersonal links (social networks, labor mobility...)
  - To what extent are they really localized ? (role geography)

Social networks and proximities

- Breschi & Lissoni, 2003, 2009 (using citations)
  - LKS are explained by network relationships (labor mobility)
  - → Without interpersonal links no spillovers ; knowledge is not in the air …
  - Knowledge flows are localized to the extent that network ties also are
  - → Networks as knowledge diffusion channels
  - → Link Network structure to economic performance

Networks, knowledge diffusion and economic performance

- Social distance significant when dist.<3 (Singh, 2005)
- Small worlds : dense networks and highly interconnected actors increase knowledge creation and knowledge diffusion
  - Ex: Silicon Valley versus Boston (Fleming, King, Juda, (2007) & Powel, White, Koput, Owen_Smith, 2005)
- The role/position of academic inventors in knowledge diffusion

Less studied question : the dynamic of network formation ?

Network formation

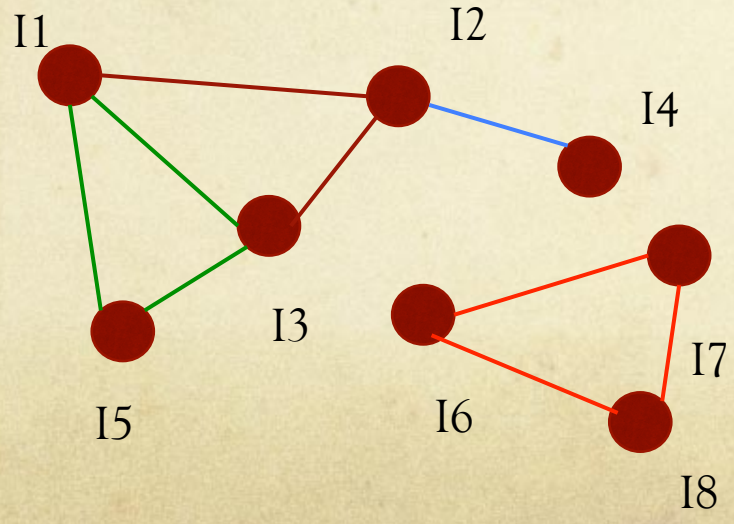
- **Cumulative mechanisms** related to historical process:
  - *Preferential attachment* : the tendency to link to the most connected actors
  - *Homophily* : the tendency to form a link with most similar actors – same organization, same technology, same region
  - *Closure* : the tendency to cluster
- **Selective mechanisms** : actors implement some strategies to gain benefit from the network
  - Actor choose to link with actors endowed with specific assets or resources (relevant knowledge)
  - Actors fill structural holes in order to benefit from other actors strategic position (control the flow of information)

Aim of the paper and methodology

- Study the determinants of co-inventor network formation
  - Consider the impact of network structure and inventor's position
  - Investigate the strategic role of applicants in the network formation and knowledge flows
- Methodology : case-control design and econometric analysis

Co-inventor networks

- **Patent data** :
  - Inventors' name and location
  - Applicants name (location cannot be used)
  - Year, IPC classes (technology), citations
  - *John Smith* problem
  - The network is limited to patent collaborations, that is only a fraction of all relevant network ties

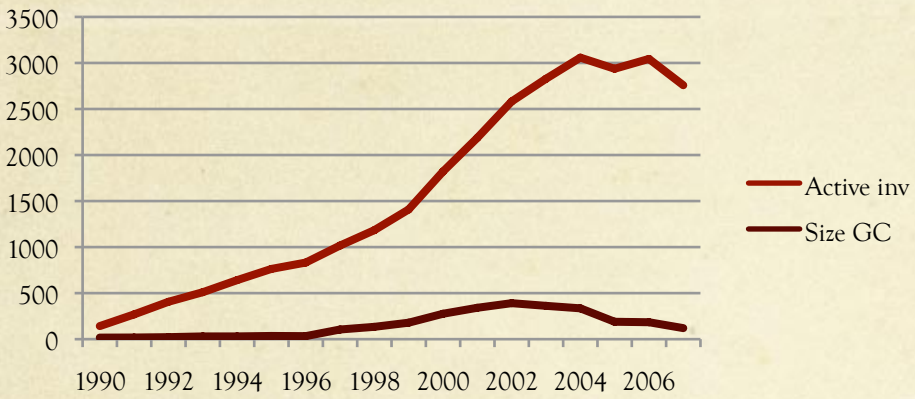


- **B1** : I1-I2-I3
- **B2** : I2-I4
- **B3** : I1-I3-I5
- **B4** : I6-I7-I8
- **Network** I1-I2-...-I8
- **Component 1** : I1-...-I5
- **Component 1 size** : 5 inventors
- **Component 2** : I6-I7-I8

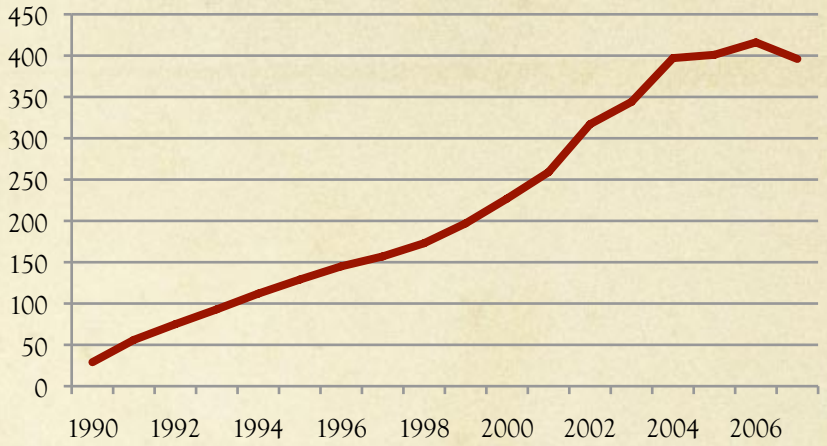
Patent networks in genomics 1990-2007

- Database PATSTAT
  - All EPO patents in genomics (1990-2007) – ANR Corpus genomic with OST-INRA-ADIS
  - All genomic patents with at least one inventor reporting a French address
    - 2401 patents
    - 5261 inventors
    - 662 applicants
- Network built using five-year windows
  - Network in 1994 is built on patents published between 1990 and 1994

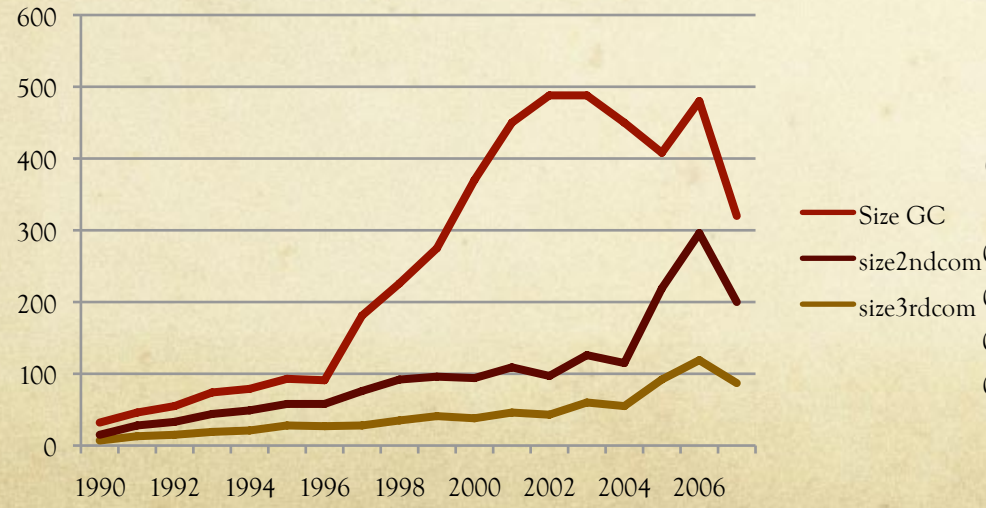
Active network and Giant component (size)



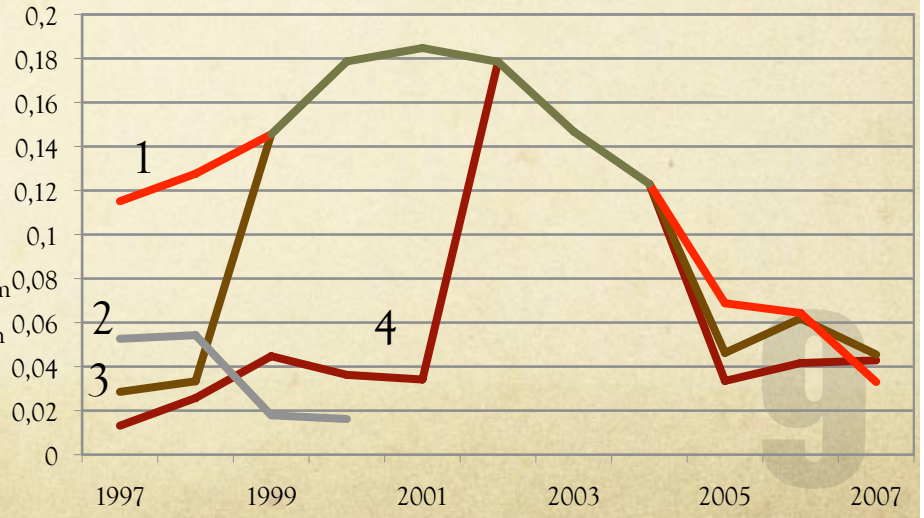
Number of components



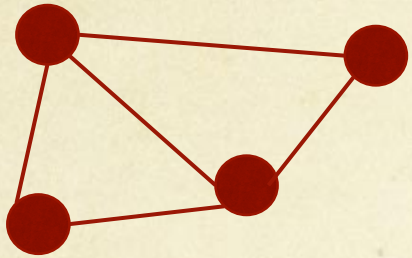
Component size

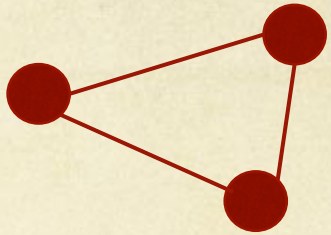


1998 first four component over time



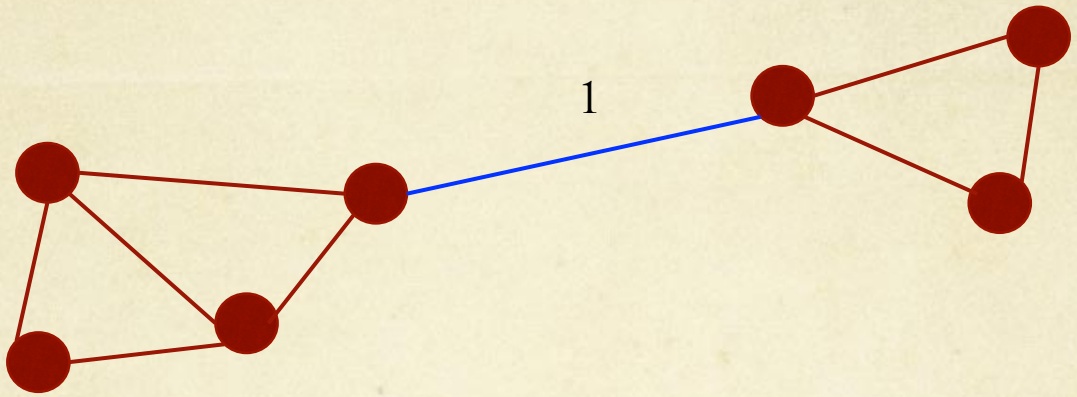
Type of network ties





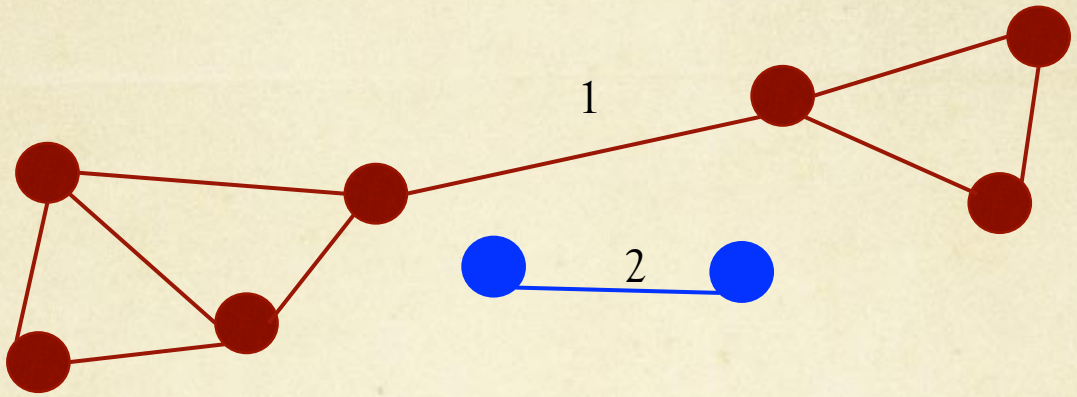
Links	Total number	%
Total	14779	100

Type of network ties



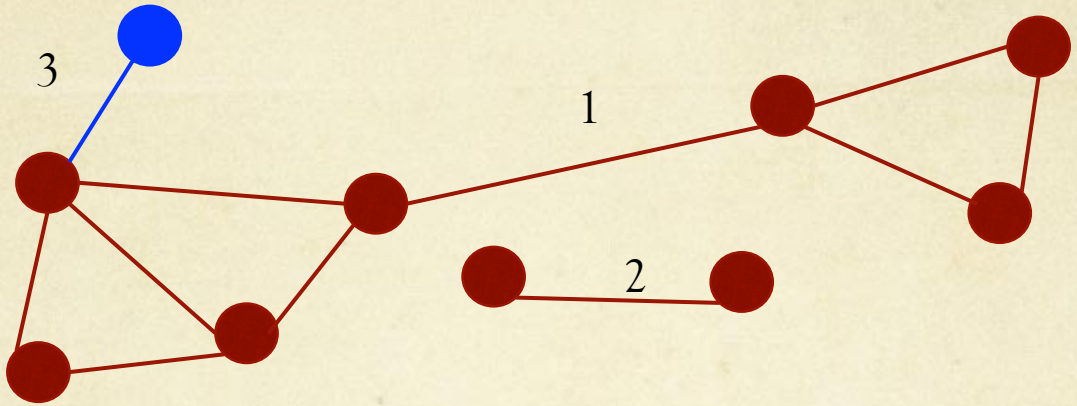
Links	Total number	%
1. Bridging links	265	1,79
Total	14779	100

Type of network ties



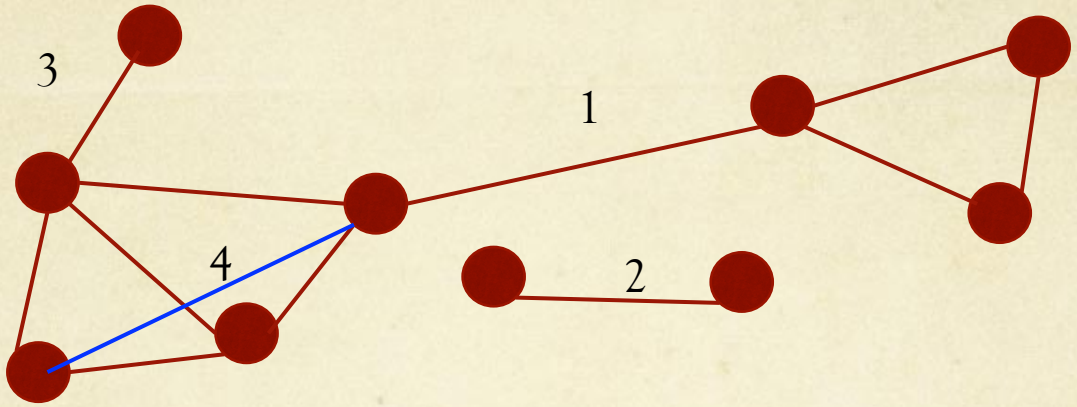
Links	Total number	%
1. Bridging links	265	1,79
2. New Component link	10037	67,91
Total	14779	100

Type of network ties

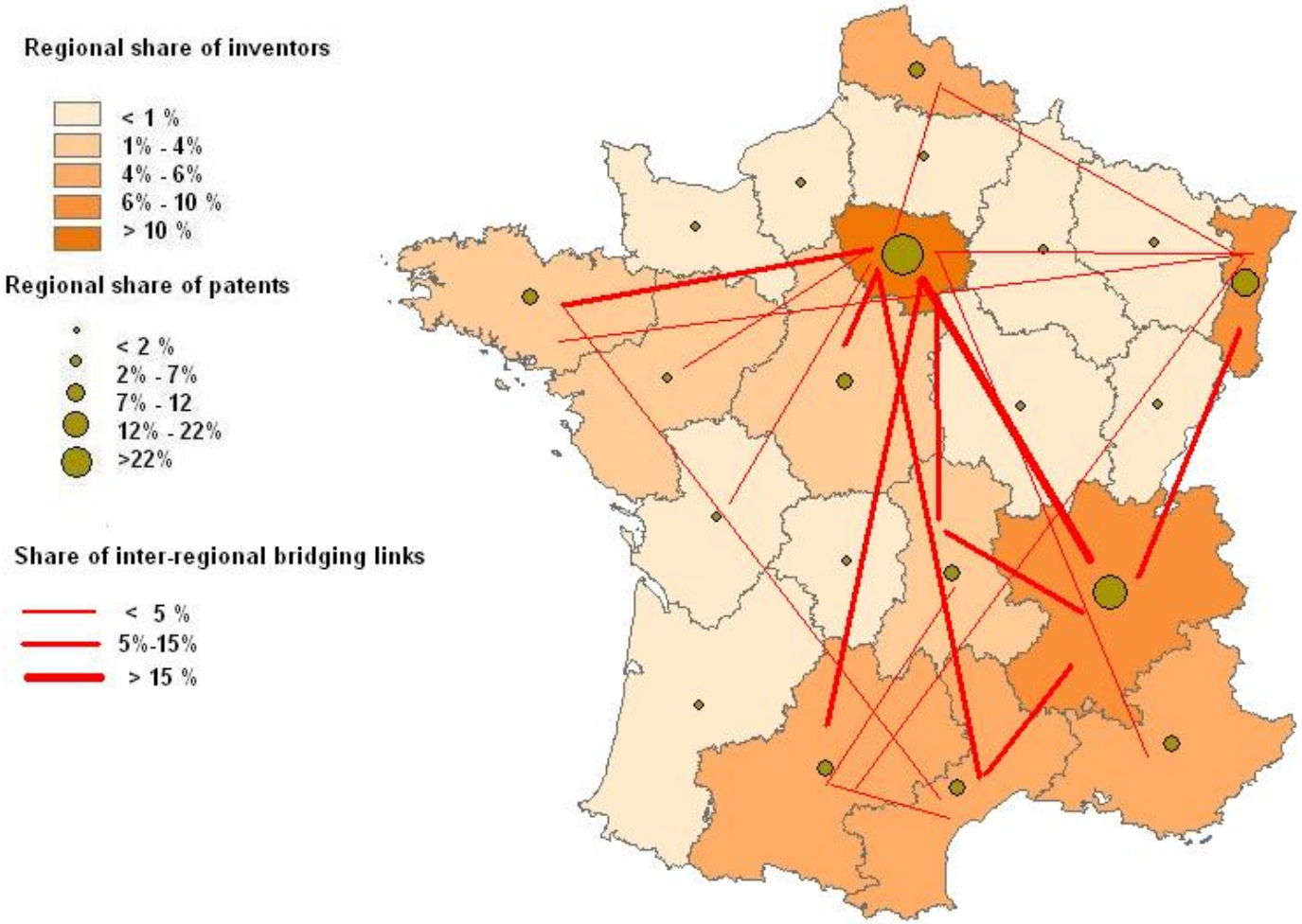


Links	Total number	%
1. Bridging links	265	1,79
2. New Component link	10037	67,91
3. Pendant links	4266	28,87
Total	14779	100

Type of network ties



Links	Total number	%
1. Bridging links	265	1,79
2. New Component link	10037	67,91
3. Pendant links	4266	28,87
4. Intra-component link	211	1,43
Total	14779	100


Regional share of inventors
< 1 %
1% - 4%
4% - 6%
6% - 10 %
> 10 %
Regional share of patents
< 2 %
2% - 7%
7% - 12
12% - 22%
>22%
Share of inter-regional bridging links
< 5 %
5%-15%
> 15 %

Organizational types		bridge	intracomp
withinacad	**25**	19	6
inter-org	**226**	176	50
betweenacad	25		
interfirm	201	152	49
betweenfirms	72	47	25
firm_acad	129	105	24
intra-firm	**224**	70	154
TOTAL	**475**	**265**	**210**

Testable Propositions and variables

- **Network effects and inventor's position**
  - The probability of forming a bridging or intra-component tie is increased when co-inventors have a higher degree centrality or belong to a larger component
  - → **Co-inventors' degree and component size (Fafchamps et al. (2005) ; Autant-Bernard et al. (2007) ;**

$$\Delta\ \text{degree}_{ij} = \left|\text{degree}_i - \text{degree}_j\right|$$

$$\overline{\text{degree}_{ij}} = (\text{degree}_i - \text{degree}_j)/2$$

$$\Delta\ \text{compsize}_{ij} = \left|\text{compsize}_i - \text{compsize}_j\right|$$

$$\overline{\text{compsize}_{ij}} = (\text{compsize}_i - \text{compsize}_j)/2$$

- **Geographical effect**
  - The probability of forming a bridging or intra-component tie is increased when co-inventors are geographically close
  - The impact of geographical proximity should decrease over time
  - → **Co-inventors' geographical proximity = 1 if same region 0 otherwise in t-1**
  - → **Foreign = 1 if the co-inventor is a foreigner in t-1**
  - → **Border =1 if a foreign inventor comes from a neighboring country in t-1**

Testable Propositions and variables

- **Technological effect**

The probability of a bridging or intra-component tie is decreased with co-inventors technological distance

→ **Co-inventors' technological distance** = Euclidian distance between each co-inventors patent portfolio in terms of IPC classifications (3digit) t-1

- The probability may decrease if co-inventors or applicants try to gain access to similar or complementary knowledge bases

- **Affiliation and "strategic effects"**
  - The probability of a bridging tie is increased when co-inventors belong to different organizations
  - The probability of an intra-component tie is increased when co-inventors belong to the same organizations.

→ Within **same type of organization** = 1 if between firms or between academics t-1

→ Within **same applicant** = 1 if the same firm or same research unit.

→ Between firm and academia

→ Between academia

Case-control design

- Unit of analysis : co-inventor dyad
- Controls : for each realized dyad we have 10 controls, 5 for the inventor 1 and 5 for the inventor 2

264 links and 2631 controls

- Estimation procedure : discrete-choice or *conditional* logit (McFadden, 1973) with robust standard errors
  - Regression for bridging ties
  - Control for largest regions

Determinants of network ties

	network b/se	Bridging b/se	Intracomp b/se	network b/se	Bridging b/se	Intracomp b/se	network b/se	Bridging b/se	Intracomp b/se
Degree (Diff)	-0.011	0.270*	-0.260**	0.252**	0.142	-0.075	0.408**	0.154	0.525+
	(0.074)	(0.115)	(0.095)	(0.088)	(0.113)	(0.193)	(0.126)	(0.151)	(0.287)
Degree (Avrg)	0.779***	0.112	1.591***	0.195	0.389*	1.148**	0.320	0.806**	0.320
	(0.121)	(0.171)	(0.179)	(0.146)	(0.192)	(0.390)	(0.213)	(0.277)	(0.515)
Component size (Diff)				-0.531***	0.337***	-25.048***	-0.346***	0.483***	-24.299***
				(0.033)	(0.080)	(0.500)	(0.037)	(0.096)	(0.778)
Component size (Avrg)				0.532***	-0.471***	4.220	0.206**	-0.834***	-13.637***
				(0.048)	(0.111)	(4.686)	(0.068)	(0.143)	(3.460)
Same region							1.744***	1.753***	1.432**
							(0.170)	(0.192)	(0.539)
Technological distance							-2.097***	-2.172***	-1.348
							(0.361)	(0.414)	(0.926)
Foreign							-1.015***	-1.720***	-0.174
							(0.205)	(0.324)	(0.590)
Border							3.294***	3.836***	2.089*
							(0.412)	(0.503)	(0.976)
Same Applicant							2.147***	1.908***	1.776***
							(0.163)	(0.213)	(0.408)
Observations	5229.00	2909.00	2320.00	5229.00	2909.00	2320.00	5218.00	2909.00	2309.00
Log Likelihood	-1125.65	-632.38	-481.87	-917.19	-625.90	-101.54	-592.12	-420.88	-67.41
LR Chi Square	108.20	31.99	100.38	327.78	45.81	2672.43	574.58	329.09	1481.78
p	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Pseudo R-Square	0.02	0.01	0.05	0.20	0.02	0.80	0.48	0.34	0.87

Determinants of network ties

	network b/se	Bridging b/se	Intracomp b/se	network b/se	Bridging b/se	Intracomp b/se
Degree (Diff)	0.334**	0.147	0.359	0.325**	0.142	0.359
	(0.108)	(0.134)	(0.263)	(0.109)	(0.135)	(0.263)
Degree (Avrg)	0.278	0.679**	0.632	0.304	0.696**	0.632
	(0.184)	(0.241)	(0.466)	(0.185)	(0.242)	(0.466)
Component size (Diff)	-0.455***	0.411***	-25.359***	-0.454***	0.401***	-25.352***
	(0.035)	(0.091)	(0.752)	(0.035)	(0.089)	(0.753)
Component size (Avrg)	0.358***	-0.714***	-12.623**	0.344***	-0.711***	-12.465**
	(0.063)	(0.136)	(4.547)	(0.064)	(0.135)	(4.412)
Same region	2.163***	2.001***	2.057***	2.179***	2.020***	2.057***
	(0.162)	(0.185)	(0.523)	(0.162)	(0.186)	(0.523)
Technological distance	-2.086***	-2.195***	-1.890*	-2.114***	-2.203***	-1.890*
	(0.319)	(0.378)	(0.950)	(0.319)	(0.377)	(0.950)
Foreign	-0.786***	-1.712***	0.237	-0.796***	-1.712***	0.237
	(0.199)	(0.312)	(0.670)	(0.200)	(0.312)	(0.670)
Border	3.272***	3.970***	1.811*	3.307***	3.988***	1.811*
	(0.391)	(0.452)	(0.900)	(0.391)	(0.443)	(0.900)
Same Applicant						
Same Organizational type	0.626***	0.402**	0.404			
	(0.129)	(0.151)	(0.429)			
Between firm and academ				-0.670***	-0.422**	-0.404
				(0.134)	(0.158)	(0.429)
Between academia				-0.900***	-0.664**	-0.874
				(0.259)	(0.249)	(1.227)
Observations	5218.00	2909.00	2309.00	5218.00	2909.00	2309.00
Log Likelihood	-684.97	-464.87	-81.14	-683.29	-464.03	-81.14
LR Chi Square	528.32	265.89	2323.19	537.01	273.88	2624.60
p	0.00	0.00	0.00	0.00	0.00	0.00
Pseudo R-Square	0.40	0.27	0.84	0.40	0.27	0.84

Determinants of inter-organizational ties

	network b/se	Bridging b/se	Intracomp b/se	network b/se	Bridging b/se	Intracomp b/se
Degree (Diff)	0.054	0.083	0.146	-0.021	0.045	0.099
	(0.106)	(0.143)	(0.176)	(0.123)	(0.169)	(0.215)
Degree (Avrg)	0.218	0.116	0.341	0.337	0.105	0.407
	(0.247)	(0.309)	(0.386)	(0.281)	(0.378)	(0.440)
Component size (Diff)	0.446***	0.136*	0.617***	0.531***	0.214*	0.684***
	(0.036)	(0.068)	(0.067)	(0.047)	(0.089)	(0.079)
Component size (Avrg)	-0.600***	-0.259*	-0.530	-0.721***	-0.348*	-0.549
	(0.082)	(0.115)	(0.360)	(0.103)	(0.150)	(0.441)
Same region	-1.681***	-1.431***	-1.980***	-1.970***	-1.679***	-2.251***
	(0.174)	(0.201)	(0.360)	(0.219)	(0.255)	(0.388)
Technological distance	1.286***	1.046*	1.886**	1.707***	1.436**	2.371**
	(0.334)	(0.423)	(0.639)	(0.387)	(0.491)	(0.753)
Foreign	0.313	0.579*	-0.101	0.088	0.342	-0.304
	(0.199)	(0.264)	(0.371)	(0.225)	(0.292)	(0.409)
Border	-1.058**	-1.782***	0.256	-1.049*	-2.441***	0.513
	(0.384)	(0.518)	(0.732)	(0.463)	(0.608)	(0.878)
Academia	2.107***	1.435***	4.195***			
	(0.212)	(0.212)	(0.812)			
Same Organizational type				-21.293***	-18.836***	-19.634***
				(0.433)	(0.287)	(0.560)
Observations	3784.00	1756.00	2028.00	3784.00	1756.00	2028.00
Log Likelihood	-628.81	-398.65	-196.22	-419.03	-249.49	-154.59
LR Chi Square	359.39	139.02	199.14	2661.73	5491.79	2013.23
p	0.00	0.00	0.00	0.00	0.00	0.00
Pseudo R-Square	0.41	0.18	0.66	0.60	0.49	0.73

Determinants of inter-regional ties

	network	Bridging	Intracomp	network	Bridging	Intracomp
	b/se	b/se	b/se	b/se	b/se	b/se
interregio						
Degree (Diff)	-0.019	0.004	0.036	-0.016	0.014	0.033
	(0.069)	(0.090)	(0.107)	(0.069)	(0.090)	(0.108)
Degree (Avrg)	0.104	0.005	0.095	0.104	-0.007	0.098
	(0.132)	(0.175)	(0.202)	(0.132)	(0.176)	(0.203)
Component size (Diff)	0.184***	-0.019	0.257***	0.185***	-0.018	0.260***
	(0.022)	(0.045)	(0.031)	(0.022)	(0.044)	(0.031)
Component size (Avrg)	-0.273***	-0.020	-0.213	-0.284***	-0.027	-0.228
	(0.053)	(0.076)	(0.158)	(0.054)	(0.076)	(0.158)
Technological distance	0.784***	0.805**	0.604+	0.790***	0.824**	0.600+
	(0.216)	(0.307)	(0.325)	(0.216)	(0.305)	(0.323)
Foreign	2.744***	3.314***	2.363***	2.742***	3.313***	2.360***
	(0.174)	(0.282)	(0.223)	(0.173)	(0.281)	(0.222)
Border	-1.700***	-2.317***	-1.071**	-1.701***	-2.323***	-1.075**
	(0.294)	(0.502)	(0.401)	(0.299)	(0.510)	(0.402)
Same Organizational type	-0.356***	-0.388***	-0.254*			
	(0.081)	(0.106)	(0.128)			
Academ				0.343***	0.409**	0.215
				(0.091)	(0.128)	(0.132)
Firm				0.314*	0.296*	0.318
				(0.129)	(0.145)	(0.321)
Observations	4943.00	2678.00	2265.00	4943.00	2678.00	2265.00
Log Likelihood	-1663.79	-898.23	-747.24	-1664.39	-898.34	-747.55
LR Chi Square	317.28	181.53	181.36	321.29	179.55	182.58
p	0.00	0.00	0.00	0.00	0.00	0.00
Pseudo R-Square	0.20	0.19	0.22	0.20	0.19	0.22

Discussion

- Role of technological and geographical proximity within networks
- Network growth through new ties
- Higher probability to bridge within the same company or research unit : very conservative behaviors
- What about bridging structural holes ?
- Control for large applicants
- Case-control : control for geography and for technology

	Region Code	#	%
ILE-DE-FRANCE	11	158	46.33
CENTRE	24	7	2.05
NORD-PAS-DE-CALAIS	31	5	1.47
ALSACE	42	24	7.04
PAYS-DE-LA-LOIRE	52	1	0.29
BRETAGNE	53	9	2.64
POITOU-CHARENTE	54	2	0.59
MIDI-PYRENEES	73	16	4.69
RHONE-ALPES	82	62	18.18
AUVERGNE	83	11	3.23
LANGUEDOC-ROUSSILLON	91	12	3.52
PROVENCE-ALPES-COTE-D'AZUR	93	1	0.29
Etrangers	99	33	9.68