Groupe de Recherche en
Économie Théorique et Appliquée

The French Aerospace Sector Collaboration Network : Structural Dynamics And Firm Performance

Johannes VAN DER POL

GREThA, CNRS, UMR 5113, Université de Bordeaux

johannes.van-der-pol@u-bordeaux.fr

Cahiers du GREThA
n° 2016-24
Octobre

GRETHA UMR CNRS 5113
Université de Bordeaux
Avenue Léon Duguit - 33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75 - Fax : +33 (0)5.56.84.86.47 - www.gretha.fr

Le réseau de collaboration du secteur aéronautique en France : Dynamique structurelle et performance des firmes

Résumé

L'objectif de ce papier est d'analyser le lien entre la position de la firme dans son réseau de collaboration et sa performance. Sous l'hypothèse que les firmes ont accès à des connaissances diverses par le biais de leur réseau, on analyse comment les firmes choisissent leurs collaborateurs et comment les connaissances peuvent diffuser dans le réseau qui résulte de ces décisions.

Dans un premier temps la structure du réseau est analysé à trois niveaux : la structure globale, les clusters et la position des firmes individuelles. Le dernier point est accompli en utilisant un Exponential Random Graph Model (ERGM). Dans un second temps, le lien entre la position de la firme dans le réseau et sa performance est analysée.

Mots-clés : Analyse réseau ; Réseau d'innovation ; ERGM ; Performance ; Small World ; Scale-free

The French Aerospace Sector Collaboration Network : Structural Dynamics And Firm Performance

Abstract

The focus of this paper is on the link between network structure and the financial performance of the individual firm. Under the hypothesis that firms access diverse and valuable knowledge through collaboration we analyse how firms pick their collaborators and how knowledge flows impact the financial performance of the firm.

First, the evolution of the structure of the collaboration network of the French aerospace sector is analysed between 1980 and 2013. The global structure is identified and, using an ERGM and clustering identification, the structure of the network is explained. Second, a panel regression identifies a link between the position of the individual firm inside the network and their financial performance.

Keywords: Network analysis ; Innovation network ; ERGM ; Performance ; Small World ; Scale-free

JEL: L25 ; C23 ; D85 ; L14 ; C20

Reference to this paper: VAN DER POL Johannes (2016) The French Aerospace Sector Collaboration Network : Structural Dynamics And Firm Performance, *Cahiers du GREThA*, n°2016-24.

http://ideas.repec.org/p/grt/wpegrt/2016-24.html.

GRETHA UMR CNRS 5113
Université de Bordeaux
Avenue Léon Duguit - 33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75 - Fax : +33 (0)5.56.84.86.47 - www.gretha.fr

1 Introduction

The technological landscape expands continuously. With the increasing diversity in technologies it becomes difficult for one firm to master all the technologies needed for innovation. Firms hence seek access knowledge through external sources, through collaboration (Pyka, 2002).

The aggregation of collaboration for a given criteria (sector, technology, region) results in a collaboration network. Within this network firms recombine their knowledge, resulting in an exchange of knowledge and ideas. The speed and efficiency with which the knowledge diffuses throughout the network depends upon its structure (Verspagen, Duysters, 2004; Cowan, Jonard, 2004). Networks with a large average path length will require more time for complete diffusion of knowledge than networks with a low average path length for example[1]. In order to characterize network structures, research has aimed at identifying different network structure and analyze their characteristics. In this light, the small world structure has been identified as being the most efficient structure (Verspagen, Duysters, 2004; Watts, 1999b; Morone, Taylor, 2004; Alghamdi et al., 2012). This observation is however the result of mostly theoretical work (Cowan, Jonard, 2004; Baum et al., 2003) even though the small world structure has been observed empirically (Ahuja, 2000). Other canonical structures, such a core-periphery structures (Barabási, Albert, 1999) and nested-split graphs have also been identified (König et al., 2009). Each structure has its own characteristics. Since firms do not collaborate with firms at random, there are factors explaining why a network has the structure that it has. Factors such as industry (Salavisa et al., 2012), types of actors included (Nieto, Santamaría, 2007) as well as geography (McKelvey et al., 2003) have shown to have an important impact.

As pointed out by (Pavitt, 1984; Hagedoorn, Narula, 1996), the sector is a defining factor in the innovation process. It would hence be interesting to study how the structure of an innovation network behaves according to the sector of analysis. This paper will focus on the innovation network of the French Aerospace sector. This sector was chosen for two reasons. First, it is a high technology sector that plays an important role in the French economy as well as the european economy. Second, the sector is organized in a particular manner, it is a value chain. This value chain has been been optimized by Airbus with its Power8 program. This particular type of sector should transpire into the structural dynamics of the collaboration network. The analysis of the structure of the network in this paper will be based on three levels of analysis: the global network level, the level of the clusters and the level of the firm. The latter will be be accomplished using an Exponential Random Graph Model (ERGM). The aim of this analysis is to identify which factors incite firms to collaborate with one firm rather than another. The sector is a large supply chain build around the European assembler Airbus. The latter has is a prime example of a modular firm in the sense that it has externalized most of its production to suppliers. The different parts of the aircraft are produced by different sections of the production chain. Each of which contain pivot firms (Frigant et al., 2006) which link the different parts of the aircraft together. In addition, since the year

[1] On a more micro level other factors, such as absorption capacity, clustering and degree distribution also have a vital role to play in the efficiency of knowledge diffusion.

2000, Airbus has been working on its "Power8" program, aiming at the optimization of its supply chain. Given these characteristics we would expect that the collaboration network of the French aerospace sector will closely resemble that of the production chain. Since the production chain is build up from a small number of highly connected firms (pivot firms) and a central assembler (Airbus) I propose the following hypothesis:

Hypothesis 1a: The structure of the collaboration network of the French aerospace sector is a core-periphery structure.

In order to better understand how this structure came to be the mechanisms that drive link creation between firms need to be identified. In other words, I want to know why did firm "i" collaborate with "j" rather than "k".
Technological proximity between firms is a requirement for cooperation. If firms are too similar, they work on the same technologies and hence would not want to collaborate. As the technological distance increases the complementarity of the knowledge bases of the firms increases. This results in an increase in the probability of observing a collaboration. This complementarity does however reach a point where technologies become too distant and the complementarity decreases. This results in turn in a decrease in the probability of cooperation. These statements induce the second hypothesis to test:

Hypothesis 1b: There is an inverted U-shape relation between the probability of a collaboration and the technological proximity of two firms.

In addition to technological proximity, social proximity is expected to play an important role when it comes to partner selection, especially since the "power8" program launched to streamline production. Reputation as well as similar work methods allow firms to work more efficiently by reducing frictions due to diverging methods. I hence propose the following hypothesis:

Hypothesis 1c: Collaborators of collaborators have a higher probability to collaborate than firms without a common connection.

Once the structure has been analyzed the focus switches to the link between the position of the firm in the network and its financial performance. As was stated earlier, knowledge flows through the network. According to the position of the firm inside the network, a firm can be exposed to more or less diverse knowledge flows, impacting its performance. Financial data on firms inside the network is used to measure the performance of the firm.
I mobilize the Schumpeterian hypothesis that innovation is achieved by the recombination of ideas. This hypothesis implies that firm exposed to a large variety of ideas will have a high potential for innovation (Dosi, 2000; Cowan, Jonard, 2007). In other terms, the advancement on the inventive trajectory will be faster when the knowledge diversity available to the firm is stronger. When diversity is low firms risk decreasing returns to innovation. A variable called "neighborhood diversity" is used which computes for each year the number of technologies in the neighborhood of the firm. Each technology is

considered to be an IPC code. The aim is to measure the diversity in the neighborhood, the IPCs of the focal firm are hence not included in the measure. This leads to the following hypothesis:

Hypothesis 2a: The technological diversity in the neighborhood of the firm has a positive impact on its performance.

Two theories claim the importance of clustering in a network. The two theories do however oppose each other when it comes to the sign of the impact. A first theory suggests that having collaborators work together results in a positive impact on innovation and performance. The cooperations allow for a better understanding of the functioning of each firm. This information will allow firms to better organize their innovative activities.The effect is enhanced when cooperations are repeated over time, the more they know about each other the more efficient the cooperation. The other theory however suggests that a social lock-in might occur when firms cooperate too often, they would rather work with people they know rather than take the risk of finding a partner that is not efficient. This may result in a reduction of the innovativeness of firms, by the means of a stagnation or even reduction of the diversity of technologies. Instead of cooperating with a firm that masters new technologies they keep cooperating with firms that master the same technologies. This leads to the following hypothesis:

Hypothesis 2b: Clustering has a positive impact on the performance of the firm due a better mutual understanding of firms.

Notice that if this hypothesis is invalid then the theory on social lock-in would be valid. A network connects firms by creating paths between them. Knowledge flows between firms that are directly or indirectly connected. A firm with a position on many of these paths has access to more knowledge flows. This position is measured by the betweenness centrality coefficient which takes into account the position of a firm on path between other firms (Wasserman, 1994). The higher the centrality of the firm, the more it is on the crossroads of knowledge flows. The higher the centrality of the firm, the more it is able to benefit from diverse sources of knowledge.

The average distance gives a measure of the average distance a firm is removed from all other firms in the network. The closer it is to all other firms the more beneficial the knowledge flows should be. An argument against this idea is that if the distance is too low there is a high risk of redundancy of information and hence low distance should have a negative influence on the performance of the firm. I hence test the following hypothesis:

Hypothesis 2c: The more central the firm, the better the performance due to an increased access to knowledge flows.

The number of patents gives an indication of the innovative dynamism of the firm. The more patents are deposited by the surrounding firms the more knowledge they accumulated. The following hypothesis is hence tested:

Hypothesis 2d: The more patents in the neighborhood of the firm the stronger the knowledge spillovers to the focal firm.

Knowledge spillovers are only useful for a firm if she is able to absorb the knowledge it is exposed to. I use the number of technologies mastered by a firm as a proxy for the absorption capacity of the firm. The more technologies mastered by the firm the easier it should be for the firm to learn new knowledge which should result in increased performance. This gives the final hypothesis.

Hypothesis 2e: The absorption capacity of the firm is positively related to its performance.

In this chapter I will first introduce the main assumptions for this sector, then the methods that will be used to determine the structure of the network, and the impact of the position of each firm in this structure on its performance.

2 Data

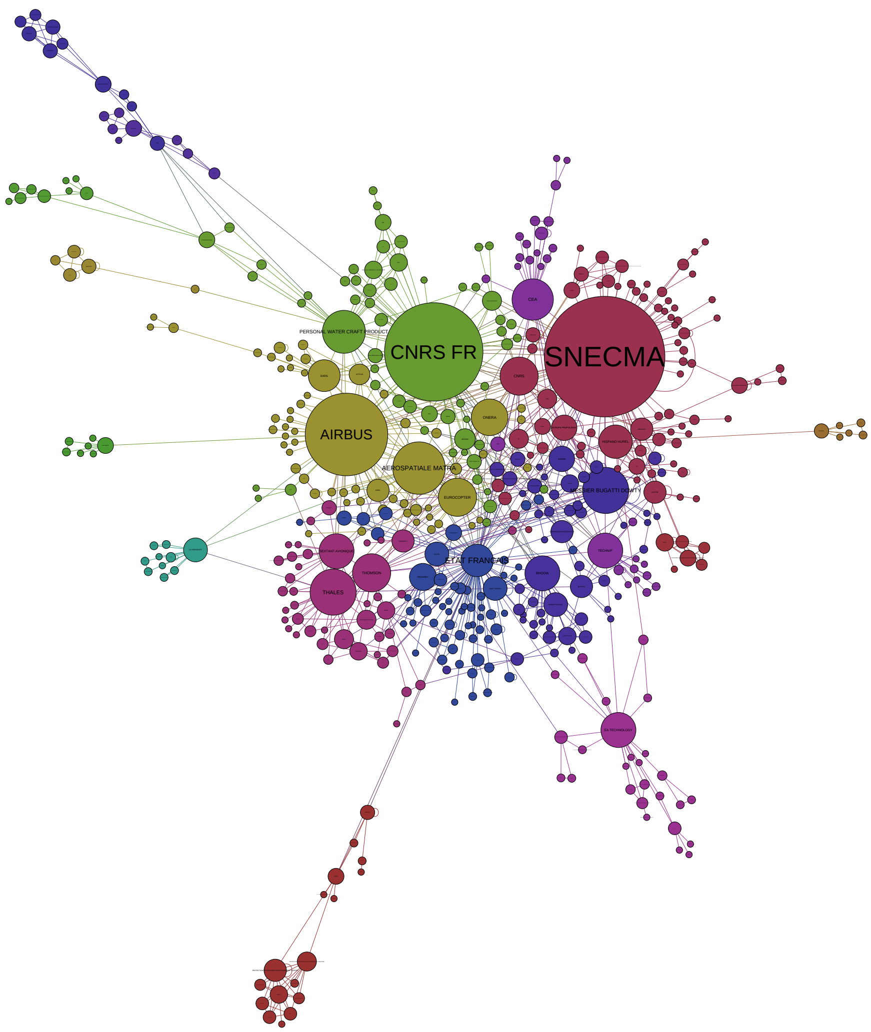


Figure 1: The aerospace collaboration network as of 2014. Node size is proportional to the number of collaborations, colors correspond to structural clusters identified by a maximization of modularity.

2.1 Patent data

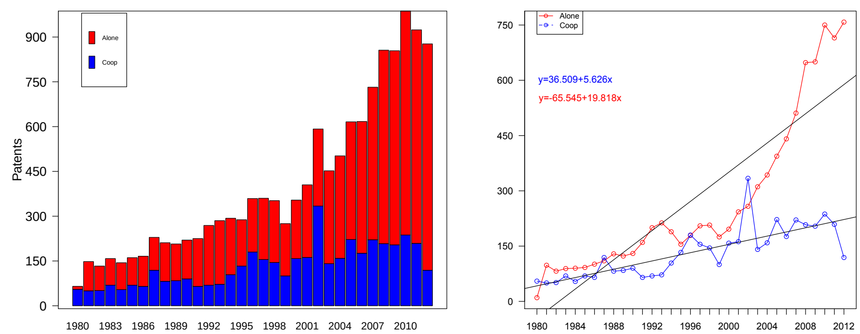
Since our focus is on knowledge flows, data on collaborations that were initialized for the purpose of creating new technologies is required. For this purpose an innovation network is created using from patent data. Whenever two or more firms are present on the same patent a link is created between the firms. All patents were extracted from the Orbit database, the firm names in the dataset were treated by hand to remove any typos and text lost in translation.

I restricted the focus on Patents deposited in France by French companies in order to avoid any problems with data from different patent offices. For instance, the USPTO tends to cite more intensely than the other offices while the German firms make a heavier use of utility models. Restricting our dataset allows us to avoid biases in these aspects.

In order to select patents relative to airplane technologies a query was constructed using a combination of keywords and IPC codes. I found that using only keywords resulted in a heavy percentage of false positives while selecting patents according to NACE codes was too restrictive. The combinatory method allows us to focus on all the different technologies that make up an airplane. After all, an airplane is the perfect example of a multi-technology product (Prencipe, 1997).

Building such a query does require specific knowledge about the technologies inside an aircraft and their corresponding keywords and IPC codes. The query used here was provided by the VIA-INNO platform[2] and is the result of repeated discussions between aircraft engineers and the platform to ensure viable results. The query resulted in a dataset of 11992 patents with a priority date between 1980 and 2013. 9544 (79.59%) patents were deposited by a single firm, 2448 (20.41%) patents were subject to a collaboration. From the 2448 patents that were identified 4369 cooperations between 1309 companies during the time period (1.78 cooperations on average per patent). Aggregation of these collaborations results in the network in figure 1.

[2]Plateforme d'intelligence économique labélisé centre d'investissement sociétale par l'initiative d'excellence de Bordeaux dans le cadre des investissements d'avenir de l'Etat Français http://viainno.u-bordeaux.fr/ (Website)



(a) Histogram of evolution of the number of patent deposits

(b) Trend in deposits

Figure 2: Evolution of the number of patents and the corresponding trend. A distinction is made between the number of patents deposited alone (red) and the number of patents deposited by collaboration (blue)

Figure 2 shows the evolution of the number of patents deposited between 1980 and 2013. In figure 2(a) I distinguish between patents deposited by one firm and patents that are the result of a collaboration. Figure 2(b) shows a clear positive trend in both patenting and collaborative patenting in the aerospace sector. Similar observations have been identified in other sectors such as biotech and software by (Pyka, Scharnhorst, 2009; Gulati et al., 2011; Salavisa et al., 2012) . The trend for patenting alone (5.626) is however much higher than for cooperative patenting (19.82).
One can observe an important increase in the number of patents from the year 2000 onwards. This can be explained partially be the commercialization of the Airbus A380. A particular aspect of the aerospace sector is the fact that there are mass patent deposits after the commercial release of an airplane which might explain some of the variance in the dataset.

2.2 Financial data

The objective of the this section is to establish a link between financial performance and structural position. The structural position of the firm is important mainly because of knowledge flows. Innovations are achieved by the recombination of knowledge (Schumpeter, 1942). Since the knowledge stock inside a firm expands slowly and diversity decreases over time, external knowledge sources are important. The position of the firm inside the network defines the number and the diversity of knowledge sources to which the firms has access.
A panel data analysis will be presented to estimate the influence of the position of the firm on it's performance.
Financial data is hence required for the identified firms. From the sample of 1309

depositors all research institutions, financial institutions and government agencies need to be removed. 676 firms were identified in the dataset of 1309 firms. The financial performance of the firm will be measured by the Return On Assets (ROA) of the firms:

$$ROA_t = \frac{Net\ Income_t}{Total\ Assets_t} \tag{1}$$

The ROA seems the appropriate measure since the denominator of the ROA includes intellectual property and all capital mobilized for R&D activities. The data will be extracted from the Amadeus database. Since we have network data over 34 years it would be optimal to have 34 years of financial data. This was however not possible due Amadeus' policy. Firms are automatically deleted from the database once they have not transferred any data for 3 years. This means that firms that changed their names during the 34 year period are no longer in the database. Using DVDs from a previous version of Amadeus (between 2000 and 2007) it is possible to extract a relatively complete dataset over the years 2000 to 2012.

3 Methods

In order to check the hypotheses about the structure of the global network, methods are required. These methods are the same as those used in the previous chapter.

3.1 Core-periphery detection

The core-periphery structure is identified from the degree distribution of the network. A core-periphery network is defined a small number of densely connected firms and a large number of firms with a low number of links. Using the Cumulative Frequency Distribution derived from the degree distribution of the network one can fit a function to the data in order to test if the network has a core-periphery structure (see Appendix D for more details).

3.2 Small-World detection

In order to check if our network has a small world structure I follow a methodology presented by (Gulati et al., 2012). Small world structures are defined by a low average distance and a high clustering coefficient. The Clustering coefficient of a network is defines as the ratio of observed triangles in the network to the number of possible triangles. The average distance is simple the average number of links between any two nodes in the network.
Since nodes can be added each year I need to make sure that a decrease in clustering is the result of less firms connecting in triangles and not the simple result of an additional node that reduces the overall clustering coefficient. The coefficients are hence normalized and compared to a random network with an identical number of nodes and links.
The theory behind small worlds is that random networks have low clustering while empirical networks have higher clustering. The latter is the results of social / economic /

geographic / ... motivations of the entities inside the network. As such, a network is a small world if its clustering coefficient is higher than that of a random graph of identical dimension (i.e same number of nodes and same number of links). This would hence imply that the graph is not random and that there are some underlying rules dictating the creation of ties in the network.
As for the average distance, it should be roughly identical to that of a random graph. I note Cr (Lr) the clustering coefficient (path length) of the random network and C (L) the clustering (path length) of the empirical data.
We hence need to observe $\frac{C}{Cr} >> 1$ and $\frac{L}{Lr} \approx 1$.
The evolution of the network was considered following two methods: using a 5-year sliding window and a method in which data was added year after year.

3.3 Exponential Random Graph Model

An Exponential Random Graph Model models the global structure of a network while allowing inference on the likelihood of a link between two nodes. It is basically a modified logistic regression, the models are modified in the sense that they do not require a hypothesis of independence between observations. For instance, if firm A is connected to B and C, there is a high probability that B knows C through its connection with A. A link between B and C has hence a higher probability than B connecting with a another, random, node. This implies that a link between two nodes depends upon the existing structure of the network. Regular logistic regressions are unable to account for these aspects since they require links to be independent upon each other. These levels of dependence are vital for the understanding of social and economic networks. The ERGM model to be estimated takes the form given in equation 2.

$$Pr(X = x \mid \theta) = P_\theta(x) = \frac{1}{k(\theta)} \cdot exp(\theta_1 \cdot z_1(x) + \theta_2 \cdot z_2(x) + \ldots + \theta_p \cdot z_p(x)) \quad (2)$$

Where X is the empirical observed network, x is the simulated network, θ a vector of parameters, z_i the different variables and $k(\theta)$ the normalizing constant. In short, the probability that the network generated by the model is identical to the observed network depends upon the given variables. If one consider that technological proximity has a role to play, it will be introduced as a variable. The model will then generate links while increasing (iteratively) the probability that nodes with higher proximity will connect. This is repeated a certain number of times. If, on average, the network generated is equal to the observed network then one can conclude that proximity plays a role the structuring of the network. For a more complete explanation of ERGM models see Chapter 2 of this thesis (or Lusher et al. (2012)).

3.4 Measuring Technological proximity

Many measures of technological proximity exist, some are based on patent citations (Chang, 2012), (Marco, Rausser, 2008), (Mowery et al., 1998) while others use IPC codes (Jaffe, 1986), (Breschi et al., 2003). The idea is that the different technologies firms work on are not chosen at random, they co-exist because they have factors in

common (Teece et al., 1994). This idea has led to different measures of technological proximity between firms, the most prominent was initiated by (Jaffe, 1986) further developed by (Breschi et al., 2003). Finer measures exist, see for instance (Bar, Leiponen, 2012) or (Bloom et al., 2013).
For the present chapter it is chosen to use an IPC based measure of technological proximity. A slightly different measure than the ones previously cited will be used, even though based on IPC codes. Our aim is to provide the likelihood of a cooperation based on the technologies mastered by firms. Therefore I assume that firms cooperate on technologies that are closely related in order to ensure proper incorporation of new technologies into an aircraft. As such having one technology in common is motive enough for two firms to cooperate. If one were to use one of the more common measures the prediction could be biased.
An IPC takes the following form: B64C1/18. Each part of the code (B, 64, C, 1,/18) indicates a practical classification. B stands for Performing operations and Transporting, B64 reduces the technologies to Aircraft, Aviation and Helicopters, B64C denotes Airplanes and Helicopters, B64C1 are Fuselages, wings etc. B64C1/14 are windows. The longer the code the more precise the technology. The full length of the IPC-codes is used in order to capture the largest amount of details of the technologies. When a firm deposits a patent one can deduce from the IPC codes what a firm has been working on and which technologies it masters. The measure of technological proximity is based on an analysis of IPC codes. The indicator of proximity computes the overlap in IPC codes between two companies. Table 1 shows two firms with 3 IPC codes. The numbers in the matrix correspond to the level of proximity. If both firms work on B they will have an overlap of one, if they both work on B64 the overlap is 2 and so-on. The proximity is maximal when firms deposit patent in the same 9 digit IPC codes. It takes the value of 0 when there are no elements in common.

		Firm B		
		B64C/19	B53D/01	C01F/03
Firm A	B64C/19	4	1	0
	B53D/01	1	3	0
	C01F/03	0	0	2

Table 1: Illustration of the proximity measure used in the ERGM

I defend the position that knowledge about one specific technology is enough to initiate a collaboration. The use of complete portfolios would induce a lot of noise in the data. In the end, firms cooperate often for a particular set of skills and not for all the skills used by a firm. A downside of this method is that the dataset is reduced to firms depositing both alone and by cooperation. One can only assume a firm masters a certain technology if it has deposited a patent alone. Cooperation data is then needed to create a network. Firms that only deposit by cooperation are hence excluded from the dataset. A proximity matrix was computed for 176 firms and generated the network that connected them.

3.5 Variable lags for the panel regression

This study uses data from two different sources. The financial data from 2012 comes from the performance in the year 2012, the patent data from 2012 does however result from cooperations that took place any time before 2012. In order to perceive an effect of the cooperation on performance lags need to be included in the patent-related variables. How far back the lags should go depends entirely on the type of information, some have a faster influence on the performance than other do. In terms of lag we will consider that a cooperation is initiated three years before the priority date of the patent. This means that the transfer of some types of information may flow from that point on. The effects of the knowledge flow should be visible at about the date of priority of the patent. The effects of the production of the patented technology should be visible (if the technology is indeed put into production) at any point in time from $t-1$ on.
Structural variables: Firms are influenced by the knowledge held within the firm at the moment of collaboration. The diversity is hence lagged to $t-3$: firms connected by a patent in 2010 cooperated in 2007 and are hence influenced by the diversity in the firm in the year 2007. However, since it takes time to absorb the knowledge and put it to use the impact on the ROA should be observed some time after the initialization of the cooperation, I will consider 3 years. Hence the variable Diversity is not lagged, the same is applied to the number of patents and the number of technologies. All the other variables are lagged at $t-3$ since the knowledge flows may influence the performance from the start of the cooperation on.

4 Results on the network structure

4.1 Cluster identification

The previously identified dataset leave us with over 4300 collaborations. The collaborations allow us to generate a network by creating a link between all firms that have deposited a patent together. The result is shown in figure 1. The bigger the size of the node the more collaborations the firm has. The coloring is the result of a community detection algorithm based on modularity. Modularity measures how well defined communities are inside a graph. Modularity gives a value between 0 and 1, the more the value tends towards 1 to more clearly defined the communities are (Newman, Girvan, 2004). For the result to be significant one expects a value of at least 0.6.
An algorithm introduced by (Blondel et al., 2008) was used to identify these communities using the open-source program Gephi (Bastian et al., 2009).
This community detection algorithm identifies communities inside a network purely based on the structural properties of the network. It starts by assigning each node with a community, it then selects a node at random and create a community with one of it's direct neighbors. The neighbor with whom it will create a community is the one that will maximize the modularity of the graph. This step is continued until maximum modularity is achieved. This method has the advantage of detecting automatically the number of communities (clusters) in the network while other methods ask the user for a fixed number of communities to be identified.

The results should however be handled with caution. The random component selects a node at random. It is possible that different results emerge if a different node is chosen at the start of the algorithm. In fact, the sequence of choice of the nodes plays an important role in the detection of the communities. I hence ran the algorithm several times to make sure the same communities were detected on average.

The results are rather interesting given that the communities were clearly defined and easy to interpret. Different communities were identified around the following firms:

- Hispano Hurel: Nacelles
- Rhodia: Chemicals
- Thompson: Seating
- Messier Bugatti: Landing and braking.
- Pechiney Rhenalu: Structural elements (aluminium)
- Alcatel Lucent: Avionics and communication systems

These clusters suggest local technological development according to different parts included in the production of an aircraft. This allows us to understand the previously identified scale-free network structure. The large assemblers (Airbus, Snecma and Thales) and the CNRS have a large number of links connecting them with first order suppliers which in turn have their own clusters in which they are densely embedded.
This observation coincides with the industrial organization of the sector, which is indeed rather hierarchical. Airbus, at the center, designs the aircrafts while externalizing large portions of the production process to first order suppliers (Frigant et al., 2006). The latter will work with other, second order suppliers. As such there are not many competitors but competition is tough between the few (Niosi, Zhegu, 2005). The sector has undergone a significant restructuring in the 90' and the 2000's resulting in the specialization of some suppliers while others diversified their production to include other sectors (Frigant et al., 2006). In addition, the sector has high barriers to entry, mainly because of high level of knowledge required. The sector need an influx of cutting-edge technologies and hence close collaboration with fundamental research. The collaboration network that I observe here reflect these sectorial aspects: in a central position the CNRS (National Centre for Scientific Research) can be found providing an influx of fundamental science to the large manufacturers and first order suppliers. While clusters exist around the first order suppliers connecting specialized and diversified suppliers. This results in a particular network structure that is made up from an interconnection of clusters. The overall structure of the network resembles a connected caveman structure (Watts, 1999a) in which each specific part of the airplane is developed in it's own cluster. In terms of knowledge these firms need to collaborate with a large number of firms from different clusters in order to assemble an aircraft. While there is no need for direct knowledge flows between the landing and braking system and the nacelle manufacturer, Airbus needs knowledge on both technologies to assemble the final product.
The exception being that some firms connect all the clusters. Airbus has this central

position since it needs to absorb knowledge from all clusters. Very little knowledge flows seem to exist between clusters, while there is a necessity for transfer intra-cluster. Innovation in the aircraft industry is the result of an interplay of technology push and market pull (Dosi, 2000). On the one side aircraft manufacturers aim at making their aircrafts more cost efficient while there is a demand for governments to reduce noise and make planes more eco-friendly.

4.2 Structural Dynamics

In order to identify the structure of the network I will track the evolution of the network from 1980 onwards. This will allow us to have a clear vision of the structuring of the network.

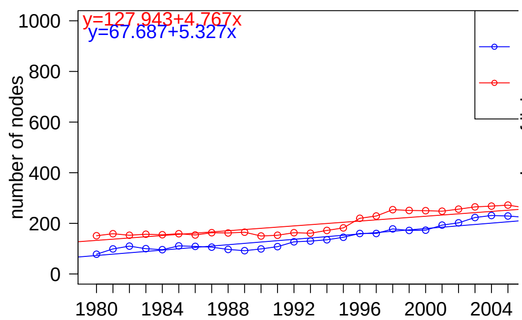


Figure 3: Evolution of the number of nodes with a 5-year sliding window (i.e 1980 → 1980-1984). "GC" is the giant component of the network, "whole" the giant component with all the smaller components

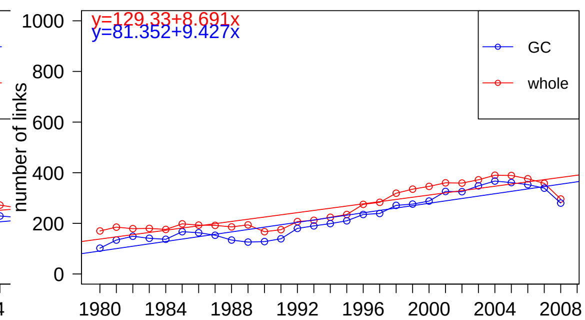


Figure 4: Evolution of the number of links with a 5-year sliding window (i.e 1980 → 1980-1984). "GC" is the giant component of the network, "whole" the giant component with all the smaller components

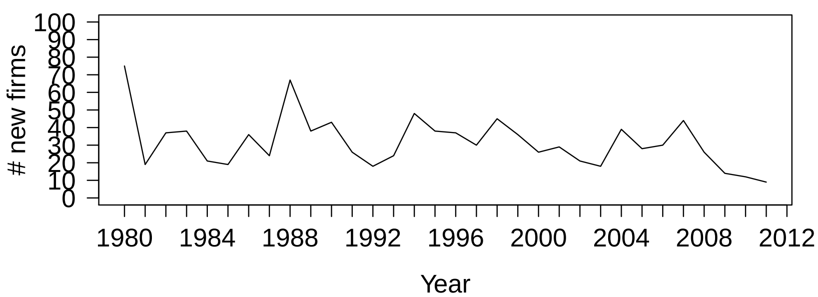


Figure 5: Evolution of the number of new firms entering the collaboration network each year.

Figure 5 reports the number of new firms that enter the network each year. The variance is explained by the previously discussed patenting behavior in the sector. The evolution of the number of nodes (figure 3) is computed using a sliding window of 5 years. This allows to keep track of the active firms in the network. This shows us that the network increases in size over the period with a decline during the last period (note that 2008 implies the frame 2008-2013). The decline can be explained by two factors. First, a small decline in the number of deposits in the last couple of years (figure 2(a)). Second, the decline in the number of firms might be explained by the "Power8" program launched by Airbus in order to optimize their production chain which resulted in a decrease in the number of suppliers. The evolution of the network was considered in two ways: using a 5-year sliding window and a method in which data was added year after year. The results are reported in figure 6 and 7.

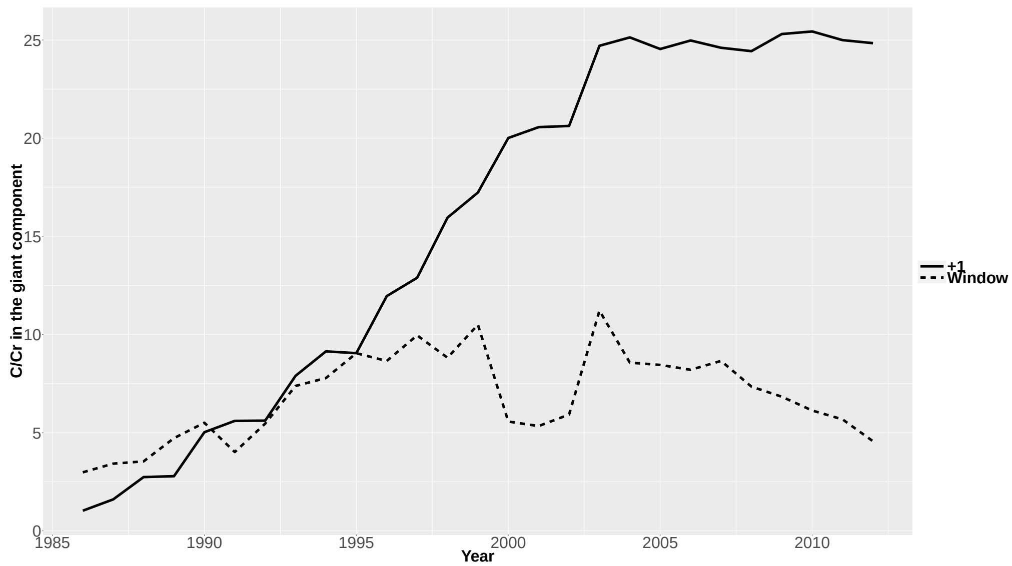


Figure 6: Adjusted clustering coefficient

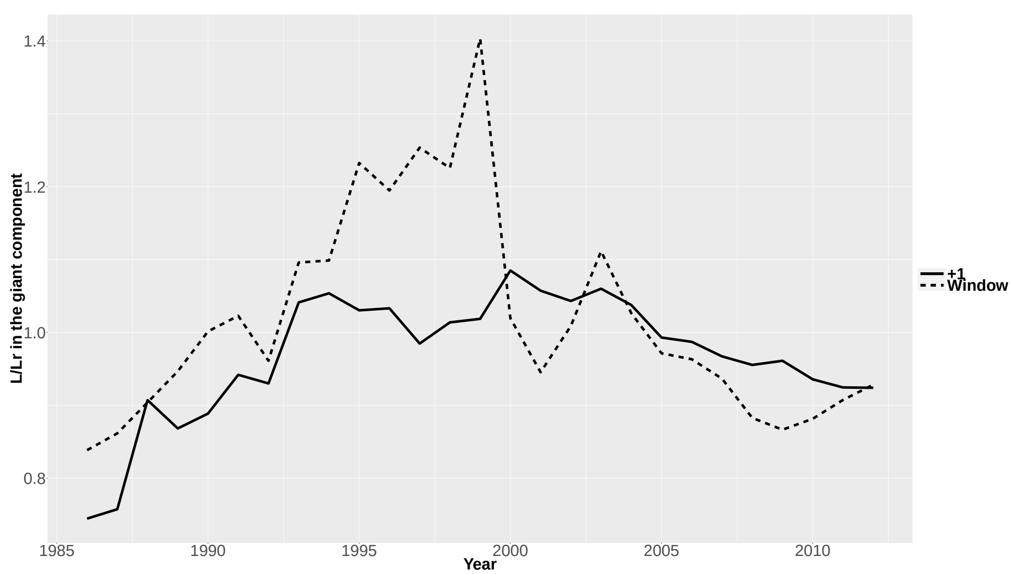


Figure 7: Adjusted average distance

Figure 6 shows that the clustering coefficient trends strongly away from 1, indicating that the clustering observed in the networks increases faster than clustering in a random network of identical dimension. This is the case for both methods, showing that even when one removes firms that are no longer part of the network, the clustering stays higher than random. This high level of clustering is due to the different clusters that build the different parts of the airplane. These clusters are highly interconnected resulting in a high level of clustering. The power8 program which had the aim of optimizing the supply chain appears to have had a significant impact on the network, creating a decrease in the clustering coefficient that remained for a couple of years. The average distance shows a similar decrease around this period, clearly showing the effects of the program. The 5-year window shows that the average distance of the network was too high (as compared to a random network) to be considered a small world. The different clusters in the network were not interconnected enough to be considered a small world. The drop in the year 2000 however, allows the network to reach the small world butter-zone. The +1 method shows that the network converges towards a small world early on and stays its course until the year 2007 where is converges towards the 5-years window. The network appears to have stabilized. I hence find converging conclusions from the results in Gulati et al. (2012) who identifies an inverted U-shape in the small worldliness of the collaboration network. The structure of the network seems to be highly correlated with the structural specificities of the aerospace sector. Indeed, knowledge stays within the clusters since specific knowledge is developed inside each cluster. Knowledge flows between clusters through pivot firms interconnecting the clusters. Communication and knowledge flows are necessary between firms inside clusters since the parts developed by firms in clusters need to interact and need to be compatible. The most central firms hence benefit from the most knowledge flows since they have to assemble the different parts of the plane.

It can be concluded here that there is a high tendency for firms to cluster which confirms our previous observation that firms where organized in interconnected clusters. The structure also appears to stay relatively stable when it comes to these two indicators, especially in the time-laps network. In the 90' has started a radical change in the organization of the sector resulting in many suppliers exiting the sector which has as a consequence a lower number of collaborators. These collaborators collaborate more intensively resulting in a more stable structure towards the end of the period.

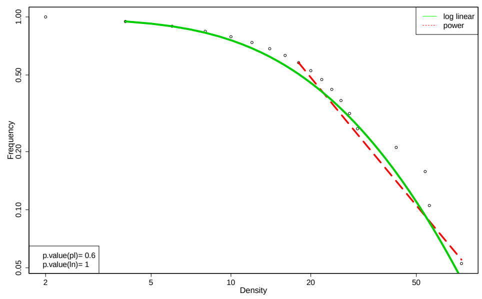


Figure 8: Power-Law and log-normal fit for 1996 (window)

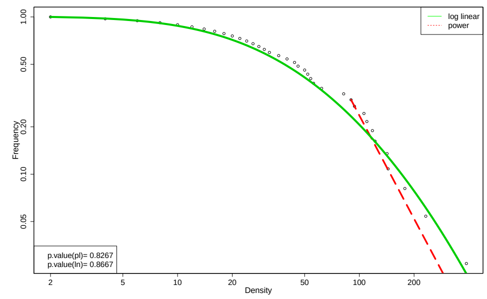


Figure 9: Power-Law and log-normal fit for 1996 (+1)

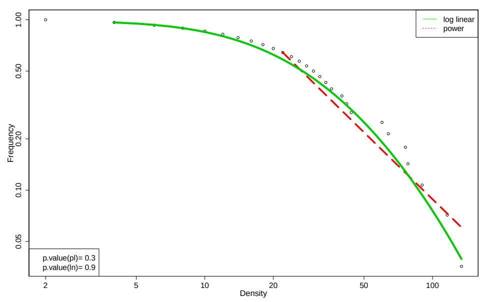


Figure 10: Power-Law and log-normal fit for 2006 (window)

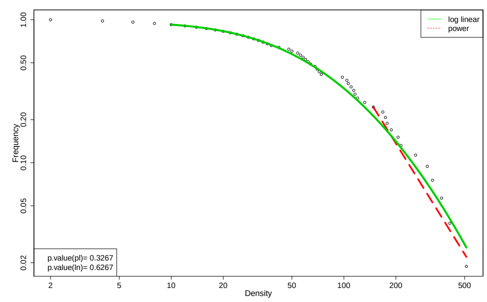


Figure 11: Power-Law and log-normal fit for 2006 (+1)

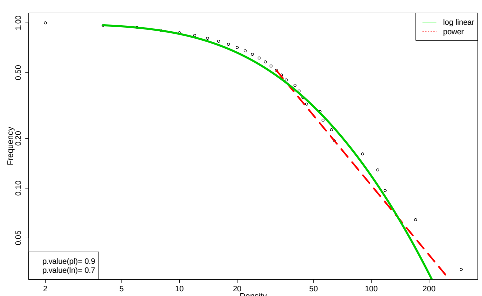


Figure 12: Power-Law and log-normal fit for 2012 (window)

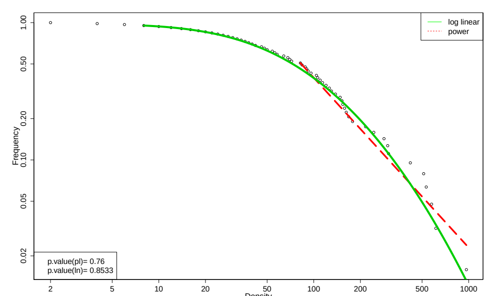


Figure 13: Power-Law and log-normal fit for 2012 (+1)

Quite interestingly, the network appears not only to have small world features but also core-periphery features. Figures 8 to 13 show the CFD of the network as well as the fitted functions. Recall that the null hypothesis (Data comes from a power-law distribution) is rejected when the P.value is smaller than 5%. Even though the power-law is significant, it is only significant starting at a high density ($x_{min} \geq 20$). It cannot be concluded here that the network is scale-free. However, the log-normal fit is significant for both the window and +1 method. This implies that the degree distribution of the network follows a log-normal distribution stabilizing around $\mu = 3.44$ and $\sigma = 0.992$ (see table 2. The distribution shows that a large fraction of the nodes of the network have a relatively low density. At the same time, there is a low fraction of the nodes that have a relatively large density. The fraction of nodes with a low density is the periphery of the network. These are the firms inside the different clusters as can be seen in figure 1. The small number of firms with a higher density are the pivot firms, Airbus and the CNRS. The latter are connected to many firms inside the clusters to oversee the production of the different part they need to assemble. In addition they are connecting different clusters. The parts they create need to be compatible with other parts of the airplane. Interactions are hence required to ensure compatibility.

These elements result in core-periphery characteristics at the level of the global network structure. The network takes this structure from the early stages of the network until the end. The results in table 2 show the parameters of the adjusted law. The structure of the network stabilizes around the year 2005 for the +1 method, and a couple of years earlier for the window.

In conclusion then, the network has both small world and core-periphery characteristics. Similar results have been found in other types of networks by Guida, Maria (2007) and Requardt (2003). From these observations hypothesis 1a can be considered verified. In conclusion then, knowledge creation in the aerospace sectors is a localized phenomenon. Knowledge is generated in different clusters in which pivot firms assure the diffusion of this knowledge to the rest of the network.

Year	Mean (window)	SD (window)	Mean (+1)	SD (+1)
1983	2.17	0.79	2.48	0.31
1984	2.55	0.56	2.28	0.56
1985	2.40	0.80	2.24	0.77
1986	2.55	0.58	2.26	0.80
1987	2.55	0.69	2.80	0.46
1988	2.30	0.90	2.47	0.87
1989	2.46	0.69	2.49	0.90
1990	2.28	0.89	2.52	0.89
1991	2.60	0.70	2.50	0.86
1992	2.65	0.67	2.78	0.76
1993	2.66	0.71	2.75	0.80
1994	2.50	0.87	2.82	0.82
1995	2.50	0.85	2.84	0.78
1996	2.73	0.57	2.95	0.80
1997	2.56	0.83	3.12	0.73
1998	2.65	0.67	3.18	0.73
1999	2.75	0.69	3.34	0.63
2000	2.68	0.78	3.10	0.84
2001	2.81	0.83	3.05	0.95
2002	2.81	0.86	3.27	0.75
2003	2.72	1.04	3.32	0.65
2004	2.63	1.04	3.30	0.74
2005	2.83	1.05	3.32	0.88
2006	2.86	1.03	3.33	0.90
2007	2.91	1.04	3.34	0.92
2008	2.86	1.07	3.28	1.04
2009	2.76	1.09	3.37	0.94
2010	-1.42	2.10	3.42	0.98
2011	2.43	1.13	3.44	0.99
2012	2.07	0.83	3.44	0.99

Table 2: Evolution of the parameters of the fitted laws.

4.3 Micro level motivations for collaboration

An ERGM model is used to determine the mechanisms that rule link creation. Table 3 shows the regression results, note that these coefficients cannot be interpreted as such. In order to compute the precise impact one needs to transform them into odds.
The results show that several factors explain the global structure of the network. It was hypothesized that technological proximity was a decisive factor in collaboration between firms in the aerospace sector. The models shows that this is indeed the case. Firms with a higher technological proximity have a tendency to work together. More precisely the odds of a link between firms that are technologically close is higher than the odds of a link between firms that are technologically far.
Moreover there appears to be an inverted U-shape to this relation as shows by the

significance of the variable proximity2. This would imply that firms collaborate if they can learn from one another but if they are too close in terms of technology then the probability of a link deteriorates. Firms that are too close in terms of technologies can consider that the other firm has nothing to offer them and hence prefer collaborating with a firm that has different technologies. Hypothesis 1b is hence verified.
The $altkstar$ parameter checks (and controls) for the core-periphery structure. Since the parameter is significant we see that the model has correctly identified the scale-free structure previously found.
Taking the $kstar2$ and $triangle$ parameter together allows for checking for triadic closure (Lusher et al., 2012) . Since both the parameters are significant I conclude that firms with a common node have a higher probability of connecting than firms with no common node. It hence seems that the trust that diffuses through the network as well as the increased performance due to common practices is a motivator for collaboration. Hypothesis 1c is verified.
Finally, co-citations are significant as well. Implying that firms that cite each-others patents will end up collaborating at some point in time.

5 Results on the impact of network position of the firm on performance

Two types of variables were included in this regression. Structural variables and technology variables. We have a panel of 1605 observations over a 10 year period. A standard linear panel regression to test the influence of the network on the performance of the firm is used. The previously discussed variables were included with the corresponding lags:

$$ROA_{t,t+1} = Clustering{*}density_{t-3} + Centrality_{t-3} + AverageDistance_{t-3} + Technologicaldiversity + Numberoftechnologies + Numberofpatents + Numberofcooperations$$

In a first regression only the variables relative to the position of the firm inside the network (model (1)) were used, a second regression includes only the technology variables (model (2)), the last model show the regression with both types of variables (model(3))
In order to assess which type of regression is adequate for the data several statistical tests were performed. The Lagrange Multiplier Test (Breusch-Pagan) showed that there is presence of panel effects in the data, simple OLS regressions are hence rejected.
I then checked for time fixed effects in the data, by adding a dummy variable for each year and compared the regression results with an F-test, the results show that no time-fixed effects have to be included in the model. A fixed, random and pooled model were then tested against each other, the fixed effects was retained as the best model. Since the data presented serial correlation and heteroscedasticity, I used robust estimates.

	Dependent variable:		
	Network		
	(1)	(2)	(3)
edges	-7.267^{***}	-1.121^{*}	
	(0.228)	(0.626)	
kstar2	0.155^{***}		
	(0.003)		
degree2		-1.336^{***}	14.467^{***}
		(0.255)	(2.970)
edgecov.citation			-20.934^{***}
			(1.090)
triangle	3.428^{***}	1.923^{***}	1.726^{***}
	(0.007)	(0.0001)	(0.0002)
gwesp	-0.439^{***}		
	(0.166)		
gwesp.alpha	0.523		
	(0.385)		
edgecov.proximity2		1.565^{***}	6.620^{***}
		(0.271)	(0.345)
altkstar.1.6		-1.864^{***}	
		(0.172)	
altkstar.1.7			-3.371^{***}
			(0.086)
Akaike Inf. Crit.	578722	617651	9813
Bayesian Inf. Crit.	578760	617689	9851
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3: ERGM model results

The results of the regression are shown in table 4. All variables have a significant impact on the ROA with the exception of the number of cooperations and the number of patents. The latter observation is rather to be expected. Not all patents have the same value and only a small portion of patents have an exploitable value. The number of cooperations shows that not all cooperations have a benefit in terms of knowledge flows. The number of collaborations being higher than the number of collaborators, it can be interpreted as the intensity of collaborations between firms, i.e how close firms are socially. The impact of social links is an order of magnitude lower than the impact of knowledge transfer by other objects and is difficult to capture.

The structural variables are all significant, showing that the position of the firm in the network does indeed have an impact on the performance of the firm. The adjusted clustering measure shows that firms with a higher clustering coefficient perform better. The collaboration of collaborators is hence a positive effect. The idea that working with people who already know each other seems to be validated.
In terms of knowledge absorption the central position of a firm is significant. The more central the firm is, the more knowledge it is able to absorb. The measure retained here is the betweenness centrality which measures the extend to which a firm is positioned on the a path between all the firms in the network. The higher the centrality the more favorable the position for knowledge absorption. The Average distance measures how far is firm is positioned from other firms, the further away the less knowledge the firm is exposed to. As such, the negative coefficient of this variable confirms the hypothesis that knowledge flows in the network have a decaying factor.

The technology related variables highlight the importance of technological diversity. Innovation literature puts forth the idea that innovations are achieved by the recombination of ideas. The diversity of technologies in the neighborhood of the firm should hence have a positive impact on the performance of the firm. The regression shows that this hypothesis is validated.
The final variable, the number of technologies mastered by the firm, has a negative impact. In our particular case, i.e the aerospace sector; the firms with the most technologies are suppliers with a specific position in the value chain. The regression show that specialized firms perform better than diversified firms, in a network. Specialized firms have to advantage of detaining valuable knowledge that can result in efficient innovations through collaboration. Diversified firms might be less interesting for cooperations and hence partner with less than optimal partners.

6 Conclusion

The production chain characteristic of the aerospace sector results in a network in which different clusters foster different technologies. These clusters are interconnected by a small number of large firms resulting in a Core-Periphery structure. The specificities of the aerospace sector play a vital role in the shaping of the collaboration network. The central position of Airbus in the networks ensure an interconnection of all different clusters. Knowledge is required to flow from each cluster this central firm. Knowledge

	Dependent variable: Return on Assets		
	Network var.	Techno. var.	Combined
Adjusted clustering	0.646**		0.623*
	(0.313)		(0.322)
Centrality	0.890*		0.941*
	(0.513)		(0.501)
Average distance	−0.328**		−0.335***
	(0.128)		(0.127)
Technological diversity		0.002***	0.001***
		(0.0004)	(0.0005)
Number of technologies		−0.005***	−0.005***
		(0.001)	(0.001)
Number of patents		0.004	0.003
		(0.004)	(0.004)
Number of cooperations		0.001	−0.001
		(0.004)	(0.003)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 4: Panel regression results

is created locally in this network and diffuses through the pivot firms to the assembler. The Power8 program instigated by Airbus in the early 2000's had for main objective to streamline the production chain, and this appears to have had as a result a small world structure in the collaboration network.
On a micro-level this chapter has shown that technological proximity explains collaborations between firms but that this behavior follows an inverted U-shape. There is hence a butter-zone for the level of proximity that leads to collaboration.
The analysis of the performance of the firm tends to indicate that a central position in the network goes hand in hand with better performance for the firm. This is explained by the access to knowledge flows by firms with a high centrality and a low average distance. The choice of partner is proven to be important for two reasons, the clustering of the firm and the specialization of the firm. If the partner evolves in an environment in which collaborators of collaborators collaborate, this will have a positive impact on it's performance. If the firm choses a specialized firm to innovate with this will also have a positive impact on the performance of the focal firm.

References

Ahuja Gautam. Collaboration networks, structural holes, and innovation: A longitudinal study // Administrative science quarterly. 2000. 45, 3. 425–455.

Alghamdi Mohamad, McDonald Stuart, Pailthorpe Bernard. The Emergence of a Small World in a Network of Research Joint Ventures. 2012.

Bar Talia, Leiponen Aija. A measure of technological distance // Economics Letters. 2012. 116, 3. 457–459.

Barabási Albert-László, Albert Réka. Emergence of scaling in random networks // science. 1999. 286, 5439. 509–512.

Bastian Mathieu, Heymann Sebastien, Jacomy Mathieu. Gephi: An Open Source Software for Exploring and Manipulating Networks. 2009.

Baum Joel AC, Shipilov Andrew V, Rowley Tim J. Where do small worlds come from? // Industrial and Corporate change. 2003. 12, 4. 697–725.

Blondel Vincent D, Guillaume Jean-Loup, Lambiotte Renaud, Lefebvre Etienne. Fast unfolding of communities in large networks // Journal of Statistical Mechanics: Theory and Experiment. 2008. 2008, 10. P10008.

Bloom Nicholas, Schankerman Mark, Van Reenen John. Identifying technology spillovers and product market rivalry // Econometrica. 2013. 81, 4. 1347–1393.

Breschi Stefano, Lissoni Francesco, Malerba Franco. Knowledge-relatedness in firm technological diversification // Research Policy. 2003. 32, 1. 69–87.

Chang Shann-Bin. Using patent analysis to establish technological position: Two different strategic approaches // Technological Forecasting and Social Change. 2012. 79, 1. 3–15.

Cowan Robin, Jonard Nicolas. Network structure and the diffusion of knowledge // Journal of economic Dynamics and Control. 2004. 28, 8. 1557–1575.

Cowan Robin, Jonard Nicolas. Structural holes, innovation and the distribution of ideas // Journal of Economic Interaction and Coordination. 2007. 2, 2. 93–110.

Dosi Giovanni. Innovation, organization and economic dynamics: selected essays. 2000.

Les territoires de l'aéronautique: EADS, entre mondialisation et ancrage. // . 2006.

Guida Michele, Maria Funaro. Topology of the Italian airport network: A scale-free small-world network with a fractal structure? // Chaos, Solitons & Fractals. 2007. 31, 3. 527–536.

Gulati Ranjay, Lavie Dovev, Madhavan Ravindranath Ravi. How do networks matter? The performance effects of interorganizational networks // Research in Organizational Behavior. 2011. 31. 207–224.

Gulati Ranjay, Sytch Maxim, Tatarynowicz Adam. The rise and fall of small worlds: Exploring the dynamics of social structure // Organization Science. 2012. 23, 2. 449–471.

Hagedoorn John, Narula Rajneesh. Choosing organizational modes of strategic technology partnering: international and sectoral differences // Journal of international business studies. 1996. 27, 2. 265–284.

Jaffe Adam B. Technological opportunity and spillovers of R&D: evidence from firms' patents, profits and market value. 1986.

A dynamic model of network formation with strategic interactions. // . 2009.

Lusher Dean, Koskinen Johan, Robins Garry. Exponential random graph models for social networks: Theory, methods, and applications. 2012.

Marco Alan C, Rausser Gordon C. The role of patent rights in mergers: Consolidation in plant biotechnology // American Journal of Agricultural Economics. 2008. 90, 1. 133–151.

McKelvey Maureen, Alm Håkan, Riccaboni Massimo. Does co-location matter for formal knowledge collaboration in the Swedish biotechnology–pharmaceutical sector? // Research Policy. 2003. 32, 3. 483–501.

Morone Piergiuseppe, Taylor Richard. Knowledge diffusion dynamics and network properties of face-to-face interactions // Journal of evolutionary economics. 2004. 14, 3. 327–351.

Mowery David C, Oxley Joanne E, Silverman Brian S. Technological overlap and interfirm cooperation: implications for the resource-based view of the firm // Research policy. 1998. 27, 5. 507–523.

Newman Mark EJ, Girvan Michelle. Finding and evaluating community structure in networks // Physical review E. 2004. 69, 2. 026113.

Nieto María Jesús, Santamaría Lluis. The importance of diverse collaborative networks for the novelty of product innovation // Technovation. 2007. 27, 6. 367–377.

Niosi Jorge, Zhegu Majlinda. Aerospace clusters: local or global knowledge spillovers? // Industry & Innovation. 2005. 12, 1. 5–29.

Pavitt Keth. Sectoral patterns of technical change: Towards a taxonomy and a theory // Science Policy Research. 1984. 13. 343–373.

Prencipe Andrea. Technological competencies and product's evolutionary dynamics a case study from the aero-engine industry // Research policy. 1997. 25, 8. 1261–1276.

Pyka Andreas. Innovation networks in economics: from the incentive-based to the knowledge-based approaches // European Journal of Innovation Management. 2002. 5, 3. 152–163.

Pyka Andreas, Scharnhorst Andrea. Innovation networks // Innovation Networks: New Approaches in Modelling and Analyzing, Understanding Complex Systems, ISBN 978-3-540-92266-7. Springer-Verlag Berlin Heidelberg, 2009. 2009. 1.

Requardt Manfred. Scale free small world networks and the structure of quantum space-time // arXiv preprint gr-qc/0308089. 2003.

Salavisa Isabel, Sousa Cristina, Fontes Margarida. Topologies of innovation networks in knowledge-intensive sectors: Sectoral differences in the access to knowledge and complementary assets through formal and informal ties // Technovation. 2012. 32, 6. 380–399.

Schumpeter Joseph A. Capitalism, socialism and democracy. 1942.

Teece David J, Rumelt Richard, Dosi Giovanni, Winter Sidney. Understanding corporate coherence: Theory and evidence // Journal of Economic Behavior & Organization. 1994. 23, 1. 1–30.

Verspagen Bart, Duysters Geert. The small worlds of strategic technology alliances // Technovation. 2004. 24, 7. 563–571.

Wasserman Stanley. Social network analysis: Methods and applications. 8. 1994.

Watts . Networks, dynamics and the small world phenomenon // American journal of sociology. September 1999a. 105, 2. 493–527.

Watts Duncan J. Networks, dynamics, and the small-world phenomenon 1 // American Journal of sociology. 1999b. 105, 2. 493–527.

Cahiers du GREThA
Working papers of GREThA

GREThA UMR CNRS 5113

Université de Bordeaux

Avenue Léon Duguit
33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75
Fax : +33 (0)5.56.84.86.47

http://gretha.u-bordeaux.fr/

Cahiers du GREThA (derniers numéros – last issues)

2016-08 : LORENZON Emmanuel, *Collusion with a Greedy Center in Position Auctions*

2016-09 : DELFAUD Pierre, Retour sur une expérience unique d'interventionnisme en économie de marché: La Politique agricole commune européenne (1955-2015)

2016-10 : MUTASCU Mihai, PEREAU Jean-Christophe, URSU Eugen : A Wavelet Analysis of the Environmental Kuznets Curve in France

2016-11 : LISSONI Francesco : Migration and Innovation Diffusion: An Eclectic Survey

2016-12 : BONIN Hubert : Les paradoxes de l'économie maritime française au début du XXIe siècle

2016-13 : CASTIGLIONESI Fabio, NAVARRO Noémi : (In)Efficient Interbank Networks

2016-14 : PEREAU Jean-Christophe, MOUYSSET Lauriane, DOYEN Luc : Groundwater management in food security context

2016-15 : FAURE Yves-André: Institutions locales et résistances au test du VIH/sida. Quelques leçons d'une enquête dans la ville de Fortaleza, Brésil

2016-16 : FAUVET Pierre : Autorisation de mise sur le marché, défaillance de la responsabilité, et groupes de pression

2016-17 : BECUWE Stéphane, BLANCHETON Bertrand : French Textile Specialisation in Long Run Perspective (1836-1938) : Trade Policy as Industrial Policy

2016-18 : LEVASSEUR Pierre: The effects of bodyweight on wages in urban Mexico

2016-19 : LUNG Yannick, JULLIEN Bernard, MIDLER Christophe : De la Logan à la Kwid. Ambidextrie, innovation inversée et fractale, design-to-cost : les recettes de la stratégie Entry de Renault

2016-20 : TROMEUR Eric, DOYEN Luc : Optimal biodiversity erosion in multispecies fisheries

2016-21 :FAURE Yves-A : Institutions locales et résistances au test du VIH/sida. Quelques leçons d'une enquête dans la ville de Fortaleza, Brésil

2016-22 :VAN DER POL Johannes : The modelling of networks using Exponential Random Graph Models: an introduction

2016-23 :VAN DER POL Johannes : Social interactions between innovating firms: an analytical review of the literature

La coordination scientifique des Cahiers du GREThA est assurée par Emmanuel PETIT.
La mise en page est assurée par Julie VISSAGUET