
Innovation networks: the effects of collaboration in the oil and gas industries

Adonis Pedro Coutinho Barboza*

PETROBRAS – Petróleo Brasileiro S/A – Rod. Pres.,
Dutra Km 142 – Jd. Diamante,
São José dos Campos – SP/Brazil
and

PUC-SP – Pontifícia Universidade Católica de São Paulo,
Rua Monte Alegre, 984 – Perdizes,
São Paulo – SP/Brazil

Email: adoniscoutinhobarboza@gmail.com

*Corresponding author

Mauro Maia Laruccia

PUC-SP – Pontifícia Universidade Católica de São Paulo,
Rua Monte Alegre, 984 – Perdizes,
São Paulo – SP/Brazil
and

FUNDACENTRO – Fundação Jorge Duprat e Figueiredo,
Rua Capote Valente, Nº 710 – Pinheiros,
São Paulo – SP/Brazil

Email: mauro.laruccia@gmail.com

Abstract: This qualitative study of the innovation performance aims to determine whether collaboration affects patent productivity given network metrics such as centrality, structural holes, and hierarchy. The sample used in this study consists of 1,811 patents and 3,213 researchers related to the C10 area of knowledge with regard to oil, fuel gases, and lubricants. The study shows that Sinopec Limited has higher patent productivity than Royal Dutch Shell plc, BP plc, and Total S.A., with a descriptive level p-value of less than 0.05 and a 95% confidence interval. In contrast, BP, Shell, and Total cannot be differentiated based on patent productivity. The regression models in this study provided an explanatory power of 59.03% for Total, 42.43% for BP, 39.10% for Sinopec, and 14.14% for Shell. Although these models are limited to those specific companies, their statistically significant results explain patent productivity, confirming the hypotheses and providing a basis for future research.

Keywords: social network; collaboration; innovation; patent.

Reference to this paper should be made as follows: Barboza, A.P.C. and Laruccia, M.M. (2017) 'Innovation networks: the effects of collaboration in the oil and gas industries', *Int. J. Auditing Technology*, Vol. 3, No. 3, pp.201–216.

Biographical notes: Adonis Pedro Coutinho Barboza has a PhD in Innovation Management (PUC-SP; 2016), MA in Knowledge Management (PUC-SP; 2004) and BA in Business Administration (MACKENZIE; 1999). He is a Black

Belt Program Coordinator for LEAN SIX SIGMA; Examiner for the National Quality Award (PNQ); 2009/2010; member of the WTC Business Club – Vale do Paraiba. He is a Coordinator of the post-graduate course in Integrated Management Systems (SENAC; 2010). He is an Accredited Instructor through the National Brazilian Micro Enterprise Support Service (SEBRAE) in the areas of quality management and business planning and Instructor at the São José dos Campos Technology School (FATEC; 2007) for the Mobilization Program of the National Oil and Natural Gas Industry (PROMINP).

Mauro Maia Laruccia has a PhD in Communication and Semiotics by Pontificia Universidade Católica de São Paulo (2002), Master in Business Administration by Pontificia Universidade Católica de São Paulo (1997) and Bachelor in Business Administration by Pontificia Universidade Católica de São Paulo (1983). He is currently a Professor of FEA/PUCSP and Faculdades Integradas Campos Salles. AdHoc avaliador INEP/MEC. He has experience in business management with focus on the following topics: administration, education, communication, information technology, marketing and business strategy. His specialties are business management, strategic planning, organisational behaviour and semiotics.

1 Introduction

The relationship between collaboration and productivity is a premise of the theory of social networks, but it has not been statistically tested sufficiently to develop a generic explanatory model. Although some studies of that relationship from the perspective of networks have been conducted, the combination of variables that best explains it has not been identified. Buckley and Alstyne (2004, p.3; 2006, p.18) correlated various network metrics – the degree of centrality, betweenness centrality, and structural holes – with the performance of teams of 39 and 47 recruiters, obtaining an explanatory power (R^2) of between 0.19 and 0.40 and between 0.16 and 0.49. Kang et al. (2011, p.1072) in turn, correlated reputation and centrality with the performance of virtual teams of 100 attendants, obtaining an R^2 of between 0.086 and 0.496. Further studies, such as this one, contribute to the literature by enabling the comparison of results and thus the selection of appropriate variables and structure for explanatory models.

This study investigates several hypotheses about the effects of collaboration in innovation networks on the patent productivity of BP plc, Royal Dutch Shell plc, Sinopec Limited, and Total S.A., aiming to answer the following questions: Does collaboration affect the patent productivity of these companies, given network metrics such as degree of centrality, structural holes, and hierarchy? If so, how much effect does it have? Are there any additional independent variables that can increase the explanatory power of the model? Can the model be used to predict patent productivity?

According to Lin (1999), social analysis is more effective if restricted to examining the extent of relationships, once the explanatory power is higher inside the relationships boundary. In this sense, the hypotheses considered in this study are tested for each of the companies in the sample: For each hypothesis, the descriptive statistical significance (p-value) and the R^2 are estimated and the hypothesis is either accepted or rejected as described in Table 1.

From a rational perspective, co-authoring of patents occurs when researchers expect to obtain gains from them. If, however, researchers expect to obtain advantages from monopolisation, they avoid collaboration. In this sense, the theory of social networks has important applications for understanding the relationships among researchers in high-tech sectors.

Table 1 Hypotheses about the effects of variables on patent productivity

<i>Hypothesis</i>		<i>Rule for acceptance of the hypothesis</i>
H0	The selected variable has no effect on patent productivity.	The hypothesis is accepted if simple regression of patent productivity on the selected variable provides a descriptive level p-value greater than or equal to 0.1.
H1	The selected variable has an effect on patent productivity.	The hypothesis is accepted if simple regression of patent productivity on the selected variable provides a descriptive level p-value less than 0.1.
H1a	The selected variable has a statistically significant effect on patent productivity.	The hypothesis is accepted if multiple regression of patent productivity on the selected variable provides an R ² greater than 0.2.
H1b	The inclusion of the selected variable improves the explanatory power of the model.	The hypothesis is accepted if multiple regression of patent productivity on the selected variables provides an R ² for the model greater than 0.7.
H1c	The model can be used to predict patent productivity.	The hypothesis is accepted if multiple regression of patent productivity on the selected variables provides residuals with an Anderson-Darling normality test p-value greater than 0.05.

The sample used in this study consists of patents obtained by the selected companies. Patents are important in innovation: They serve as a source of technological information. They also are useful research subjects on which hypotheses and scientific developments in the area of business administration can be tested. Centrality, hierarchy, and structural holes are attributes of social networks and was selected as the main variables in the explanatory models proposed in this study.

2 Social relationships

The cohesion of social relationships was the focus of Coleman's study (1998). He suggested that the best social structure consists of dense, interconnected networks. He also stated that interconnection produces synergy and strengthens the groups that recognise social exchanges and share conviviality rules. Building close relationships assumes that people will maintain ties with people with whom they have something in common.

Granovetter (2009) also contributed to the discussion of cohesion in social groups. He proposed the concept of social immersion, which was the basis for the establishment of economic sociology. He questioned the application of rationality to economic choices.

The similarity of contexts brings together social actors who then interact with one another and exchange experiences via dialogue, learning, and joint action. In similar contexts, relationships are easily formed because they are based on good faith. Granovetter highlighted collaboration, suggesting that social immersion influences people to adopt voluntary behaviours based on affinities.

The initiation of relationships, on the other hand, assumes that prolonged contact with the same people is stressful and impairs such relationships. Simmel (2010) was a pioneer in the study of social relationships with his empirical research on life in the metropolises. He stated that the expansion of cities is a response to the dull effect of its inhabitants' continuous interaction with their peers. He pointed out that the search for new relationships is a response to the human trait of learning from differences.

Granovetter (1983) also empirically researched the importance of diversity in social relationships. In his study, he advocated the promotion of weak ties to gain new knowledge from relationships. For him, redundant ties do not add new information and therefore have little value in social relations.

Burt (1992), in turn, suggested that the optimal social structure includes gaps caused by actors who are connected to other networks. The ties resulting from collaborative networks are important vehicles in obtaining external knowledge. Weak ties and structural holes are strategic resources that enable effective collective action produced by social capital (Lin, 1999). On the other hand, the openness influences behaviour reciprocity because distant relationships are based on rules with less social context and are more streamlined.

3 Network analysis

Network analysis, a tool for the study of social action, is applicable to structural, relational, and individual levels. It includes metrics that enhance research on social relationships. At the structural level, connectivity defines and detects the components of a network, the places that exhibit specific cohesion properties such as high density. At the individual level, the main concept is centrality, a family of properties related to the structural importance or distinctiveness of an individual in the network. Finally, at the relational level, structural equivalence is the extent to which pairs of actors take on similar roles in the network.

Centrality is an important concept in social network theory. It gives the individual a strategic advantage. The greater the number of connections an individual has, the more network resources he or she can call on. The central actors have more favourable social situations in the network than do the peripheral actors. In this regard, the examination of relationships considers obligations and expectations. They support the structure and, in their concentration, provide the agents with the confidence to establish and maintain social groups.

Betweenness, an aspect of the broader concept of centrality, measures the potential of an individual to control others' relationships. A broker can arbitrate contacts, gain advantages, isolate other actors, and prevent contacts. Betweenness is a particular type of centrality that identifies how often an individual mediates others' relationships in these ways. It is a measure of the power that actors can claim to stop flows or distort information to serve their own interests.

The performance of an individual is due, in part, to the resources that their social ties allow them to access. The network's structure and the position of the individual in the network are crucial factors in predicting outcomes (Burt, 2004). Because it is by brokerage action that networks are engaged, expanded, and promoted, brokers play a significant role on the outcome of a social network.

Brokers also play the key role of linking peripheral actors to central actors. They are representatives of the systemic guidelines of social network. Brokers in dense social networks are noteworthy because they are privileged channels of communication. They are responsible for the isolation or interconnectedness of the networks. Brokerage, in addition to being a structural characteristic of networks, is a process that involves motivation and opportunity. Opportunities are derived from connections in the network, which enable the creation of new relationships by which individuals obtain information and knowledge (Burt, 1992).

4 Social constraints

Actors are direct agents in a network's structure because they build, change, and break social ties. The behavior of members of a social group can be evaluated based on the characteristics of the ties among the members. In this sense, constraint measures the effect of the actions of brokers because they are third parties who determine the structure of relationships (Burt, 1992; Obstfeld et al., 2014).

However, the value of a relationship is not defined within the limits of immediate connections. The distribution of ties in the social network must be considered. Constraint is a summary measure that incorporates social context (Burt, 1992). It measures the concentration of direct and indirect ties in a network. If constraint is high, ties with the individual were cultivated from a few number of relationships. If constraint is low, ties with the individual were cultivated from a higher number of relationships.

4.1 Social roles

In networks, actors can act as communicators, opportunists, or facilitators (Obstfeld et al., 2014). Communicators are the drivers of information. Their presence may be a non-redundant tie, that is, the only contact path. They stand out because new relationships with them are important sources of news and innovation is driven by the ability to establish relationships. Opportunists filter and treat information before passing it on to other people. They try to prevent the establishment of new ties, to maintain their unique positions. They may also distort information to influence other actors' perceptions of one another. By doing so, they favour social situations that interest them. Facilitators promote new ties among the actors with which they connect. They restructure networks, promote integration, and strengthen social capital. By doing so, they achieve visibility and recognition from actors who gain advantage from the newly established ties. They may also serve as the central node of a new structure composed of the new relationships. These roles played by individual actors in a network deserve special attention because they affect the continuity or interruption of the social cycle.

4.2 *Privileged position*

The privileged position is related to the role played in the social structure. The number of ties that an actor creates, unlike centrality, has no relation to this structural position. Actors in privileged positions are the brokers of dense relationship networks. They are the only channels of contact between networks. They influence communication, promote relationships, or prevent relationships among other actors. Brokerage and constraint are stages of social action that reconfigure networks. Opportunistic attitudes constrain the formation of new ties and hinder the development of the network and innovation.

Lin (1999) proposed the individual location as a concept that relates the concentration of social capital with the position of the individual in the social structure. It applies Burt's (1992) concept of structural holes and Granovetter's (1983) concept of weak ties to social network analysis. It enables the identification of the concentration of social capital in a social network. Brokerage restricts the effectiveness of social groups and contains the seeds of misappropriation of advantages gained by the individual efforts of members. In highly fragmented groups, social capital is highly concentrated in delegates or representatives.

Given the level of brokerage and constraint, the actions of actors may be isolated, peripheral, opportunistic, or collaborative. Actors with a low level of brokerage ability are isolated in the dynamics of social structures. They are not driven to create new relationships because they do not represent the shortest connection path among other actors. Increased immersion and the strengthening of their social context are dependent on other actors.

With the strengthening of social context, an actor approaches the centre of the social network. This favoured individual location allows the individual's actions to have peripheral effects on the social structure. Such actors may brokerage relationships that provide them with bargaining power. If such actors choose to keep their privileged positions, they hinder the relationships that they mediate. Their opportunistic approach maintains *structural holes* that separate actors and disrupt social dynamics. These empty spaces represent the distance between an individual's indirect connections (Burt, 1992). Finally, individuals in privileged positions may choose to strengthen social capital by facilitating relationships. Such collaborative action eliminates gaps and allows the social cycle to fulfil its purpose.

5 **Research method**

This paper investigates the effects of collaborative innovation networks on patent productivity in the oil and gas industries by analysing centrality, structural holes, and hierarchy metrics of those networks. Initial data collection and development of the analysis approach took place during January to October 2014 at the UCINET (Borgatti et al., 2002), and secondary data were collected during August to December 2014. Statistics for confidence intervals and regressions were obtained via use of MINITAB 17.

Table 2 Patents in the sample

<i>Company</i>	<i>Number of patents</i>	<i>Share</i>
BP plc	378	21%
Royal Dutch Shell plc	484	27%
Sinopec Limited	715	39%
Total S.A.	234	13%
	1,811	100%

Source: Google Patent Search (2014)

The sample consists of patents obtained by the selected companies, which were chosen among the seven energy companies with the greatest net revenues in 2013. There were collected 1,811 patent grants related to the C10 area of knowledge which is concerned the oil, fuel gases, and lubricants (Table 2).

Patents are excellent subjects for research because they are validated and constitute legal contracts between the company and the inventor (Lemos, 2011). As a source of technological knowledge, patents have important application to the study of innovation in networks.

Data processing is critical in studies that use network analysis tools because errors affect the composition of the network and the calculation of its metrics (Borgatti et al., 2013). In this research, the following potential errors were considered: duplication of records and the spelling of the inventors' names. Duplication of records was eliminated by sorting and grouping observations in the database. The records were examined to ensure that each patent number was a unique identifier and that there were no remaining duplicates. Similarly, errors in the spelling of the inventors' names were observed but corrected. The presence of a dot in the abbreviation of a name or was interpreted as a new researcher.

The degree of centrality, structural holes, and hierarchy are the independent variables that were selected as measures of cooperation in the networks. The degree of centrality and hierarchy were calculated in the UCINET 6 routine, named *structural holes* in the method *ego network model* (Borgatti et al., 2002). The variable structural holes was calculated in the UCINET 6 routine, named *reinforced structural holes*. Hierarchy, structural holes, and constraint were measured according to the algorithm proposed by Burt (1992). The variable indirect ties was calculated in the UCINET 6 routine, named *structural holes* in the method *whole network model*.

6 Analysis and results

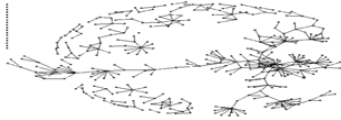
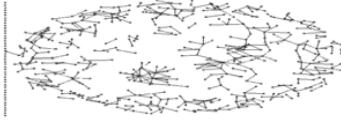


The innovation networks for the patents in the sample are composed of 3,213 researchers. As shown in Table 3, Sinopec has 1,902 researchers, or 59% of the sample, followed by Shell (640, or 20%), BP (382, or 12%), and Total (289, or 9%).

Centrality provides an actor with a strategic advantage. The results shown in Table 3 suggest that Sinopec's innovation network has the greatest degree of centrality (2.606), followed by Total (2.035), BP (1.974), and Shell (1.3). Structural holes are empty spaces that separate social actors from their indirect connections (Burt, 1992). Sinopec's network has the greatest level of structural holes (0.154), followed by Total (0.135), BP

(0.127), and Shell (0.087). Finally, hierarchy represents social ties to a sole individual. Sinopec’s network has the greatest hierarchy (0.630), followed by BP (0.629), Total (0.563), and Shell (0.572).

Although this descriptive analysis takes into consideration the differences among the innovation networks, it does not enable us to make conclusive statements. We must perform statistical tests to determine the confidence intervals of these averages and regressions to estimate the correlations among the variables and the explanatory power of the models. Although the social network theory suggests that centrality, structural holes, and hierarchy influence productivity (Freeman, 1979; Burt, 1992), only statistical models can determine to what extent the independent variables explain this theoretical causality.

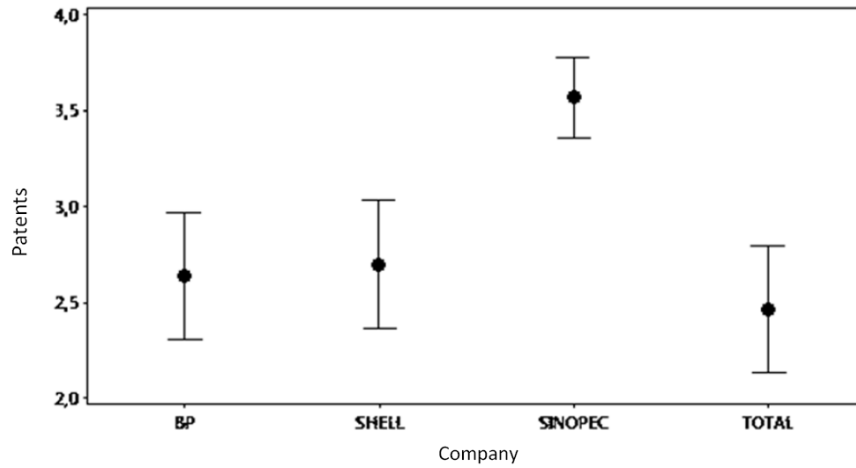
Table 3 Network metrics for the sample

BP plc (382 researchers)		Centrality 1.974 Structural holes 0.127 Hierarchy 0.629
Royal Dutch Shell (641 researchers)		Centrality 1.300 Structural holes 0.087 Hierarchy 0.572
Sinopec Limited (1,902 researchers)		Centrality 2.606 2.606 Structural holes 0.154 Hierarchy 0.630
Total S.A. (289 researchers)		Centrality 2.035 2.035 Structural holes 0.135 Hierarchy 0.563

6.1 Comparison studies

The results of comparison studies support the conclusion of the averages for the networks. Comparison of the variance of patent productivity confirms that Sinopec’s network has the greatest patent productivity in the sample, with a descriptive level p -value < 0.05 and a confidence interval of 95% (Figure 1). The patent productivity of BP, Shell, and Total cannot be differentiated based on the descriptive analysis.

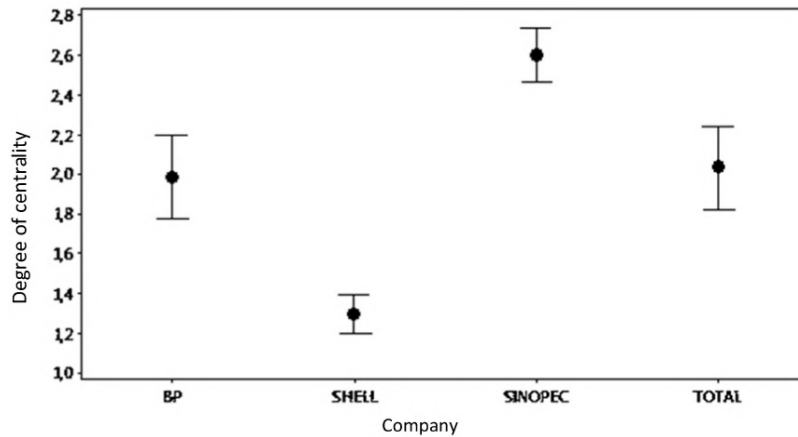
Figure 1 Analysis of the variance of patent productivity



Source: Prepared by the author on MINITAB 17 using the research data

Sinopec’s network clearly has the greatest centrality, followed by those of BP and Total (Figure 2). Shell’s network has the least centrality. Sinopec, the company with the greatest patent productivity, also has the greatest network centrality. Therefore, by observing confidence intervals, it is possible to deduce a relationship between centrality and patent productivity.

Figure 2 Analysis of the variance of the degree of centrality

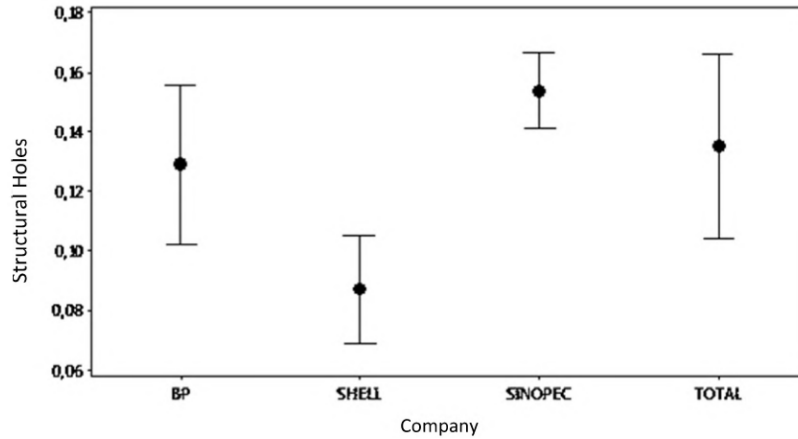


Source: Prepared by the author on MINITAB 17 using the research data

Shell’s network has the lowest level of structural holes (Figure 3). BP, Sinopec, and Total have similar confidence intervals for structural holes. It is important to take into consideration that, of the companies in the sample, Shell has the most researchers (120 researchers with no co-authoring). It is followed by BP (18 researchers), Total (14), and Sinopec (18). Furthermore, Shell’s network is divided into 126 groups, followed by

Sinopec (98), BP (42), and Total (31). Although Shell’s network has a low level of structural holes, it is fragmented and exhibits less co-authoring among subgroups of researchers.

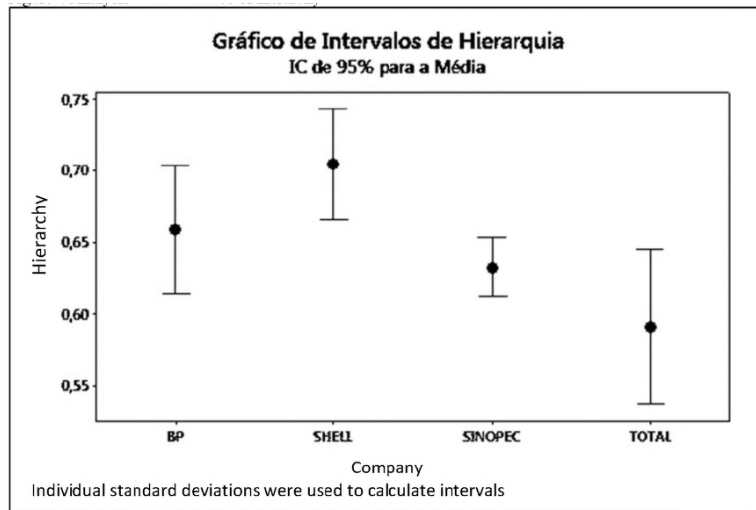
Figure 3 Analysis of the variance of structural holes



Source: Prepared by the author on MINITAB 17 using the research data

Finally, confidence intervals show that Shell has the greatest hierarchy level (Figure 4). BP, Sinopec, and Total, in turn, share similar confidence intervals for hierarchy. By observing confidence intervals, it is possible to infer that of the sample companies Shell focuses the least on promoting ties among groups within its innovation network because it has the lowest level of structural holes and the greatest hierarchy.

Figure 4 Analysis of the variance of hierarchy



Source: Prepared by the author on MINITAB 17 using the research data

Although the analysis of confidence intervals provides comparative conclusions, it does not characterise the relationships among the variables or the extent to which they influence patent productivity. To that end, correlation and regression analyses are required.

6.2 Correlation and regression

56%), followed by BP (33%), Sinopec (29%), and Shell (11%) (Table 4). The R² for structural holes is highest for BP (34%), followed by Sinopec (30%), Total (29%), and Shell (10%). The R² for hierarchy is highest for BP (35%), followed by Total (33%), Sinopec (31%), and Shell (18%). For Total, centrality (R² of 56%) is a much stronger driver of patent productivity than are structural holes (29%) or hierarchy (33%). For the other companies, explanatory power was relatively similar across independent variables. At this point, it is possible to consider the first question posed at the beginning of this paper: Does collaboration affect patent productivity for these companies? The answer is yes. The effects can be measured by the network’s centrality, structural holes, and hierarchy. Regression of patent productivity against these network metrics provides statistically significant results, with all coefficients having descriptive level p-value less than 0.001.

Table 4 Regression R² for descriptive level p-value < 0 001

<i>Company</i>	<i>Centrality</i>	<i>Structural holes</i>	<i>Hierarchy</i>
BP plc	33%	34%	35%
Royal Dutch Shell plc	11%	10%	18%
Sinopec limited	29%	30%	31%
Total S.A.	56%	29%	33%

Source: Prepared by the author on MINITAB 17 using the research data

Note: Dependent variable: patent productivity.

Although isolated correlation of independent variables with patent productivity identifies the effects, it cannot measure the effects. The combination of all of the variables in a model enriches the analysis by enabling estimation of the effects and prediction. The equation proposed for the company models in this study is as follows:

$$\text{Number of patents} = B_0 + B_1^* \text{Centrality} + B_2^* \text{Structural holes} + B_3^* \text{Hierarchy} \quad (1)$$

The equation relates the patent productivity of an innovation network to that network’s centrality, structural holes, and hierarchy, with the coefficients B_1 , B_2 , and B_3 representing the effects (Table 5).

The coefficients on the independent variables are statistically significant. The effects in the company models have a descriptive level p-value < 0.10, variance inflation factor (VIF) < 5, and Maslow Cp < = number of factors + 1. The VIF and Maslow Cp are statistical measures that check the consistency of the model by determining whether the independent variables are correlated with one another and whether the model includes too

many factors (Hair et al., 2006). The replication of identical factors in the company models is a research strategy that enables the comparison of coefficients across companies.

Table 5 Regression coefficients

<i>Company</i>	<i>y-intercept (B_0)</i>	<i>Centrality (B_1)</i>	<i>Structural holes (B_2)</i>	<i>Hierarchy (B_3)</i>
BP plc	0.182	0.7619	4.691	0.622
Royal Dutch Shell plc	2.668	0.9530	3.394	-0.674
Sinopec Limited	-0.582	0.8113	6.345	1.690
Total S.A.	-0.976	1.0906	3.119	1.444

Source: Prepared by the author on MINITAB 17 using the research data

Note: Dependent variable: patent productivity.

The Shell model has the highest y-intercept (2.668), followed by BP (0.182), Sinopec (-0.582), and Total (-0.976). Shell's high y-intercept is related to the number of its researchers in its network who do not co-author patents. The Shell network includes 120 isolated actors (of its total 641 researchers), followed by BP (18), Total (14), and Sinopec (8). A high y-intercept (B_0) is a characteristic of a network composed of isolated actors because isolated actors have null values for centrality, structural holes, and hierarchy.

At this point, it is possible to consider the second question posted at the beginning of this paper: How much do these network metrics affect patent productivity? For centrality, Total has the highest coefficient (1.0916), followed by Shell (0.9530), Sinopec (0.8113), and BP (0.7619). For hierarchy, Sinopec has the highest coefficient (1.690), followed by Total (1.444), BP (0.662), and Shell (-0.674). For structural holes, Sinopec has the highest coefficient (6.345), followed by BP (4.691), Shell (3.394), and Total (3.119).

Regarding explanatory power, the Total model has the highest R^2 (59.03%), followed by BP (42.43%), Sinopec (39.10%), and Shell (14.14%) (Table 6). Although the independent variables in the models do not fully explain patent productivity, they are valid in a statistically significant sense and can be considered in managing innovation networks.

Table 6 Explanatory power of the models

<i>Company</i>	R^2	<i>Adjusted R^2</i>
BP plc	42.43%	41.98%
Royal Dutch Shell plc	14.14%	13.73%
Sinopec Limited	39.10%	39.00%
Total S.A.	59.03%	58.60%

Source: Prepared by the author on MINITAB 17 using the research data

Note: Dependent variable: patent productivity.

Taking into account that these three independent variables are a starting point, data for Total were processed using a stepwise method in Minitab 17 to verify to what extent the model explains patent productivity. Total was chosen as the test case here because its model provided the highest R^2 . In this new model, we adopted a 95% confidence interval, the adjusted method type III for the sum of squares, and α of 0.15 for entry and removal of 23 factors related to network centrality, brokerage, structural holes, and fragmentation:

$$\begin{aligned} \text{Number of patents} = & 1.589 + 2.061 \text{ Centrality} + 4.578 \ln(\text{Constraint}) \\ & - 6.590 \text{ Indirect} + 4.059 \text{ Structural holes} - 2.111 \text{ Efficiency} \end{aligned} \quad (2)$$

At this point, it is possible to consider the third question that was posed at the beginning of this paper: Are there any additional variables that can increase the explanatory power of the model? The answer is yes. The hierarchy variable was removed from the initial model and variables for constraint, indirect ties, and efficiency were added. The variable $\ln(\text{Constraint})$ is the Napierian logarithm of the social investments of the actors in the network. Indirect is the number of indirect ties of the actor. Efficiency relates centrality with the number of indirect ties. In the new model, the explanatory power is satisfactory (R^2 of 71.48%, adjusted R^2 of 69.56%). The model's results have descriptive level p -value = 0, VIF < 5, and Maslow $C_p \leq p + 1$.

7 Discussion

The innovation networks of BP and Total shared similar characteristics. They had similar averages in the descriptive analysis and network graphs, showing a similar geodesic distribution for researchers. The analysis of confidence intervals confirmed the networks' similarity, but the regression analysis showed that Total's innovation network exhibits better adjustment to its model.

Sinopec's innovation network is noteworthy, with the most researchers, the most patents, and the greatest centrality. The analysis of confidence intervals confirmed these findings and corrected the conclusions that might be drawn, erroneously, from its position in terms of structural holes and hierarchy relative to other companies. The only company for which patent productivity is highlighted by confidence intervals is Sinopec. Measures of centrality, structural holes, and hierarchy had good explanatory power for Sinopec's patent productivity, although the best company model in terms of explanatory power is the model of Total.

In the study, the results for Shell differed from those for the other companies. The descriptive analysis of Shell showed the lowest measures of centrality and structural holes. The analysis of confidence intervals confirmed those results and found that the Shell network had the highest measure of hierarchy. In this regard, the negative coefficient on hierarchy in the Shell model is worthy of further research.

8 Conclusions

We have found statistically significant effects of innovation networks' centrality, structural holes, and hierarchy on patent productivity. In other words, in the innovation networks of *BP*, *Shell*, *Sinopec*, and *Total*, collaboration affects patent productivity (*H1*). The statistical significance of these effects (*H1a*) is evident from a R^2 greater than 0.2 for the networks of *BP*, *Sinopec*, and *Total*. An R^2 less than 0.2 for *Shell*'s network indicates a somewhat statistically significant relationship between collaboration and patent productivity.

Table 7 Regression R^2 of models in various studies

<i>Model</i>	R^2
Total S.A. (this study)	71%
Kang et al. (2011)	50%
Buckley and Alstyne (2006)	49%
Buckley and Alstyne (2004)	40%

Source: prepared by the author with the research data

In addition, the final model developed in this study, the more complex one for *Total*, has an R^2 of 71%, which is higher than the R^2 of models in other studies (*H1b*) (Table 7). However, it is important to point out the significant difference in the size of the networks in the samples used in this study and in the studies of Bulkley and Alstyne (2004; 2006) and Kang et al. (2011). Research regarding *Big Data* for patents has enabled large network size and greater explanatory power of the statistical results. Although Buckley and Alstyne (2004; 2006) and Kang et al. (2011) analysed the behaviour of 39, 47, and 100 actors, respectively, in this study we collected data on networks that totalled 3,214 researchers.

Although the models presented in this paper do not fully explain the relationship between the selected network metrics and the patent productivity of those innovation networks, their results are sufficiently statistically significant to confirm the hypotheses *H1*, *H1a*, and *H1b*.

9 Limitation

Models are designed, ideally, to predict outcomes in order to guide the management of innovation networks. However, there are statistical requirements that must be met to ensure the reliability of predictions. At this point, it is possible to consider the fourth question that was presented at the beginning of this paper: Can the model be used to patent productivity? The answer is no.

The models in this study generate a significant number of large residuals and atypical values, which prevents them from being characterised as predictive (*H1c*). A descriptive

level p-value < 0.005 was identified via the Anderson-Darling test of the models for all four companies. That results in the acceptance of the null hypothesis and the conclusion that the residuals are not normally distributed.

However, it is important to note that the significant number of large residuals and atypical values does not reduce the explanatory power of the models with regard to the investigated phenomenon. Although residuals and atypical values indicate predictive ability, statistical significance (p-value and R^2) indicate explanatory power.

Therefore, the results of the analysis statistically validate the hypotheses *H1*, *H1a*, and *H1b*. However, although BP, Shell, Sinopec, and Total are among the seven companies with the greatest revenues, they are insufficient to represent the whole oil and gas industries. Regardless, although the conclusions presented in this study are limited to these four companies, they explain patent productivity, confirm hypotheses, and provide a basis for future research.

References

- Borgatti, S.P., Everett, M.G. and Freeman, L.C. (2002) *Ucinet 6 for Windows: Software for Social Network Analysis*, Analytic Technologies, Harvard, MA.
- Borgatti, S.P., Everett, M.G. and Johnson, J.C. (2013) *Analyzing Social Networks*, Sage Publications, London.
- Buckley, N. and Alstynne, M. (2004) *Does Email Make White Collar Workers More Productive?*, NAACSOS (North American Association for Computational Social and Organization Sciences), June 2004 [online] <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.61.8179&rep=rep1&type=pdf> (accessed on July 2015).
- Buckley, N. and Alstynne, M. (2006) *An Empirical Analysis of Strategies and Efficiencies in Social Networks*, SSRN (Social Science Research on Network), February 2006.
- Burt, R.S. (1992) *Structural Holes: The Social Structure of Competition*, Harvard University Press, Cambridge, New York.
- Burt, R.S. (2004) 'Structural holes and good ideas', *American Journal of Sociology*, Vol. 110, No. 2, pp.349–399.
- Coleman, J.S. (1998) 'Social capital in the creation of human capital', *The American Journal of Sociology*, Vol. 94, No. 1988, pp.95–120.
- Freeman, L.C. (1979) 'Centrality in social networks: conceptual clarification', *Social Networks*, Vol. 1, No. 3, pp.215–239.
- Google Patent Search (2014) [online] <http://www.google.com/advanced patent search> (accessed on 15 December 2014).
- Granovetter, M. (1983) 'The strength of weak ties: a network theory revisited', *Sociological Theory*, Vol. 1, No. 1983, pp.201–233.
- Granovetter, M. (2009) 'Economic action and social structure: the problem of immersion', in A.C. Martes (Ed.): *Networks and Economic Sociology*, EdUFSCar, St. Charles.
- Hair, J., Black, B., Babin, B. and Anderson, R. (2006) *Multivariate Data Analysis*, Prentice Hall, New Jersey, Upper Saddle River.
- Kang, M., Kim, B., Gloor, P. and Book, G. (2011) 'Understanding the effect of social networks on user behaviors in community-driven knowledge services', *Journal of the American Society for Information Science and Technology*, Vol. 62, No. 2, pp.1066–1074.
- Lemos, R. (2011) *Intellectual Property*, Fundafao Getulio Vargas, Rio de Janeiro.
- Lin, N. (1999) 'Building a network theory of social capital', *INSNA – Connections*, Vol. 22, No. 1, p.2851.

- Obstfeld, D., Borgatti, S.P. and Davis, J. (2014) 'Brokerage as a process: decoupling third party action from social networks structure', in Brass, D.J., Labianca, G., Mehra, A., Halgin, D.S. and Borgatti, S.P. (Eds.): *Contemporary Perspectives on Organizational Social Networks*, , Vol. 40, pp.135–159, Research in the Sociology of Organizations, Emerald Publishing, Bradford.
- Simmel, G. (2010) 'The metropolis and mental life', in Bridge, G. and Watson, S. (Eds.): *The Blackwell City Reader*, Wiley-Blackwell, Malden, MA.