Mapping the innovation Landscape: A Network Analysis of R&D Collaborations in the Province of Alicante

Mariola Sánchez Romero
José A, Belso Martínez
María J. López Sánchez

UNIVERSITAS Miguel Hernández
Cátedra de Transformació del Model Econòmic

Xarxa Càtedres de Transformació del Model Econòmic
GENERALITAT VALENCIANA
Conselleria d'Hisenda i Model Econòmic

Universitat de València
Universitat Politècnica de València
Universitat d'Alacant / Universidad de Alicante
UJI Universitat Jaume I
UNIVERSITAS Miguel Hernández

1. Introduction

Successful innovation is a key source for regional development. Therefore, the number and quality of collaborations among the partners involved in the innovation system are crucial for achieving a high rate of technological change in a specific region [3]. In this sense, the importance of University-Industry Collaboration (UIC) has been widely recognized both in theory and in practice [8].

Through the various activities encompassed by UIC, university knowledge enhances the productivity of R&D and the innovative capacity of a company, leading to increased financial gains [7]. Additionally, for companies, gaining access to university expertise and facilities can accelerate the speed of new product development. By collaborating with an external partner, companies can share the costs and risks of R&D [6]. In particular, collaboration with universities and other research institutions provides companies with access to cutting-edge knowledge and advanced scientific facilities, serving as a significant complement to internal R&D efforts [4]. On the other hand, UIC is also beneficial for universities or technological centers as it allows researchers to learn from industrial practice and gain reputation within industry communities [1].

This piece of work explores the R&D collaboration network between public universities located in the province of Alicante (UMH and UA), research and etchnological centers (Fisabio, AITEX,..),including companies, or self-employed individuals. Given the importance of such collaborations, a network analysis appears as a suitable approach to examine the evolution of regional flows of collaborations among the different participants in the creation of R&D. Once we get the collaboration network, we yield the metrics associated with node's centrality and perform an ANOVA analysis to explore which factors (location and type of organization) are crucial to explain the "central" role of nodes in the collaboration net.

The importance of R&D collaborations, as well as their benefits, has been extensively studied in the literature. However, innovation collaboration is not without its challenges. In practice, the UIC process is often associated with significant costs. This is especially the case in emerging economies, where SMEs are mostly at the lower end of the industrial chain and may not be able to make substantial investments in R&D collaborations with universities [5]. It's not only SMEs that face difficulties in participating in UIC. Large companies can also be disadvantaged when collaborating with universities. [2] indicate that R&D collaboration with universities is characterized by a high level of uncertainty since it's difficult to project expected outcomes and timing. [1] further argue that academic research can be overly theoretical and somewhat irrelevant to commerce. Publicly traded large companies prioritize their short-term financial performance (Geyskens et al. 2002), suggesting that they may consider R&D collaboration with universities too risky if short-term financial benefits cannot be guaranteed. It is precisely these disadvantages that drive the need for different public policies to strengthen this type of activity, which, moreover, becomes decisive for the economic development of regions.

This report also have implications for policymakers and stakeholders involved in regional development. The insights into collaboration dynamics can guide the formulation of policies that support and strengthen collaborative activities, ultimately contributing to the economic development of the region.

The rest of the paper is organized as follows. Section 2 presents the methodology and data. Section 3 reveals the collaboration network's big picture and the evolution of the net over the years in the province of Alicante. ANOVA measures are introduced by Section 4. Finally, Section 5 offers the conclusions.

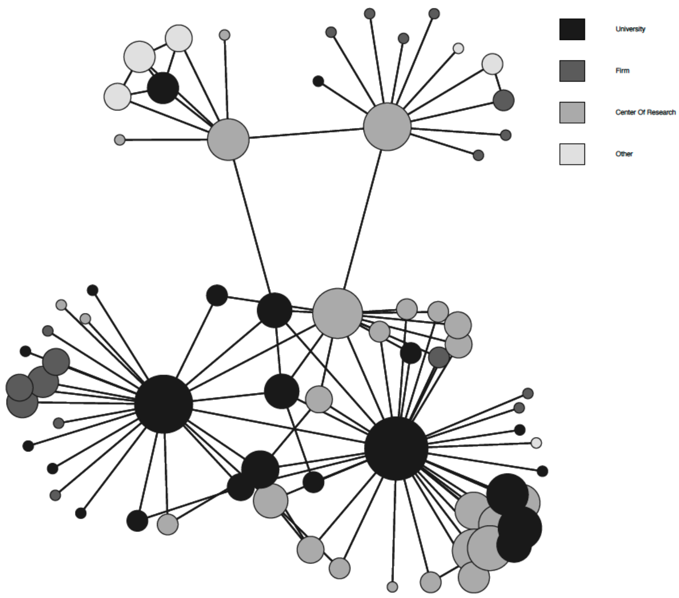
2. Methodology and data

In this study, we focus our attention on the collaboration network among organizations of different nature (research centers, universities, companies, or even individuals acting autonomously) that share ownership of an innovation registered with the OEPM with one or several other players. Given the that we have access to OEPM (Oficina Espanola de Patentes y Marcas)[1], we can complete the network for the two public universities UA, UMH, and the research centres, AITEX, FISABIO and INESCOP.[2] The following table 1 shows the data collected, where we filtered to focus on collaborations, and not in UMH patents' solo. Finally, we get 116 colaborations, that results in 195 collaborations due to the multiple (higher than 2 actors) participants in the patent.

Table 1: Total Collaborations

UMH	52
UA	31
FISABIO	16
AITEX	11
INESCOP	5

Figure 1: Network Collaborations from 1992 to 2023



3. The evolution of the network

Figure 1 shows the overall network collaboration. Firstly, we have categorized our net depending on the type of organization. We distinguish among University (type 1), Firms (type 2), Other research Centre (type 3) and Other (type 4, mainly independent professionals). In the figure, these type of organizations are differentiated by the tone of grey (see the legend). On the other hand, nodes size reveals the degree of the node. A higher node size represents higher fraction of nodes it is connected to. We can see that the two public universities have too many connections, therefore, they have high degree. Of course, there are other type of organizations with higer degree, that are other research centre, AITEX and FISABIO, and CSIC:

The big picture:

Table 2: Structural indicators of the network

Density	0.0563
Diameter	4
Average path length	2.8017
Transitivity	0.2591
Centralization	1.1076
Edge	136
Dyad	2415
Size	70

[1]To access to the OEPM: https://worldwide.espacenet.com/

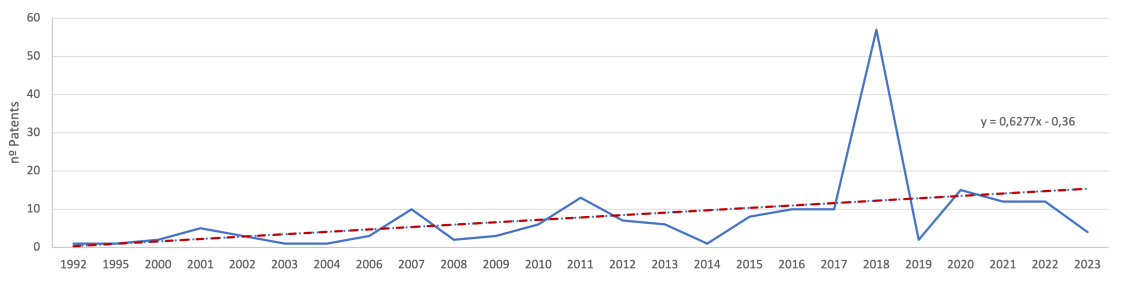
[2]We looked for others centres, as AIJU, that is also taken into consideration in the studio.

At first sight, the visualization and main descriptive statistics of Table 1 reveal interesting insights. The network density is 0.0563, meaning that 5.6% of all possible ties are activated out of 2415 potential relationships, while the average path length is 2.80. Centralization expands the concept of density by examining how cohesion is organized around specific focal points. In this case, the index value of 1.1076, which is above 1, indicates that the observed network is centralized rather than more dispersed and is likely to generate a center. The global clustering coefficient or transitivity value is 0.2591. Along with the low diameter value, these indicators may suggest a smooth circulation of information and resources in the collaboration network. A centralized network indicates that one or a few actors hold an important position. Highly central actors act as a resource in the network. In our study, as the network is built through the two universities of the province and some research centers, it was clear that the network visualization (Figure 1) shows the two egos (UA, UMH), and their alters. Furthermore, greater centralization means that information and resources flow through one or a few actors, potentially leading to greater efficiency. To increase the functionality of the network, the participation of key actors or decentralization is necessary. Therefore, this initial approach suggests that the universities and these provincial centers may play a significant role in the rapid dissemination of knowledge [9].

3.1. The evolution of the net

Given that Espacenet allows us to have information about the year of publication of patents, we can observe a great asymmetry of collaborations over the years. Figure 2 shows the evolution of the net.

Figure 2: Patents each years



The database shows us the first collaborations in 1992 and they extend to the present day, 2023. The red line establishes the total average of patents collected through Espacenet, which is 7.8 patents per year. As can be seen, from 2015, the number of patents is around that average, with the observed peak exceeding 40 patents being of great importance. Specifically, the figure for 2018 is striking, with 57 patents.

A simple regression performed with Excel shows a growing trend over the years (red dashed line in the graph). This leads us to affirm that the trend to collaborate in the province is increasing, and this fact is crucial to incentivize this activity. Strategies to promote collaboration may include fostering a community working environment, encouraging open communication, highlighting individuals' strengths, and investing in collaboration tools.

Collaboration, particularly at a regional level, has been shown to address problems that individual entities may not be able to solve on their own. It allows for the pooling of resources, knowledge, and skills, thereby enhancing problem-solving capabilities and fostering innovation. Moreover, collaboration can lead to increased efficiencies and cost savings, as resources can be shared and utilized more effectively.

4. ANOVA measures

The Analysis of Variance (ANOVA) is a statistical method that is used to examine differences among group means by comparing variances. It is a versatile tool that allows for the investigation of significant differences in the mean of a dependent variable, which is typically a numerical variable that we aim to explain, across different levels of an independent variable. These independent variables are categorical and are defined in factor levels. ANOVA assumes that the data is normally distributed, the variance among the groups should be approximately equal (homogeneity of variance), and the observations are independent of each other.

ANOVA can be used to study not only fixed effects but also interaction effects among factor levels of two independent variables. Interaction effects represent the combined effects of factors on the dependent measure. When an interaction effect is present, the impact of one factor depends on the level of the other factor. This means that the interpretation of the main effects is incomplete or misleading without considering the interaction effects.

The independent variables in ANOVA must be categorical (nominal or ordinal) variables, while the dependent variable must be a continuous (interval or ratio) level of measurement. ANOVA assumes that the data is normally distributed, the variance among the groups should be approximately equal (homogeneity of variance), and the observations are independent of each other[3].

In this study, we consider that our dependet variables are the centrality measures of nodes normally used in networks (Eigen Centrality, Betweeness Centrality and Closeness Centrality). On the other hand, we consider as an independent variables the type of organizations participating in the network and the territorial ubication of the entities.

Specifically, the **questions** that we aim to contrast with ANOVA are:

1. Is the effect of the type of organization a significant aspect for it to play a more central role in the network? Or is it to be close to other nodes?
2. Are centrality and closeness dependent on the spatial location of network entities?
3. Could we conclude that with a certain type of organization, different levels of centrality are achieved? And depending on the location?

The main dependent variables we want to examine are three linked to centrality metrics: Degree Centrality, Betweenness Centrality, and Closeness Centrality:

- Eigen centrality: Eigenvalue centrality, also known as eigenvector centrality or prestige score, is a measure of the influence of a node in a network. It assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. In other words, a node with high eigenvector centrality is connected to many nodes that themselves have high eigenvector centrality scores.

- Closeness centrality: in a connected graph, closeness centrality (or closeness) of a node is a measure of centrality in a network, calculated as the reciprocal of the sum of the length of the shortest paths between the node and all other nodes in the graph. Thus, the more central a node is, the closer it is to all other nodes.

- Betweenness centrality: is a measure of centrality in a graph based on shortest paths. For every pair of vertices in a connected graph, there exists at least one shortest path between the vertices such that either the number of edges that the path passes through (for unweighted graphs) or the sum of the weights of the edges (for weighted graphs) is minimized. The betweenness centrality for each vertex is the number of these shortest paths that pass through the vertex.

Figure 3 shows the big picture of the dependent variables.

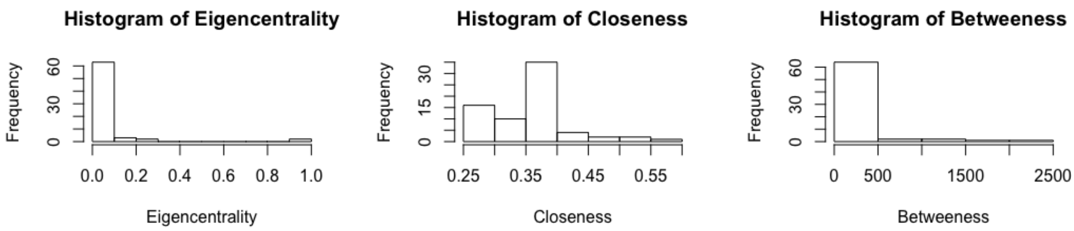
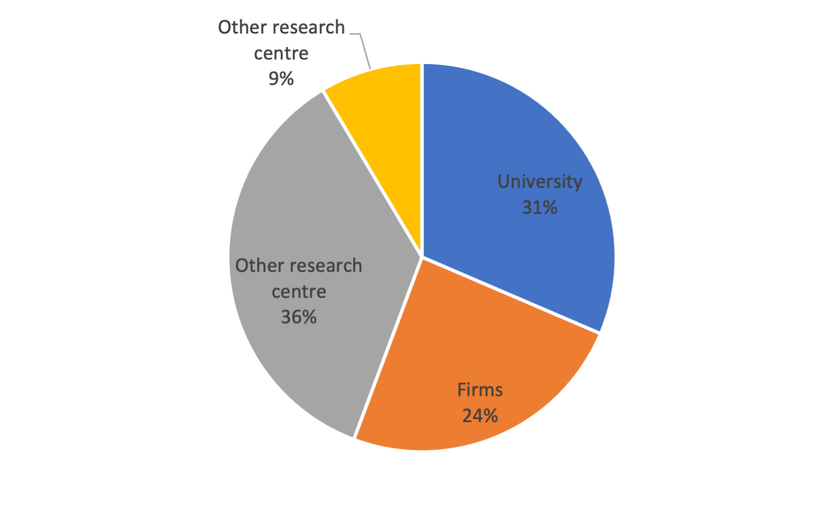


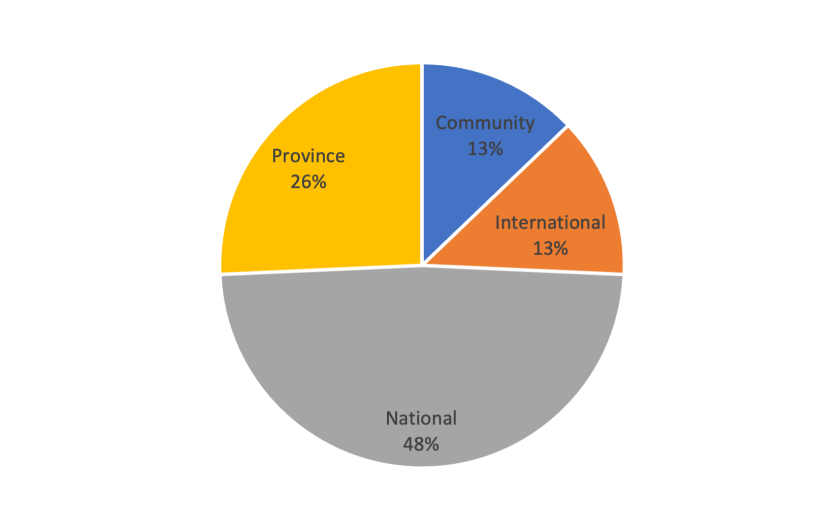
Figure 3

Our independent variables are mainly two: i) the type of organizations, where we distinguish among University, Firms, Other research centre (e.g., AITEX, INESCOP, FISABIO, CSIC...) and other (autonomous persons), and ii) we study the implications of the territorial ubication of entities as a factor that may determine how well are organizations connected. Territorial ubication of entities are split in four possibilities, bearing in mind that we focus firstly on the provincial level, community level. And lately, on entities located at national and international levels.

Our data of the independent variables shows the following information:



(a)



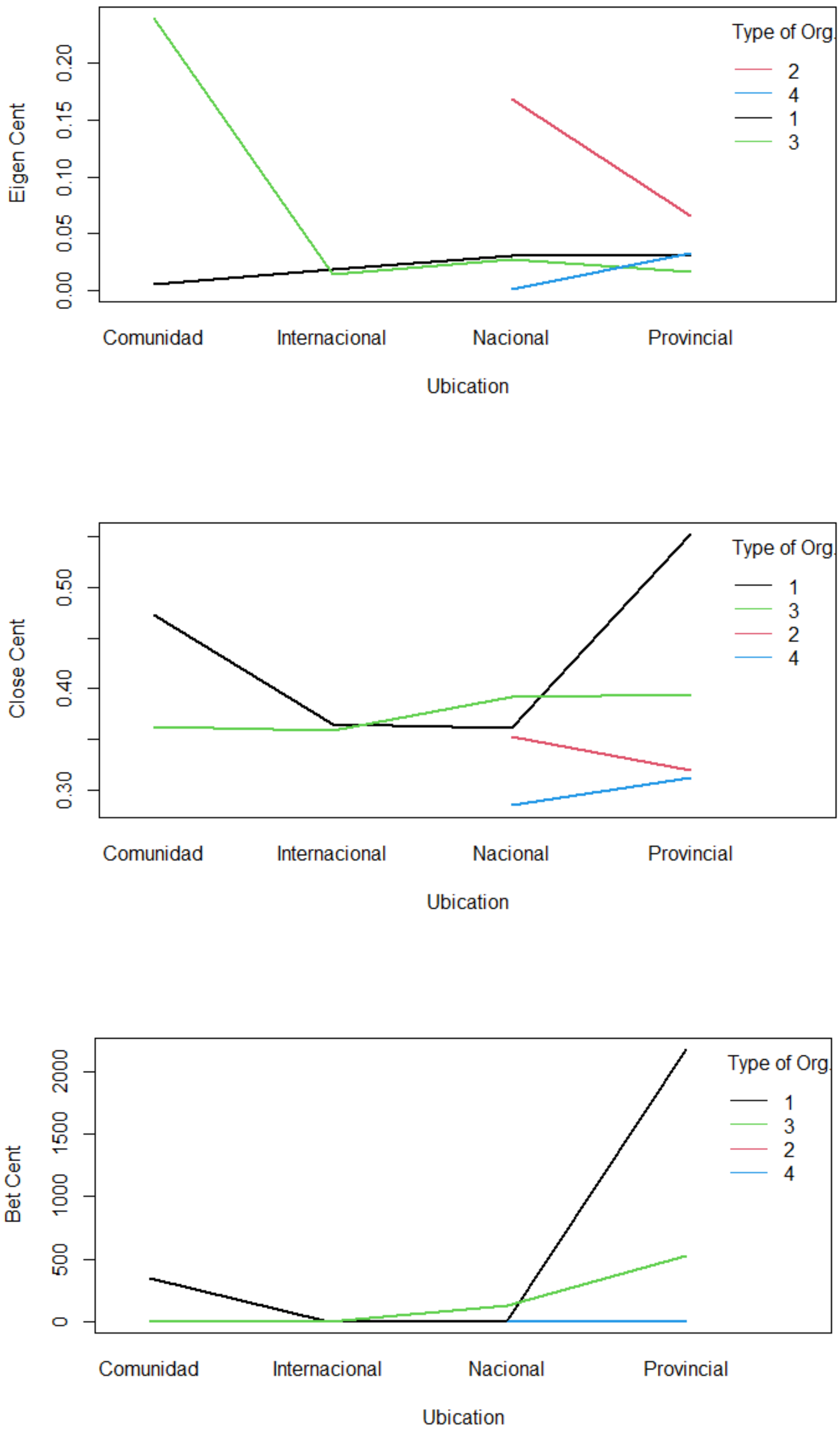
(b)

Figure 4: Pie graphs of type and ubication distributions, (a) and (b), respectively.

The next step to compute ANOVA is to check if the data meet the minimum assumptions for comparing variances. We then test normality (the data follows a normal distribution) and homocedasticity requirements (variances do not vary). For the first variable, Eigencentrality, Levene test reveal a $p-value = 0.4111$ for type of organizations, and a $p-value = 0.4138$ for ubication of organizations. Thus, homocedasticity is met. For the Closeneess variable, Levene test reveal a $p-value = 0.2506$ for type of organizations, and a $p-value = 0.0191$ for ubication of organizations, significant al 1%. Finally, the betweenness variable, Levene test reveal a $p-value = 0.3595$ for type of organizations, and a $p-value = 0.0657$ for ubication of organizations, significant al 5%. Normality checks are met by the Central Limit Theorem.

We are interested in fixed effects but also in the existence of iterations effects. In Figure 5, we plot the interaction effects between factors. A quick look shows as that lines cross for certain combinations of different levels for the two factors.

Figure 5



Recall that we distinguish among Type of organizations with this code: University (type

1), Firms (type 2), Other research Centre (type 3) and Other (type 4, mainly independent professionals). Furthermore, the interaction plots show mean values for each dependent variable. Thus, for example, it is easy to check that firms get higher eigen mean values at both national and provintial levels than the other types of entities. Therefore, one may expect that ANOVA test shows significant differences depending on type of entities and locations.

4.1. ANOVA Regressions

First, we compute ANOVA for the eigen values as a dependent variable. The regression shows that the type of organization and their locations are not significant to be well positioned in the network and to guarantee the potential connections in the net.

Dependent variable: eigen vector

	Dependent variable	**Dependent variable**
Predictors	*p*	*p*
type	0.750	0.753
location	0.466	0.479
Residuals		
type:location		0.754
Observations	70	70
R^2 / R^2 adjusted	0.040 / -0.019	0.058 / -0.049

* *p<0.05* ** *p<0.01* *** *p<0.001*

However, the figure is different if we focus on how closer is a node to all other nodes. Location does not explain the closeness score, however the fixed effect for the type reveals that the nature of entities is crucial for the closeness to other nodes.

Dependent variable: closeness vector

	Dependent variable	**Dependent variable**
Predictors	*p*	*p*
location	0.961	0.953
type	0.055	**0.041**
Residuals		
location:type		**0.013**
Observations	70	70
R^2 / R^2 adjusted	0.060 / 0.002	0.209 / 0.120

* *p<0.05* ** *p<0.01* *** *p<0.001*

Furthermore, the interaction effect of type and locations together is significant, meaning that a precise type of organization in a specific location play a more important role than other possible combinations. Organizations might find that the specific geographical location of entities involved in collaborations does not play a substantial role in the observed variations. The significant p-value (0.0413) for "type" indicates that the type of organization has a noteworthy impact on the dependent variable.

A Tukey analysis...

Dependent variable: Betweeness vector

	Dependent variable	Dependent variable
Predictors	*p*	*p*
type	0.321	0.292
location	**0.030**	**0.018**
Residuals		
type:location		**0.014**
Observations	70	70
R^2 / R^2 adjusted	0.140 / 0.087	0.275 / 0.193

* *p<0.05* ** *p<0.01* *** *p<0.001*

The "type" variable does not show a significant effect on the dependent variable, as its p-value (0.2918) is greater than the common significance level of 0.05. This fact suggest that the intrinsic characteristics of an organization might not be the sole driver of shortest paths. Therefore, organizations should focus on collaborative strategies that go beyond their organizational type.

The "location" variable has a significant effect on the dependent variable, with a p-value of 0.0182, suggesting that the spatial location of entities influences the dependent variable. The significant effect of spatial location implies that the geographical context plays a crucial role in shaping R&D collaborations. Organizations should consider local dynamics, industry clusters, and regional strengths when devising their collaboration strategies.

The interaction between "type" and "location" is also significant, with a p-value of 0.0136. This indicates that the combined effect of "type" and "location" is different from what would be expected by considering their individual effects. Since the type of organization alone does not have a significant impact, there might be opportunities for cross-sector collaborations. Organizations can explore partnerships with entities from different sectors, fostering a diverse and innovative R&D ecosystem.

In summary, the type of organization alone does not significantly influence the dependent variable, but both the spatial location and the interaction between type and location have significant effects. The interaction effect suggests that the relationship between the type of organization and the dependent variable varies across different spatial locations.

5. Conclusion

The study reveals a dynamic and evolving Research and Development (R&D) landscape in the province of Alicante. The collaboration network has witnessed substantial growth since 1992, indicating an increasing trend in collaborative activities among universities, research centers, companies, and independent entities.

University-Industry Collaboration (UIC) emerges as a critical catalyst for innovation and regional development. The study underscores the mutual benefits of UIC, with firms gaining access to advanced knowledge and facilities, while universities enhance their reputation and learn from industrial practices.

Despite the recognized benefits of R&D collaborations, challenges persist, particularly in cost implications. Small and Medium Enterprises (SMEs) in emerging economies and large companies face uncertainties and potential conflicts, necessitating strategic interventions and policy support to encourage their active participation in UIC.

Through network analysis, the study uncovers valuable insights into the collaboration dynamics. The examination of centrality measures highlights the pivotal roles played by certain organizations in facilitating the flow of information and resources. The degree and betweenness centrality metrics offer a nuanced understanding of organizational importance within the collaboration network.

The spatial location of entities is a significant factor shaping collaboration patterns. Analysis of Variance (ANOVA) results indicate that both the type of organization and territorial location exert a notable impact on centrality measures. Understanding these spatial dynamics can aid in tailoring strategies to foster collaboration among diverse entities.

Regression analysis points towards a growing inclination towards collaboration in Alicante, emphasizing the need for continued support and strategies to promote collaborative endeavors. This trend underscores the importance of initiatives that foster a collaborative environment, encourage open communication, and invest in tools that facilitate cooperation.

The findings of this study hold implications for policymakers and stakeholders involved in regional development. The insights into collaboration dynamics can guide the formulation of policies that support and strengthen collaborative activities, ultimately contributing to the economic development of the region.

In conclusion, the exploration of R&D collaborations in Alicante provides a comprehensive understanding of the network dynamics, challenges, and opportunities. By addressing these findings, stakeholders can actively shape an environment conducive to innovation, knowledge exchange, and sustained regional growth.

6. Acknowledgments

Financial support from the Chair GVA-UMH for the Transformation of the Economic Model is gratefully acknowledge.

References

[1] Ankrah, S. N., Burgess, T. F., Grimshaw, P., and Shaw, N. E. (2013). Asking both university and industry actors about their engagement in knowledge transfer: What single-group studies of motives omit. *Technovation*, 33(2-3):50–65.

[2] Caraça, J., Lundvall, B.-Å., and Mendonça, S. (2009). The changing role of science in the innovation process: From queen to cinderella? *Technological Forecasting and Social Change*, 76(6):861–867.

[3] Freeman, C. (1987). Technical innovation, diffusion, and long cycles of economic development. In *The Long-Wave Debate: Selected Papers from an IIASA (International Institute for Applied Systems Analysis) International Meeting on Long-Term Fluctuations in Economic Growth: Their Causes and Consequences, Held in Weimar, GDR, June 10–14, 1985*, pages 295–309. Springer.

[4] Frenken, K., Ponds, R., and Van Oort, F. (2010). The citation impact of research collaboration in science-based industries: A spatial-institutional analysis. *Papers in regional science*, 89(2):351–271.

[5] Handoko, F., Smith, A., and Burvill, C. (2014). The role of government, universities, and businesses in advancing technology for smes' innovativeness. *Journal of Chinese Economic and Business Studies*, 12(2):171–180.

[6] Martino, G. and Polinori, P. (2019). An analysis of the farmers contractual preferences in process innovation implementation: A case study in the italian poultry context. *British Food Journal*, 121(2):426–440.

[7] Siegel, D. S., Waldman, D. A., Atwater, L. E., and Link, A. N. (2003). Commercial knowledge transfers from universities to firms: improving the effectiveness of university–industry collaboration. *The Journal of High Technology Management Research*, 14(1):111–133.

[8] Stenbacka, R. and Tombak, M. (2020). University-firm competition in basic research and university funding policy. *Journal of public economic theory*, 22(4):1017–1040.

[9] Valente, T. W. (2010). *Social networks and health: Models, methods, and applications*. Oxford University Press.