



Endogenous growth factors and their empirical verification in the Colombian business context by applying fuzzy measurement techniques

Melva Inés Gómez-Caicedo¹ · Mercedes Gaitán-Angulo² · Anderson Quintero² · Jenny Paola Danna-Buitrago¹

Received: 8 March 2021 / Accepted: 30 August 2021 / Published online: 20 October 2021
© The Author(s) 2021

Abstract

This paper centers on the role and economic effect of the factors that affect the endogenous growth of organizations. To this end, it analyzes the case of Colombian companies, whose economy has been characterized in Latin America by maintaining a traditional economic policy and sectors that seek to maximize the opportunities offered by international markets. This study seeks to identify and analyze the factors that influence the endogenous economic growth of organizations and that allow them to be sustainable over time. For this purpose, adaptation algorithms based on fuzzy logics and supervised and unsupervised learning methods were used, generating an adaptation strategy that allows classifying and knowing the endogenous growth of companies in the Colombian context and identifying the most relevant aspects to take into account. Finally, it was found that most Colombian companies focus their business policies on strengthening human capital, followed by innovation, financial resources and, to a lesser extent, strategic alliances.

Keywords Endogenous growth · Adaptive algorithms · Organizational growth · Decision making · Colombia

1 Introduction

The dynamics of the markets, and particularly of the companies, has facilitated the development of processes that tend to strengthen the productive activity. Hence, several elements have been identified which, when used efficiently, promote the growth and competitiveness of products in the market, aspects that are considered relevant for sustainability in organizations, generating interest on the part of researchers and professionals in recent years. Its

study can be carried out endogenously (internal) and exogenously (external), which facilitates the analysis and generation of impact measures for the fulfillment of objectives.

This research addresses the analysis of internal conditions based on the concept of endogenous growth theory, identifying factors such as innovation, financial resources, human capital and alliances, which are determinants for their measurement and the generation of institutional and governmental policies.

Furthermore, technology and education-related elements are incorporated to strengthen the processes, as evidenced in the research conducted by Pan et al. [35], which indicates that in countries such as China, the number of higher education institutions at the provincial level has a positive impact on business innovation.

In other countries such as Vietnam, investments in physical and human capital and technological progress are determinants of economic growth in that nation, as evidenced by the research of Grillo and Nanetti [18], who indicate that in emerging economies it is essential to implement growth policies related to physical and human

✉ Melva Inés Gómez-Caicedo
migomezc@libertadores.edu.co

Mercedes Gaitán-Angulo
mercedes.gaitana@konradlorenz.edu.co;
mercedes@correo.ugr.es

Anderson Quintero
andersonq9716@gmail.com

Jenny Paola Danna-Buitrago
jenny.danna@libertadores.edu.co

¹ Fundación Universitaria los Libertadores, Bogotá, Colombia

² Fundación Universitaria Konrad Lorenz, Bogotá, Colombia

capital and to support innovation and development activities.

Yazgan and Yalçinkaya [49] have studied the effects of innovation and development investment variables in OECD countries, concluding that innovation investments are substantially sufficient to change the long-term economic growth performance and income levels of OECD countries.

Authors such as Fagiolo et al. [13] studied through the agent-based model, the nexus between finance and growth in industries that produce homogeneous goods and carry out innovation and development activities, finding that part of the financing for this innovation is obtained through resources provided by banks, it is concluded that banking activity has a positive impact on organizational growth; however, excessive financing can hinder growth.

The analysis of these conditions generates qualitative and quantitative information that, when analyzed at the sector or market level, reveals the needs, strengths and aspects to be improved of the companies in general and, therefore, the measures that should be adopted to achieve the objectives.

Currently, experiment data or records of a population possess information flows and distribution that are difficult to process [30], so conventional theoretical methods to establish correlations between variables have become obsolete due to the enormous effort they represent in execution, opening the possibility of generating computerized methods that apply techniques that mix the analytical precision provided by an individual and contemporary digital speed.

The above, in order to establish the categorical degrees that qualitatively measure the level of endogenous growth of the Colombian business sector (application of unsupervised intelligence), to subsequently identify the most relevant characteristics in the development of the Colombian industry (application of supervised intelligence); as well as to determine the specific bases from the DANE data that allow generating reliable and easily adaptable estimators to the evolution of endogenous growth. Likewise, to build an easily recognizable guideline for the entities.

The objective of the research is to identify and analyze the endogenous growth factors and their empirical verification in the Colombian business context through adaptation structures validated through statistical analysis. Therefore, the document is divided into five parts: the first one presents the conceptual framework of the endogenous growth theory; the second part identifies the variables of analysis for the basis of endogenous economic growth; the third part explains the case study of Colombia; the fourth part presents the methodology used and the results obtained; and finally, some conclusions of the research are presented.

2 Conceptualization of the endogenous growth theory

The economy evolves according to the needs of the markets that are required to position itself and prevail in the production and marketing of goods and services. Therefore, its results depend on the implementation of a series of measures to strengthen it.

Hence, the concept of economic growth is analyzed, which is conceived as a phase through which countries grow from the progress originated by technological improvement, investment and the adequate use of resources [44].

The basis of the neoclassical growth system, established by Solow [44], is based on exogenous criteria, being the basis of the model: neoclassical, Keynesian and competition theory.

Other economists suggest the importance of analyzing the measures adopted by companies, individuals and the state from an endogenous context to generate practices that lead to the country's economic growth. Thus, authors such as Arrow [1] and Romer [41] considered that production, being cumulative, will allow making use of different internal variables, generating productivity in companies and therefore, influencing the economic system, in order to promote better results for the country.

In this sense, the endogenous growth theory arises as a result of the importance given by authors such as Romer [40], Lucas [29], Arrow et al. [2], Rebelo [39], among others, to the analysis of factors related to knowledge, human capital and innovation [24, 36].

Other authors such as Laincz and Peretto [26] consider that the use of concepts related to economies of scale, product quality improvement and the promotion of creativity are required to give rise to endogenous long-term policies.

Thus, the endogenous growth theory is based on internal growth to improve economic development and as a consequence, the indicators that dynamize the activity of companies, economic sectors, the benefits that redound to consumers [28, 32] and the role of the policy implemented by the state to increase productivity [5].

Therefore, based on the endogenous growth theory, long-term models have been proposed that facilitate economic growth without causing or generating market failures [9]. Additionally, the generation of a social benefit may have an impact [15].

Consequently, growth cannot be attributable only to technological progress phenomena as Solow [44] stated, nor is it due only to responses to the incidence of exogenous variables [22], since the influence of both scenarios is important for the generation of competitive processes.

In this sense, according to Hernández [20], the main differences between the exogenous and endogenous economic growth models are: first, technological progress, since it is considered exogenous when technology is determined and adheres to the processes that are analyzed from the external context. On the contrary, it will be endogenous when elements are used to incorporate technological progress into fruitful activities, leading to greater productivity.

Secondly, Hernández [20] considers that physical capital is one of the variables analyzed in the exogenous growth model, while for the endogenous model, greater importance is given to human capital.

Finally, in the third place, according to Arrow [1], convergence for the exogenous growth model is determined by external conditions, while when the investment variable is taken into account in the increase in per capita income, reference is made to endogenous growth [20].

This situation leads to consider four of the main variables that influence the endogenous growth of the economy: innovation, human capital, financial resources and alliances, which will be explained below.

3 Variables of analysis for the basis of endogenous economic growth

The endogenous growth theory is based on the analysis and development of internal conditions that favor the economic development of companies and countries in general. Hence, several factors are considered for its measurement, such as innovation, human capital, financial resources and alliances. The theoretical foundations that will be used for the analysis of the case under study are presented below.

3.1 Innovation

Innovation is a necessary process for strengthening companies, which is influenced by several factors that are essential elements for economic growth [37].

Likewise, innovation has been considered a central element for economic, social and organizational growth in different areas of organizations. Its understanding is a challenge for academics and professionals who seek to understand how to incorporate improvements in processes, products and systems in organizations [31].

Its study and implementation in companies must introduce concepts based not only on the improvement or on the process developed, but also on the bases that relate it to the knowledge of each economic activity [38]. Hence, the business practices developed can incorporate elements that are favorable for the productivity of organizations.

3.2 Financial resources

The proper use of resources can be applied not only to physical or intangible resources. Also, the capital that allows the functioning of the activities developed by the company and the operability is part of the elements that can contribute to the growth of the economy.

Hope and Vyas [21] consider that it is essential to consider the way through which companies are financed, a situation that starts from the investment process, to the supply of resources required for its operation.

Authors such as Eniola and Entebang [12] indicate the importance of financial knowledge for the company's decision making and productivity, the direction of measures aimed at improving processes and products, and the management of resources that facilitate the organization's positioning in the market depend on it.

3.3 Human Capital

Human capital is considered as the set of skills, knowledge, abilities and attributes that people have, which can be translated into productivity [14] and is a crucial factor in measuring the ability of organizations to innovate [27, 45].

Authors such as Ehrlich and Pei [11] indicate that human capital can be perceived as skills between different generations, which allows achieving growth and sustainable development, all this based on the use of labor and capital. This interaction is fundamental in this new context, so it is recommended that innovative ecosystems for the development of technological entrepreneurship be strengthened.

3.4 Alliances

This process allows a company to generate joint activities that favor the interests of each of the stakeholders of the productive cycle, it is framed in the main aspects that denote the concept of alliance.

O'Dwyer and Gilmore [34] state that alliances allow access to the improvement of conditions that facilitate the exploitation of opportunities that arise in the markets. So, they can be considered not only to improve commercial activity, but also to favor internal processes.

4 A case study in Colombian companies

Latin America's poor performance and the region's lack of progress are associated with its low productivity, high informality, insufficient export diversification and growth,

which makes it difficult to create jobs and finance the growing demand for more and better public goods [23].

In this regard, there are different indicators for analyzing the economic context of countries. One of them is the Global Innovation Index (GII), which aims to capture the multidimensional facets of innovation and for this purpose ranks the world economies according to their innovation capabilities.

Colombia in particular ranked 68th among the most innovative economies in 2020, dropping 6 places compared to the 63rd place it reached in 2018, scoring below average in the pillars of human capital, knowledge and technology products, and creative products [17].

The above reflects the state of Colombian industry, which has presented difficulties in promoting technological development and generating innovation strategies [43].

Colombia is considered the third economy in Latin America, between 2009 and 2019 it registered a growth of 1.75% in GDP, going from 1.2 to 3.3% according to data from the World Bank [48]. In this sense, a positive behavior of the economy is evidenced, which implies challenges especially in areas that are influential for obtaining benefits that redound at an economic and business level.

It is evident that there is an important challenge where science, technology and innovation culture play a fundamental role in the reduction of social gaps, not only as transmitters of knowledge, but also as receivers, which allow their influence to be much more effective, in order to overcome exclusion [47].

Hence, Gómez, Hernando and Mitchell [17] indicate that the Colombian economy is undergoing a transformation process since 2014, which involves improving productive conditions, strengthening the productive apparatus, incorporating technological processes, among other endogenous aspects such as human capital, innovation, financial resources and alliances, which allow responding to the requirements of international markets and in which it is required to be competitive. Some of the areas of improvement and actions carried out are shown in Table 1.

5 Methodology

The growth theory established in the previous section indicates that economic evolution in Colombia is a highly volatile dynamic system that is governed by innumerable virtually independent variables; however, there is evidence of the existence of certain conditions in which endogenous evolution is favorable, despite the fact that the country's economic development strategies only focus on strengthening certain specific qualities.

Thus, identifying the ideal characteristics that have a direct impact on the endogenous growth of an economic entity would allow it to focus its resources optimally in order to develop evolutionary policies that have a greater impact on its profits.

Therefore, in order to identify the variables with the highest growth in a given period of time, this section will focus on explaining in detail the origin of the data used for the analysis of growth in Colombian industry and the experimental process implemented.

5.1 Entity studied

Prior to 2004, Colombia did not have periodic metrics to study and monitor technological capabilities, nor potential innovative developments with which to indicate characteristics that identify an entity as innovative [7].

The absence of indicators did not allow economic entities to identify the general capabilities of a company in constant endogenous growth and the positive repercussions it implies. Eventually, with the transition to globalization, improvement policies led to a regional need to establish growth themes that would compete with the most developed entities.

5.2 Data description

In 2007, the National Department of Statistics (DANE, for Spanish Departamento Nacional de Estadística) established an instrument for general monitoring of the characteristics of companies, which “serve as a reference in the process of measuring the processes of development and technological innovation” [7]. The objective of the data collection was in the first instance to generate in the industry, bottom-up policies aimed at a 4.0 revolution of technology in Colombia, thus facilitating the introduction of science in the business environment.

Due to the annual periodicity of the data provided by DANE and the complete monitoring used in the Colombian manufacturing sectors, it becomes a reliable and complete reference to perform descriptive studies from computational algorithms.

For the adjustment of the methods applied in this paper, the data were obtained from the survey of Technological Development and Innovation-EDIT-Industry carried out by DANE in 2016 and 2017, which consists of 510 variables, as well as 8377 and 7867 records, respectively, which is distributed by departments; the amount of data for 2017 of each department can be seen in the following table.

The variables are mixed between binary (Boolean), discrete numeric and continuous values, in addition, the survey questions are focused on evaluating the following 4 categories:

Table 1 Improvement areas. *Source:* Taken from the study of Gómez, Hernando and Mitchell [17]

Action Areas	Actions	Objectives to be achieved
Institutional environment	More commitment from the government and local authorities to Science, Technology and Innovation (ST&I)	Reorient the roles of the public entities that are part of the TC&I System
Financing	Increase public resources, especially to finance business innovation projects	Improve the process of allocation and distribution of resources for ST&I from the General Royalties System
Innovation culture	Make the company the center of the CT&I strategy	Promote entrepreneurship; Create a virtual and comprehensive open innovation platform that would integrate the innovation needs of companies in Colombia, including solutions provided by researchers, universities, individuals and other companies. In addition, it could be integrated with other open innovation networks in Colombia
Science and technology	Fund, support and connect research groups	Promote world-class sectoral and cross-sectoral research centers
Human capital	Increase the participation of basic sciences, design and engineering in higher education	Promote the incorporation of researchers in companies, and technicians and technologists in SMEs

Table 2 Distribution of data by department. *Source* Own elaboration

Departments	Total
Vichada	16
Casanare	18
Sucre	19
Córdoba	24
Cesar	29
Nariño	44
Magdalena	50
Quindío	53
Meta	55
Huila	62
Boyacá	72
Cauca	86
Tolima	103
Norte de Santander	115
Bolívar	127
Caldas	137
Risaralda	169
Atlántico	326
Santander	350
Cundinamarca	652
Valle del Cauca	1010
Antioquia	1707
Bogotá D.C.	2642

- Innovation
- Financial resources
- Human capital
- Alliances

5.3 Data processing

Because some variables did not have sufficient information to present a noteworthy characteristic in the study, the cleaning of the database focused on two actions: The first one concentrated on omitting variables with insufficient records for the study whose nature was quantitative (monetary contributions or unbounded values) and the second one was applied to ordinal qualitative variables, such as scales or Boolean responses (yes or no questions), which were autocompleted from the lowest scale of the criterion or negative responses, respectively. Integer values assigned to monetary quantities were discretized to maintain a uniform distribution across variable characteristics. After cleaning the database, the remaining variables used for the study were 213, which are distributed as follows:

As shown in Table 3, it is possible for a human to group each variable studied into categories and relate them individually in classes; however, establishing the dynamics of all the questions to measure the degree of innovation and also to identify the determining variables involved in the endogenous evolution requires a global analysis of the entire set of data, as well as establishing reliable and simple

Table 3 Distribution of variables. *Source* Own elaboration

Category	Number of variables
Innovation	29
Financial resources	32
Human capital	92
Alliances	60

metrics for any entity, since replicating an analysis of this type would require resources, time and human talent specialized in statistical analysis and data mining.

Likewise, there is no precedent in Colombia for establishing an annual self-correction guide based on DANE data, which would allow the identification of the main components based on the country's endogenous evolution. Although there is research in other developing countries on the characterization used to measure technological growth, the evidence shown to ensure that the sociopolitical and socioeconomic conditions of these territories are similar to those of Colombia is insufficient.

Therefore, it is necessary to establish specific bases adapted to the Colombian context in order to generate reliable estimators, easily adaptable to the permanent evolution of technologies, which establish an easily recognizable guideline for the entities.

5.4 Scales used to measure endogenous growth

For an adequate study of endogenous growth, it is necessary to establish a classification method based on the overall performance of the entities that relates the data obtained by the DANE survey with a value or label that is naturally identifiable by the companies and that generates a performance categorization as a final solution.

The literature indicates that nominal categorical scales guarantee better comprehension and understanding, since previous research focused on developing countries; Chitachoua et al. [6], Romero et al. [42] applied entrepreneurial performance studies, whose results were identified from scales based on nominal variables (low, medium, high). The results acquired in ascending entrepreneurs established a great understanding of and attachment to the evaluation scales.

Due to the above, the most convenient option to measure annually the progress of the Colombian business sector is from nominal labels; therefore, it was decided to establish a categorical measurement, which allows dividing the total number of entities into sets, which establish their level of growth.

5.5 Approach to the solution

Determining the metrics of comparison that will be used to divide the population into different categorical sets allows identifying the most relevant characteristics in endogenous development [25], for this A mechanics will be developed that combine the advantages of supervised and unsupervised learning by carrying out an empirical evaluation from two phases:

The first phase consists of solving a classification problem among the DANE data, in which it is required to

identify a number of k categories; these will be defined from a distance function between data, which will allow to establish relationships of closeness; therefore, each classification method depends on three factors: the distance function, the number of categories and the autocorrection strategy (algorithm). For the solution of this phase, it is necessary to establish a categorization or clustering method based on the multidimensional vector system analysis.

The second phase focuses on relating the categories obtained in phase 1 with the DANE data from participation analysis based on a hierarchy of variables; in particular, it seeks to determine which characteristics best sample the population, which allows to determine the characteristics that most influence endogenous growth.

6 Study description

6.1 First phase of the study

The first phase is established from a set of estimation strategies E_i with $i \in \{1, 2, \dots, n\}$, which provide a distribution of the data into k categories of the form B_i , then the set is sought:

$$T = \{B_i : B_i \sim T_k\}$$

T_k being an expected distribution of the data under a set of k categories. From the set T it is intended to make a set of estimators of the data that provide categorical characteristics, which represent the value of the growth of the evaluated company, being monitored as a dependent metric of the values used to adjust the algorithm and as a territorial model of measurement, allowing to accommodate the scales to the constant technological growth and restricting an overestimated or underestimated evaluation of the data.

The methods that will be implemented to generate the set of estimators T are adaptive classification algorithms [3] that refer to a special type of structures or strategies that do not require specific modifications in the process to be adjusted to a specific context.

Due to the nature of the data and the nominal characteristics of the labels that will be used as metrics to evaluate the endogenous growth of the companies, it is chosen to use methods whose application focuses on unsupervised learning methods and processes based on fuzzy sets of perception, since these clustering models are characterized by classifying based on the context in which the data are developed, an ideal structure if it is intended to implement metrics within a territory based on the same socioeconomic and sociopolitical structures.

Among the algorithms applied, the following were used:

6.2 K-means

A Euclidean metric is used to adjust a number of barycenters (Center of equal distribution of population density), in order to recognize the point with the smallest n-dimensional distance from the data [19]. Therefore, the k-means estimation problem is summarized by solving the following equalities:

$$\arg \min \left\{ \sum_{n=1}^k \sum_{X_i \in S_n} \|X_i - m_i\|^2 \right\}$$

The literature has established that the appropriate

number of classes k for which the result of the estimator is optimal is set at a scale of 5 classes, since, according to Viloría and Pineda [46], the optimal number of Clusters that measure endogenous growth in the database provided by DANE is between 3 and 6; however, a scale of 3 and 4 categories would not allow identifying with sufficient breadth companies that have a good technological development but barely establish growth policies.

Likewise, it is a broad enough scale to identify shortcomings among innovators’ requirements, which is supported by Romero et al. [42], Chiatchoua et al. 6 and Viloría and Pineda [46] and research that seeks to provide measures of the same nature.

The amount of information obtained on the number of categories used is not clear; however, to make the number of variants in the algorithm uniform, 5 classes will be used in all other algorithms, so the only parameters that will allow the estimator to vary will be the autocorrection strategies and distance metrics.

6.3 Agglomerative clustering

According to Gan et al. [16], the method seeks to establish hierarchies between elements, based on distance measures, which base the distinction and hierarchical position on the closeness they have with the other points of the set, so this algorithm provides multiple estimators as long as it is suitable for different distance metrics; in this case, two will be used:

Euclidean

Measure $d(x, y)$ is established from two points $x, y \in R^n$ and is defined as:

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \dots + (x_n - y_n)^2}$$

Manhattan

Measured is established from two points $x, y \in R^n$ and is defined as:

$$d(x, y) = \sqrt{|x_1 - y_1| + |x_2 - y_2| + |x_3 - y_3| + \dots + |x_n - y_n|}$$

The general objective of the model is to find a series of n points or clusters that are the closest under the distance functions.

6.4 Fuzzy-C

The concept of fuzzy logic clustering was established by Dunn [10] and generalized by Jim Bezdek in 1981. Consists of applying the Euclidean distance to group data into clusters according to the degree of belonging to the set [33]. It consists of an iterative self-correcting process whose solution is obtained by finding the appropriate parameters that minimize the equation.

$$\sum_{i=1}^n \sum_{j=1}^k U_{i,j}^m d(x_i, C_j)$$

The values $U_{i,j}^m$ represent the degree of belonging of the element x_i in the set j , k is the number of clusters, n the number of records or data, on the other hand, the element C_j is known as the center of the class or set and is a vector that shares dimension x_i and are the parameters that are adjusted iteratively to know the elements of each cluster.

6.5 Consensus of solutions

Once the set of estimators T has been obtained, a consensus opinion process will be performed, which for this purpose will be considered as an application $R^k \rightarrow R$ with an associated weight vector W of dimension k , with:

$$\sum_{j=1}^k w_j = 1$$

of which each weight, $w_j \in [0, 1]$, de modo que según Baez-Palencia et al. [4] el método Group Decision Making is defined as: se define como:

$$\text{Class}(b) = \sum_{j=1}^k w_j b_j$$

Here b_j is a categorical integer value scale defined for the company or entity. In order for the representative values of each estimation method to be equally considered and to be a fully interacting group, each of the values in the set W of weights will be considered equal.

6.6 Second phase of the study

For the second phase of the study, the global categorical scale of endogenous growth provided by the first phase of the study will be used to establish categorical relationships between the variables, using supervised learning methods that receive the vector data from the DANE survey as input and the scales from the first phase as output.

Specifically, the classification method to be applied is the decision tree method, according to Dey [8], which is a type of algorithm that groups attributes, classifying them according to their values from a binary branching system. The depth with which the tree is worked will ensure a much more accurate estimator.

7 Results

The values obtained by the clustering algorithms are mathematical processes based on the dimensional proximity of the data, and for these cases, the proximity was established for different geometries evaluated by metrics.

Therefore, the relationship between categories is established to patterns of proximity established by the distances chosen in the algorithms adapted to the Colombian context. With this in mind the following categories are defined:

- Category 1—Low: Companies whose levels of innovation, financial resources, human capital and alliances are kept at minimal amounts compared to the entities with the greatest influence on technological growth. A category 1 company maintains its main focus on the development of its economic activity and turns to new areas of technological expansion only when the situation requires it.
- Category 2—Low–Medium: These are entities similar to the previous case, they are mainly focused on developing their economic activity, but for this they

need to set minimum investment policies in the growth of innovative topics for the solution of difficulties.

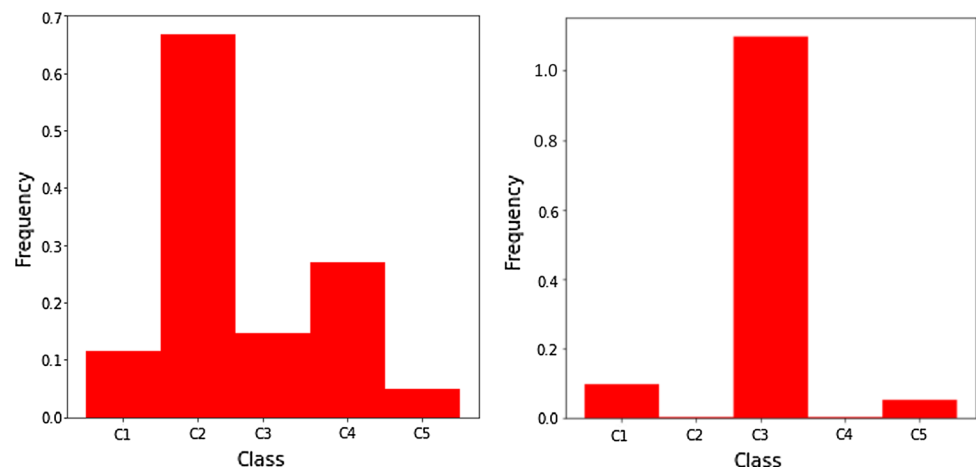
- Category 3—Medium: The companies in this category present an endogenous development and growth that is not limited by economic development; however, they are overshadowed by much more effective growth strategies presented in the subsequent categories.
- Category 4—High–Medium: Entities whose focus is on hiring skilled personnel, investments, alliances with skilled entities that support growth and resources to ensure they are at the forefront of technology in Colombia.
- Category 5—High: A company belonging to this category develops by applying efficient growth strategies, based on external contact networks, together with high investments in knowledge, research and increased dynamics that favor the company's profits, as well as the evolution of constant technological advances in Colombia.

In this way, based on the categories, the categories obtained by the algorithms were adapted and grouped when evaluated with the data, obtaining categories of the companies whose histogram achieves similar distributions. The histograms obtained by the k-means algorithm are shown in Fig. 1.

As can be seen, the categories in 2016 and 2017 have an approximate distribution, which reflects that Colombian companies have a low level of technological growth regardless of the year.

In 2016, there were a greater number of companies in category 2—Low–Medium, which means that they focused their practices on developing their economic activity. However, for 2017 they were placed in category 1, as the algorithm did not perceive significant growth in any of the four endogenous growth variables: innovation, financial resources, human capital and alliances.

Fig. 1 Distributions of categories in 2016 and 2017, respectively. *Source* Own elaboration



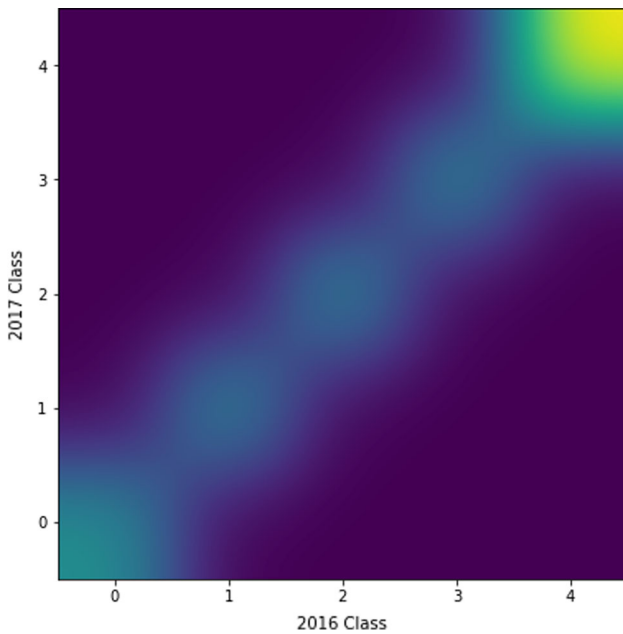


Fig. 2 Two-dimensional histogram of the proposed categories established by the k-means algorithm in 2016 and 2017. *Source* Own elaboration

In 2017, category 5 maintained its percentage of business participation with 3% of the 8192 companies studied. Likewise, category 4 remained in its participation with 9%, which indicates stability in policies oriented toward endogenous growth, and particularly focused on the factors studied.

On the other hand, the correlation between distributions of annual categories presents a high correlation as shown in Fig. 2, indicating that the algorithm maintains an invariant standard over time and that Colombia maintains a decreasing index.

Similarly, for the agglomerative clustering algorithms, the following distributions were obtained using Euclidean and Manhattan metrics:

As shown in Fig. 3, in 2016 there was a trend of higher endogenous growth in corporate policies; however, for the following year the practices remained stable.

In this way, it is possible to observe how the algorithms discriminate the behavior of the year 2017 with respect to 2016, which indicates that for the year 2017, the mechanics of technological growth presented the previous year progressively were forgotten or became obsolete, standardizing what was considered innovative and standardizing improvement strategies, a fact that is consistent with the themes of business evolution, since technology is constantly growing, leaving outdated the little progressive policies that do not suit and standardizing the advances that became distinctive in past times.

Finally, applying Fuzzy-C on the data, the following distributions are obtained, which characterize the data from the distributions shown in Fig. 4.

For the distributions obtained from fuzzy clustering, the results show how more than 60% of the data tend to have higher endogenous growth, due to the similarity found when comparing the data globally. On the other hand, the distribution in 2017 maintains a comparative standard with the previously used methods, in which the highest participation of the entities has a medium–low category maintaining the trend that had been seen applying the other algorithms.

Once the estimators were obtained from the methods explained above, an expert judgment was generated with the results obtained in the categories, with the objective of generating the decision tree, with which the hierarchical systems between the variables will be established to

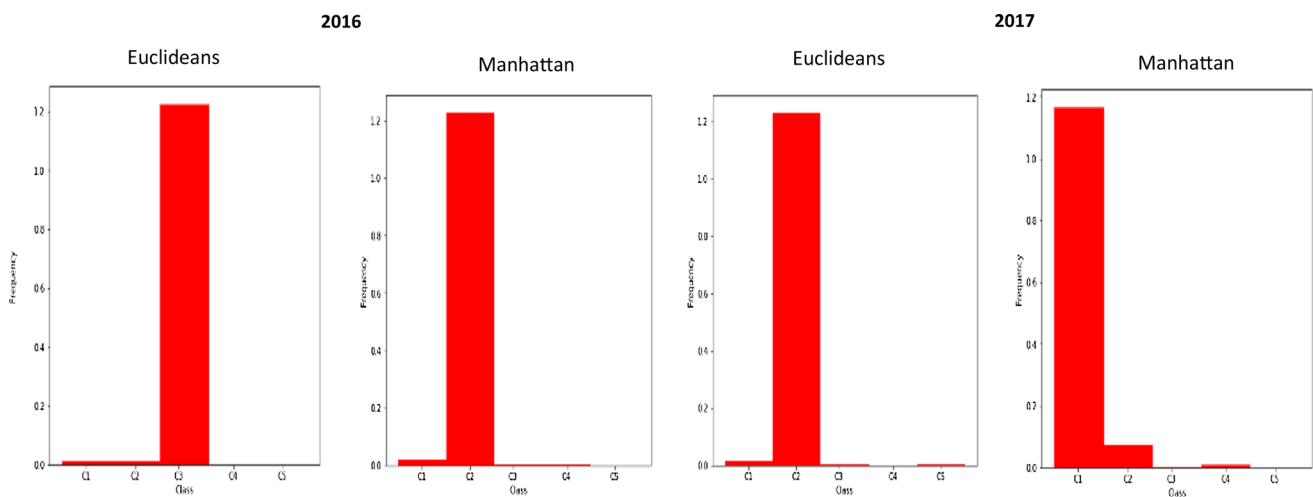


Fig. 3 Distributions of categories in 2016 and 2017 using Euclidean and Manhattan metrics, respectively. *Source* Own elaboration

Fig. 4 Distributions of categories in 2016 and 2017 from Fuzzy-C clustering. *Source:* Own elaboration

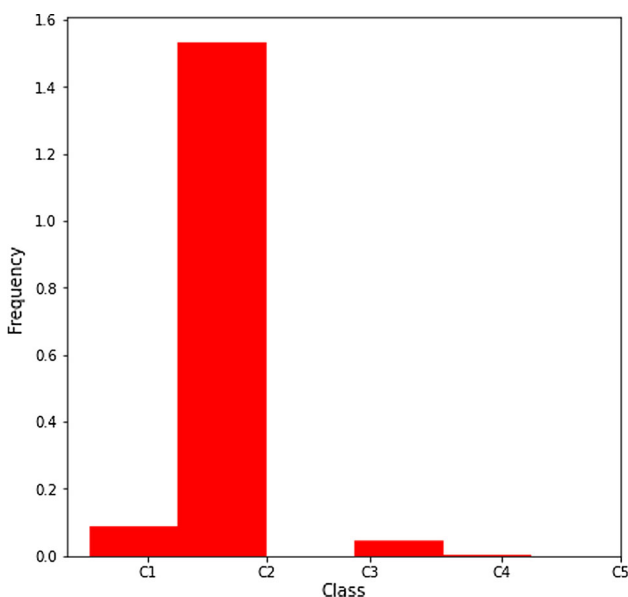
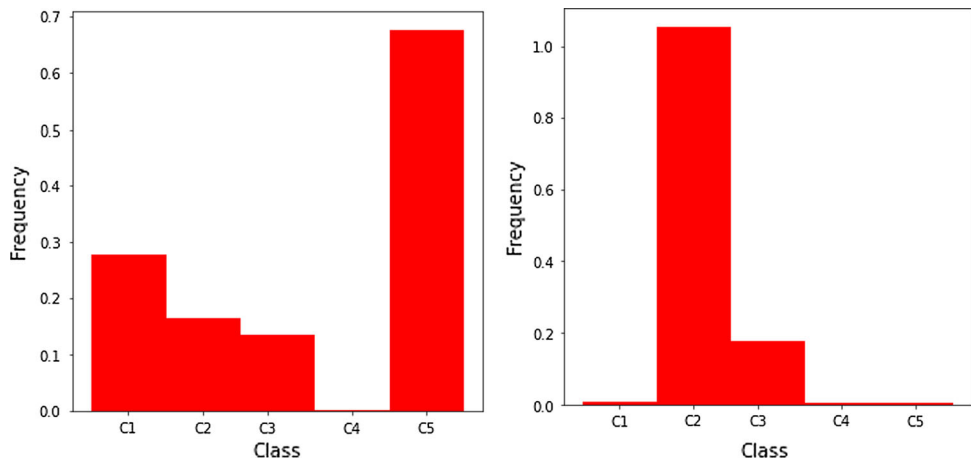


Fig. 5 Category distributions in 2017 using Group Decision Making from the four established methods. *Source* Own elaboration

determine the most determinant characteristics in the endogenous growth.

Figure 5 presents the results of the expert judgment and Fig. 5 contains the first two branches of the decision tree used to relate the categories to the companies (Fig. 6).

In this sense, the first branches of the trees present the variables that allow to obtain a better sampling of the population from the categories. Likewise, the copyright information system has a greater relationship with the established categories of endogenous growth.

According to the results obtained, there is a strong relationship between endogenous growth measures and the elements that are required to be productive internally, a situation that companies take into account for the implementation of their policies.

Likewise, the influence that the Regional Productivity Centers have on the generation of endogenous growth in the companies is related to the progress of the factors: innovation, financial resources, human capital and alliances; normally used for their development, which contributes to the improvement of the entrepreneurial dynamics.

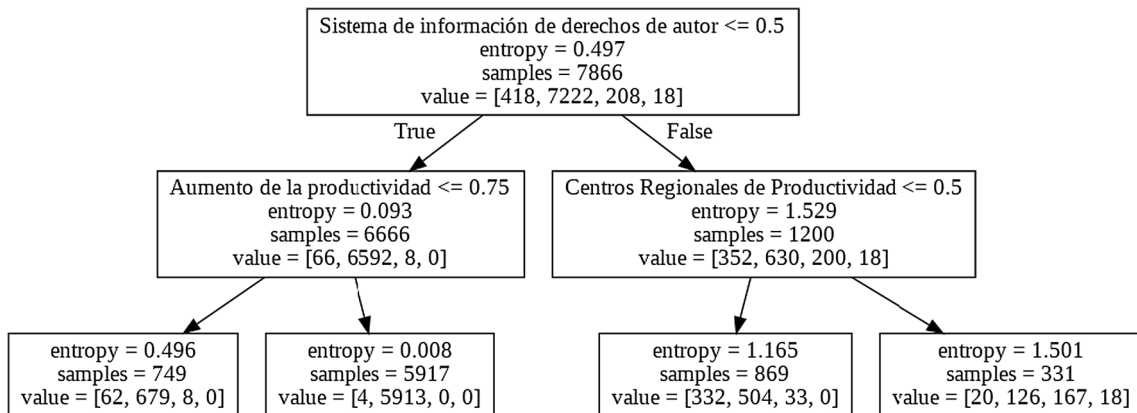


Fig. 6 First initial branches of the decision tree from the global estimator for the year 2017. *Source* Own elaboration

This analysis can be deepened as more complex trees are generated per variable analyzed. The maximum depth is 14 degrees and requires taking into account all the characteristics that support endogenous growth, making all the elements studied have a participation in the categories.

8 Conclusions

The study concludes that innovation is supported by human capital, with the departments of Antioquia, Atlántico, Bogotá, Bolívar, Boyacá, Caldas, Cauca, being the ones that have used the most business policies around this factor in favor of their endogenous growth.

In Colombia, the Department of Nariño registered the highest score in the financial resources factor, which indicates that it is one of the elements that presents deficiencies in the use of its productive capital.

Similarly, in the alliances factor, the departments with the highest scores were Quindío and Vichada; in which case the results indicate that this is a variable that can be enhanced in other regions.

On the other hand, thanks to the information provided by the survey conducted by DANE, it is clear that the endogenous progress of a developing country like Colombia requires better structuring in the areas of registration and control, since only 41% of the questions answered by the companies had enough information to generate a territorial comparison metric, this together with the volatile dynamics of growth that generate two main problems for the endogenous evolution of the territorial entities: first, the impossibility of establishing theoretical baselines to be applied to adjust analytical models for a specific context. Second, lack of knowledge of the general environment in which they develop and lack of precision in establishing improvement policies based on global references.

The aforementioned contradictions make the algorithms applied in this work ideal, since:

First, the DANE questions with which the algorithms were adapted do not focus on questioning exact quantities, but rather on establishing answers from labels or Boolean value questions, sufficiently general characteristics that facilitate the response by the entities and the precision of this, at the cost of possessing less accuracy.

Second, the more or less complex study techniques generate more flexible strategies of adjustment to real data, ideal aspects in contexts whose characteristics may have atypical behaviors.

Third, the ease of application and simplicity of the implemented metrics allow them to be understood by more experienced entities or those that are currently entering the Colombian business context.

9 Practical conclusions and future research

From the above, this study can be replicated to other Latin American economies and complemented by structuring export and import models provided by analytical models such as graph theory.

On the other hand, to identify under what conditions the endogenous growth of Latin America in the business context could become competitive with more developed countries such as Asia Pacific based on simulations generated by neuro-fuzzy detection systems.

This study generates great contributions for the support of public policies in the years after 2021, since it allows identifying and monitoring the entrepreneurial development of innovative projects in comparison with the industrial standard, encouraging the main characteristics that lead to endogenous growth, which leads to greater economic benefit and self-sustainable development.

Funding Regarding the funds used, they state that the research that led to these results received partial financial support from the Fundación Universitaria los Libertadores and Fundación Universitaria Konrad Lorenz.

Declarations

Conflicts of interest The authors declare that they have no potential conflicts of interest in the preparation of the manuscript; likewise, it is declared that open access databases were used for the development of the research.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Arrow KJ (1969) Classificatory notes on the production and transmission of technological knowledge. *Am Econ Rev* 59(2):29–35
2. Arrow KJ, Cline WR, Maler KG, Munasinghe M, Squitieri R, Stiglitz JE (1996) *Intertemporal equity, discounting, and economic efficiency*. Cambridge University Press, Cambridge, UK, New York and Melbourne, pp 125–144
3. Athey S (2019) The impact of machine learning on economics. In: *The economics of artificial intelligence*. University of Chicago Press, pp 507–547

4. Baez-Palencia D, Olazabal-Lugo M, Romero-Muñoz J (2019) Toma de decisiones empresariales a través de la media ordenada ponderada. *Inquietud Empresarial* 19(2):11–23
5. Bhattacharyya C, Gupta MR (2020) Unionised labour market, environment and endogenous growth. *Int Rev Econ Financ*. <https://doi.org/10.1016/j.iref.2020.10.005>
6. Chiatchoua C et al (2016) Construcción de una herramienta para medir la innovación tecnológica en las Pymes de la Región XI Texcoco. *Paradigma Económico* 8(1):51–76
7. Dane (2020) Encuesta de desarrollo e innovación tecnológica (EDIT). <https://www.dane.gov.co/index.php/estadisticas-por-tema/tecnologia-e-innovacion/encuesta-de-desarrollo-e-innovacion-tecnologica-edit>
8. Dey A (2016) Machine learning algorithms: a review. *Int J Comput Sci Inf Technol* 7(3):1174–1179
9. Dinopoulos E, Thompson P (2000) Endogenous growth in a cross-section of countries. *J Int Econ* 51(2):335–362. [https://doi.org/10.1016/S0022-1996\(99\)00015-X](https://doi.org/10.1016/S0022-1996(99)00015-X)
10. Dunn JC (1973) A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *J Cybern* 3(3):32–57. <https://doi.org/10.1080/01969727308546046>
11. Ehrlich I, Pei Y (2020) Human capital as engine of growth—the role of knowledge transfers in promoting balanced growth within and across Countries. *Natl Bur Econ Res*. https://doi.org/10.1162/adev_a_00155
12. Eniola AA, Entebang H (2017) SME managers and financial literacy. *Glob Bus Rev* 18(3):559–576. <https://doi.org/10.1177/0972150917692063>
13. Fagiolo G, Giachini D, Roventini A (2020) Innovation, finance, and economic growth: an agent-based approach. *J Econ Interact Coord* 15(3):703–736
14. Fulmer IS, Ployhart RE (2014) “Our Most Important Asset” a multidisciplinary/multilevel review of human capital valuation for research and practice. *J Manag* 40(1):161–192. <https://doi.org/10.1177/0149206313511271>
15. Galindo-Martín MÁ, Méndez-Picazo MT, Castaño-Martínez MS (2016) Crecimiento, progreso económico y emprendimiento. *J Innov Knowl* 1(1):62–68. <https://doi.org/10.1016/j.jik.2016.01.006>
16. Gan G, Ma C, Wu J (2020) Data clustering: theory, algorithms, and applications. SIAM, Philadelphia
17. Gómez HJ, Mitchell D (2014) Innovación y emprendimiento en Colombia—Balance, perspectivas y recomendaciones de política: 2014–2018. *Global Innovation Index 2020 (GII)*
18. Grillo F, Nanetti R (2020) Innovation and democracy: the twin paradoxes. *Area Dev Policy* 5(3):233–255. <https://doi.org/10.1080/23792949.2020.1777880>
19. Gupta MK, Chandra P (2020) An empirical evaluation of K-means clustering algorithm using different distance/similarity metrics. In: *Proceedings of ICETIT 2019*. Springer, pp. 884–892
20. Hernández C (2002) La teoría del crecimiento endógeno y el comercio internacional. *Cuadernos de estudios empresariales* 12:95
21. Hope OK, Vyas D (2017) Private company finance and financial reporting. *Account Bus Res* 47(5):506–537
22. Ireland PN (1994) Supply-side economics and endogenous growth. *J Monet Econ* 33(3):559–571. [https://doi.org/10.1016/0304-3932\(94\)90043-4](https://doi.org/10.1016/0304-3932(94)90043-4)
23. Jiménez A, Geldes C (2019) Los desafíos de la innovación en Latinoamérica. *J Technol Manag Innov* 14(4):3–5. <https://doi.org/10.4067/S0718-27242019000400003>
24. Kaldor N (1957) A model of economic growth. *Econ J* 67(268):591–624. <https://doi.org/10.2307/2227704>
25. Karlos S, Aridas C, Kanas VG, Kotsiantis S (2021) Classification of acoustical signals by combining active learning strategies with semi-supervised learning schemes. *Neural Comput Appl*. <https://doi.org/10.1007/s00521-021-05749-6>
26. Laincz CA, Peretto PF (2006) Scale effects in endogenous growth theory: an error of aggregation not specification. *J Econ Growth* 11(3):263–288. <https://doi.org/10.1007/s10887-006-9004-9>
27. Lenihan H, McGuirk H, Murphy KR (2019) Driving innovation: public policy and human capital. *Res Policy* 48(9):103–791. <https://doi.org/10.1016/j.respol.2019.04.015>
28. Li Z, Chu Y, Gao T (2020) Economic growth with endogenous economic institutions. *Macroecon Dyn* 24(4):920–934. <https://doi.org/10.1017/S1365100518000536>
29. Lucas RE Jr (1988) On the mechanics of economic development. *J Monet Econ* 22(1):3–42. [https://doi.org/10.1016/0304-3932\(88\)90168-7](https://doi.org/10.1016/0304-3932(88)90168-7)
30. Meyer G et al (2014) A machine learning approach to improving dynamic decision making. *Inf Syst Res* 25(2):239–263. <https://doi.org/10.1287/isre.2014.0513>
31. Montalvo C (2006) What triggers change and innovation? *Techovation* 26(3):312–323. <https://doi.org/10.1016/j.techovation.2004.09.003>
32. Nguyen NT (2020) Endogenous economic growth: the arrow-Romer theory and a test on vietnamese economy. *Wseas Trans Bus Econ*. <https://doi.org/10.37394/23207.2020.17.37>
33. Nwadiugwu MC (2020) Gene-based clustering algorithms: comparison between Denclue, Fuzzy-C, and BIRCH. *Bioinform Biol Insights* 14:1177932220909851. <https://doi.org/10.1177/1177932220909851>
34. O’Dwyer M, Gilmore A (2018) Value and alliance capability and the formation of strategic alliances in SMEs: the impact of customer orientation and resource optimisation. *J Bus Res* 87:58–68. <https://doi.org/10.1016/j.jbusres.2018.02.020>
35. Pan X, Gao Y, Guo D, Cheng W (2020) Does higher education promote firm innovation in China? *Sustainability* 12(18):7326. <https://doi.org/10.3390/su12187326>
36. Petrakis PE (2020) Entrepreneurship. In: *Theoretical Approaches to economic growth and development*. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-50068-9_8
37. Popa IL, Preda G, Boldea M (2010) A theoretical approach of the concept of innovation. *Managerial Challenges of the Contemporary Society*. Proceedings, 151
38. Quintane E, Mitch Casselman R, Sebastian Reiche B, Nylund PA (2011) Innovation as a knowledge-based outcome. *J Knowl Manag* 15(6):928–947. <https://doi.org/10.1108/1367327111179299>
39. Rebelo S (1991) Long-run policy analysis and long-run growth. *J Polit Econ* 99(3):500–521. <https://doi.org/10.1086/261764>
40. Romer C (1986) Spurious volatility in historical unemployment data. *J Political Econ* 94(1):1–37. <https://doi.org/10.1086/261361>
41. Romer PM (1994) The origins of endogenous growth. *J Econ Perspect* 8(1):3–22. <https://doi.org/10.1257/jep.8.1.3>
42. Romero MC, Rebori A, Camio MI (2010) Un índice para “medir” el nivel de innovación tecnológica en empresas intensivas en el uso de tecnología. *INMR-Innov Manag Rev* 7(1):03–20
43. Rubiano Riaño CA (2020) Revisión de los factores que limitan la innovación en las empresas
44. Solow RM (1956) A contribution to the theory of economic growth. *Quart J Econ* 70:655–94. <https://doi.org/10.2307/1884513>
45. Teixeira AA, Tavares-Lehmann AT (2014) Human capital intensity in technology-based firms located in Portugal: does foreign ownership matter? *Res Policy* 43(4):737–748. <https://doi.org/10.1016/j.respol.2014.01.001>
46. Vilorio A, Pineda OB (2019) Improvements for determining the number of clusters in k-means for innovation databases in SMEs. *Proc Comput Sci* 151:1201–1206. <https://doi.org/10.1016/j.procs.2019.04.172>

47. Williams LD, Woodson TS (2019) Enhancing socio-technical governance: targeting inequality in innovation through inclusivity mainstreaming. *Minerva* 57(4):453–477. <https://doi.org/10.1007/s11024-019-09375-4>
48. World Economic Forum. The global competitiveness report 2020. Available at: <https://www.weforum.org/reports/the-global-competitiveness-report-2020>
49. Yazgan Ş, Yalçinkaya Ö (2018) The effects of research and development (R&D) investments on sustainable economic growth: evidence from OECD countries (1996–2015). *Rev Econ Perspect* 18(1):3–23. <https://doi.org/10.1515/revecp-2018-0001>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.