



# New Insights on the Allocation of Innovation Subsidies: A Machine Learning Approach

Mónica Espinosa-Blasco<sup>1</sup> · Gabriel I. Penagos-Londoño<sup>2</sup> · Felipe Ruiz-Moreno<sup>3</sup> ·  
María J. Vilaplana-Aparicio<sup>4</sup>

Received: 22 February 2022 / Accepted: 25 February 2023  
© The Author(s) 2023

## Abstract

Gaining more insights on how R&D&i subsidies are allocated is highly relevant for companies and policymakers. This article provides new evidence of the identification of some key drivers for companies participating in R&D&i project selection processes. It extends the existing literature by providing insight based on sophisticated, accurate methodology. A metaheuristic optimization algorithm is employed to select the most useful variables. Their importance is then ranked using a machine learning process, namely a random forest. A sample of 1252 cases of R&D&i subsidies is used for more than 800 companies based in Spain between 2014 and 2018. The study contributes by providing useful knowledge into how the value of received subsidies are associated with some firm characteristics. The findings allow the implementation of transparent public innovation policies and the reduction of the gap between the aspects that are considered important and those that actually determine the destination of these subsidies.

**Keywords** Innovation · Subsidies · R&D&i strategy · Valencian companies · Random forest

---

✉ Mónica Espinosa-Blasco  
monica.espinosa@ua.es

Gabriel I. Penagos-Londoño  
penagosi@javeriana.edu.co

Felipe Ruiz-Moreno  
felipe.ruiz@ua.es

María J. Vilaplana-Aparicio  
maria.vilaplana@ua.es

<sup>1</sup> Department of Financial Economics and Accounting, University of Alicante, Alicante, Spain

<sup>2</sup> Department of Economics, Pontificia Universidad Javeriana, Bogotá, Colombia

<sup>3</sup> Department of Marketing, University of Alicante, Alicante, Spain

<sup>4</sup> Department of Communication and Social Psychology, University of Alicante, Alicante, Spain

## Introduction

Research and development (R&D) is an essential element of advanced societies because of its contribution to tackling social challenges (Mei et al., 2020; Wanzenböck & Frenken, 2020) such as climate change and population ageing. Innovation is among the 17 United Nations (UN) Sustainable Development Goals (SDG). Specifically, SDG number nine refers to sustainable investment in infrastructure and innovation, two key drivers of economic growth and development. Therefore, this matter is of concern to public administrations around the world (Casanova et al., 2018), which must develop and implement public R&D policies. In fact, innovation policies by public administrations aiming to increase firms' R&D investment is a topic that has received increasing interest and attention in the academic literature (Dimos & Pugh, 2016; González & Pazó, 2008; Kim et al., 2021).

Innovation policy consists of a set of policy instruments, many financial in nature. R&D policies are a subset belonging to innovation policies, therefore although R&D policy overlaps considerably with innovation policy, the latter is much broader, including commercialization policy and various demand-side policies for innovation (Martin, 2016). These R&D policies stem from a need to counter market failures (Arrow, 1962a; Heijs, 2002) and resolve the problems of appropriability (Leiponen & Byma, 2009) faced by companies. These policies are implemented using a diverse set of tools such as non-refundable aid, tax incentives, and subsidized loans. One of the more traditional R&D policy tools consists of R&D subsidies to firms. The findings from the literature are ambiguous as to whether public funding is a substitute for, or a complement to, private R&D funding (David et al., 2000). Tax incentives is another R&D policy instrument that has received much attention from economists. They are effective in increasing R&D intensity (Bloom et al., 2002).

Such is the importance of R&D that the major world powers allocate substantial funds to this area, despite differences in their scope and mechanisms. According to the most recent data for the year 2020 (Eurostat, 2022), South Korea leads the ranking of R&D investment, allocating 4.64% of its GDP to this area. South Korea is followed by Japan, with 3.2%, and the USA with 3.08%. Although Europe has tried to improve its position, its allocation remains below 2.32% of its GDP. Within Europe, Sweden (3.53%), Belgium (3.48%), Austria (3.20%), and Germany (3.14%) lead in R&D investment, far outspending other countries, including Spain (1.41%).

The scientific community has also shown an interest in this issue, and there is a rich literature on public R&D policies (Jugend et al., 2020). Previous research has shown the need for public administrations to develop tools that drive innovation (Chapman & Hewitt-Dundas, 2018; Kang & Park, 2012; Torregrosa-Hetland et al., 2019) to compensate for the associated uncertainty (Dosi, 1988; Hao et al., 2019) and risks (Chalioiti & Serfes, 2017; Klette et al., 2000). This support is essential to incentivize companies and promote collaboration between firms (Guisado-González et al., 2016).

Despite the importance of R&D policies, very few studies have examined the factors that help companies access public funds. The effect of public R&D

policies has been extensively studied (Afcha, 2012; Álvarez-Ayuso et al., 2018; González & Pazó, 2008; Guellec & De La Potterie, 2003; Vanino et al., 2019), as has its effectiveness (De Jorge & Suárez, 2011). However, only a small portion of the literature examines the factors that determine a company's success in obtaining public support. Therefore, the general objective of this study is to analyze the characteristics that help companies to get higher values of public R&D funding. Accordingly, 807 companies located in the Valencian Community (Spain) that received public funding for R&D projects from 2014 to 2018 were analyzed.

The method used in this study entails a combination of advanced econometric techniques, which are a genetic algorithm (GA) and a machine learning process, named random forest (RF) model. There are clear advantages of using this combination of these procedures. The GA allows us to select the optimal set of factors to improve the overall accuracy of the model, lessen the computational burden, and ensure that the final predictor subset leads to a more intuitive and interpretable model. Furthermore, the machine learning process implemented, a RF model, is a nonparametric ensemble learning algorithm developed by Breiman (2001) that shows high accuracy in the analysis of high-dimensional data, such as classification, regression, and other tasks based on a number of decision trees consisting of a number of decision nodes. Methods to measure variable importance in RF are better researched than in other machine learning methods (Grömping, 2015).

The remainder of the paper is structured as follows. First, a review of the scientific literature relevant to the area under investigation and research objectives are provided. In the "Method" section, the variables used, and a description of the methodology employed are provided. Then, our findings are outlined in the "Results" section, followed by the Discussion, where the managerial and policy implications together with the limitations of the study are shown. Finally, the "Conclusion" section presents the most important contributions of the research.

## Literature Review

In the current context of knowledge-based economies, innovation is fundamental, given its contribution to the creation of high-productivity jobs. It is broadly agreed that research and development (R&D) is one of the main factors enabling companies to increase their productivity (Acharya & Xu, 2017; Shinkle & Suchard, 2019; Zhang & Guan, 2018). For instance, Bravo-Ortega et al. (2014) found that companies that invest in R&D are more likely to export, increase their productivity, and complement private resources with public funding where possible.

The spread of innovation policies around the world reflects the importance of encouraging investment in innovation via R&D spending by private firms. The huge importance of public support for R&D stems from the existence of market failures that create a gap between the social and private benefits of R&D activities (Arrow, 1962b). That is, the objective of public organizations is to fund R&D projects that would not otherwise be implemented because of failures in the capital market or externality problems. The associated costs and uncertainty can discourage companies from investing in innovative activities. Therefore, innovation policies should

influence firms' behavior by stimulating innovation, despite its costs, primarily in relation to appropriability failures (Gök & Edler, 2012; Hall & Lerner, 2010). To this end, most countries implement support programmes to encourage private R&D. Public funding programs aim to influence learning processes and companies' innovative capabilities (Clarysse et al., 2009). The supply and distribution of funds among selected R&D projects and firms targets a variety of goals. Blanes and Busom (2004) identified some common patterns on some firm's characteristics on program participation. Previous experience in R&D is always positively associated with participation. They also found that firm's size and age are important barriers, showing that agencies only partially succeed in attracting relatively smaller and younger companies. It is therefore crucial to highlight the role of clearly identifying the public sector evaluation criteria and selection methods behind R&D projects and financing procedures.

Several articles have focused on evaluating the utility of R&D programs (Dimos & Pugh, 2016; Meyer-Krahmer & Montigny, 1989; Ormala, 1989; Roessner, 1989; Zúñiga-Vicente et al., 2014). Establishing whether the objectives of R&D programs have been met is crucial. The literature on the effects of public funding on the innovation behavior and private investment of firms is extensive. However, the conclusions vary for both developed (Marino et al., 2016) and developing countries (Berrutti & Bianchi, 2019; Crespi et al., 2016; Laplane, 2021). One explanation for these inconclusive results, besides the choice of econometric method, relates to the sector or national context and the specific design and implementation of instruments. Montmartin et al. (2018) developed a simple regional R&D investment model based on the work by Howe and McFetridge (1976). They demonstrated that the regional impact of public R&D support may vary greatly according to the local economic structure and policy design. The role a firm's sector plays in innovation is also the subject of debate in a number of articles. Some suggest that firms in the same sector have similar patterns of innovation activities (Malerba, 2002; Malerba, 2005a; Malerba, 2005b; Pavitt, 1984). However, other empirical studies have shown considerable variation within sectors. The conclusion, therefore, is that sector has a limited capacity to explain differences between companies' innovation behaviors (Coad, 2009; Leiponen & Drejer, 2007; Srholec & Verspagen, 2012).

From the perspective of firms' innovation input, scholars have assessed the effect of R&D incentives on qualified employment, tangible assets, and investments in R&D (Becker, 2015; Zúñiga-Vicente et al., 2014). Other studies have focused on the impact of subsidies on firms' innovation output. Aiello et al. (2019) suggested that public R&D subsidies succeed in stimulating private R&D expenditure, especially in small firms. However, public support does not affect the probability of patenting or the number of patents that a firm would have registered in the absence of this endorsement.

The government's criteria in selecting projects have been the focus of a number of articles (Hsu et al., 2003; Lee & Om, 1996, 1997). The literature indicates that selecting R&D projects for public evaluation is a difficult task because of the existence of multiple objectives and diverse stakeholders. Santamaría et al. (2010) proposed a model that helps understand the criteria used for public selection of cooperative R&D projects. They analyzed the factors underlying project selection by the

Spanish Government, the main goals of the funding program and the reasons why they implemented two instruments (credits and subsidies) in the same call. They found that the type of projects is very relevant in the selection process. In addition, they observed that the agency uses the two financial tools to address different objectives. There are also differences depending on the sector and in terms of the year of the call. This study complements the studies of Acosta Ballesteros and Modrego Rico (2001) and Blanes and Busom (2004), which focus on organizations' decisions to apply for public support for R&D.

Firms are interested in identifying the most important variables when participating in a selection process for cooperative R&D projects. The literature suggests that several company characteristics are related to public R&D funding. Factors such as belonging to a group, export experience, financial structure, innovation, and R&D efforts in previous years have been shown to have an impact on public funds.

Afcha (2012) identified the key determinants of R&D innovation strategies and the link with the concession of R&D subsidies. A two-step treatment effects model was used for a sample of Spanish firms for the period 1998 to 2005. Consistent with previous research, this paper focuses on the influence of technological cooperation, R&D efforts in previous years, recruitment of qualified personnel, and the percentage of foreign capital and exports as determinants of public subsidies. Cantner and Kösters (2012) used logistic regression to analyze the allocation of R&D subsidies to German start-ups. They showed that the initial capital of the firm and the work team influence the public funds received.

A number of papers have shown that firm size is a positive determinant of the probability of securing public funds. However, this finding may be interpreted as contradicting the declaration of many programs that their aim is to support SMEs. Duguet (2003) used matching methods to show that firm size has a direct effect on the probability of achieving public R&D funding, with larger firms being more likely to receive funds. Using a non-parametric matching approach, Almus and Czarnitzki (2003) found that sector, firm size, the presence abroad, and the existence of an R&D department affect the ability to secure public funds. Mardones and Zapata (2019) focused on the factors explaining the granting of public funding for innovative activities. They used both pseudo-panel and cross-sectional data, finding that smaller firms are less likely to obtain funds. García-Quevedo and Afcha Chávez (2009) found that large firms are more likely to access public funding from the central government, while SME companies are more likely to receive regional funds. Much like Blanes and Busom (2004), they concluded that the objectives of the innovation policies of regional and central administrations differ.

Several authors agree that having already received a grant can affect the possibility of receiving R&D funding. For example, Duguet (2003) found that the probability of receiving funds increases with the debt ratio, the existence and magnitude of past public support, and the ratio of private investment in R&D to sales. Antonelli and Crespi (2013) also found that having obtained funds in the past increases the probability of securing additional funding. Moreover, using parametric and non-parametric techniques, Duch-Brown et al. (2011) demonstrated that experience in R&D projects increases the amount received. This line of argument is also supported by Hussinger (2008), who used parametric and semi-parametric two-step selection

models to show that the German government chooses to finance the most promising candidates, which are usually characterized by having obtained funding in the past.

The study contributes to this debate by providing new evidence of the most relevant variables associated with the size of the received subsidy when companies participate in selection processes for R&D projects. This paper extends previous research by presenting new results using a unique and sophisticated combination of techniques, GA and RF model. Several variables, such as the value of intangible resources, are shown to influence the probability of obtaining R&D funds. By contrast, others, such as total assets, does not exert a significant effect.

## Method

This section describes the model and estimation procedure, as well as the data collection process used in this research. These areas are described separately below.

### Model/Estimation Procedure

The main contribution of the estimation procedure employed in this research is the combination of two advanced econometric techniques to achieve the aims of this study more accurately. Therefore, a metaheuristic optimization algorithm (the GA) was used to select the most important variables, and a RF model was used to rank the importance of the selected variables.

The set of potential variables that could affect the size of the subsidy received by a company can be very large. Therefore, this research is interested in the selection of the optimal set of predictors out of  $d$  possible variables such as to enhance the overall accuracy of the predictive model, lessen the computational burden, and ensure a more intuitive and interpretable model. The straightforward approach to use is Best Subset Selection, i.e., fitting all possible models. Although exhaustive, this methodology is time- and resource-consuming, not feasible for large values of  $d$ , and it might also lead to overfitted models. Consequently, an efficient algorithm to explore the subsets of predictors is needed. Particularly, this method distinguishes the parameters of a model that are the result of learning the data behavior and the hyper-parameters of the model that controls the learning process. The model hyper-parameters are related to the model selection, and in this estimation procedure, model selection is equivalent to features/variables selection.

To obtain the best subset of predictors we use a wrapper method (Guyon et al., 2010), which are useful for model selection. The method involves training the model with feature subsets. The main objective is to obtain the subset of features that best perform a specific desired prediction model, the learning machine. Wrapper methods use a search algorithm or sampling algorithm to explore the hyper-parameter space (Guyon et al., 2010), that is, the possible subsets of features. And the learning machine, the prediction model, is seen as a black box to which hyper-parameters are adjusted to reflect the data structure. In the wrapping algorithm used in the paper, the inner loop performs model selection (i.e.,

hyper-parameter selection), and the outer loop is used to estimate the generalization error. For this purpose, the data set  $X$  was split into three parts. One part was used as the training sample. The subset of data devoted to adjusting the hyper-parameters was the validation set. And the last part of the data set, the test set, was used to evaluate the final model overfitting.

To achieve a better performance of the Best Subset Selection technique, mostly concerned to the search algorithm, this study uses a heuristic method to obtain accurate results by trading optimality for computational resources such as processing time. In a heuristic method, rules are established to decide the best branch to follow at each step of the algorithm to reach the optimal value. This rule considers the accuracy of the predictors to decide which branch to follow. Some heuristics have the potential to find the global maximum  $\hat{\theta}$ , although there is no certainty that this will be the case; all that is known is that  $\hat{\theta}$  is the best solution so far. Evolutionary algorithms, particularly GAs, are suitable for this task.

The evolutionary paradigm is based upon mimicking natural selection, where a population of individuals evolves according to the maxim of the survival of the fittest. One type of evolutionary algorithm is a GA, for which candidate solutions are represented by a finite alphabet (in this case, 1 if the variable is present and 0 otherwise). The evaluation function is the core of the selection process. It assigns a measure of adequacy (fitness) to a specific solution, also quantifying improvement across generations. The next generation is populated with individuals through mating selection. The best individuals are selected to create offspring through crossover and variation operations (e.g., mutation that modifies old individuals to create new ones). The best offspring are selected for the next generation. The process is repeated until a termination condition is triggered (e.g., a maximum number of generations) or there is no substantial improvement in fitness.

As a learning machine, a RF model with mixed categorical and continuous data was used. A decision tree stratifies the predictor space into well-defined regions. RF improves the prediction accuracy further, through the use of uncorrelated trees. In such an approach, each tree is grown using a random sample of  $mtry < d$  variables, where  $mtry$  is usually set to  $d^{1/2}$ . Because this technique averages predictions across several low-biased and high-variance trees, it reduces variance while keeping bias low. In the model, a measure indicating importance was assigned to predictors so that they could be ranked.

The procedure was implemented in R, with the use of the CARET (Classification And REgression Training) library (Kuhn, 2019). The function GAIFS (Genetic Algorithm Feature Selection) was chosen to use a GA as the search algorithm with a wrapper method and cross-validation. The learning machine was the RF model, and the internal measure of performance to guide the search was the out-of-bag root mean square error (RMSE) estimate. The external performance measure was also RMSE. 10-fold cross-validations (number) repeated 10 times (repeat) were employed. For the GA, the roulette wheel was used as the selection operator, uniform crossover, and uniform random mutation. It started with a population of 100 individuals. A stopping criterion of 250 generations of the GA (iterations) without improvement in the best fitness value was established. The optimal value was found to be 10% for the mutation probability and 3 for the elitism.

We further analyzed two hyper-parameters associated with the RF: the number of sample predictors,  $mtry$ , and the number of trees. The latter should be large enough to ensure that the ensemble is stable. Moreover, the model should not become overfitted as more trees are added. Instead, an upper bound should be reached for the generalization error (Breiman, 2001). The  $mtry$  parameter was found with the aid of the grid search R algorithm in CARET. The possible values ranged from 1 to  $d = 13$ , which was the number of available variables.

## Data collection and Sample Characteristics

The sample comprised 807 companies and 1252 cases of R&D&i subsidies over a five-year period (2014–2018). The combined value of these subsidies totalled almost 25 million euros. Empirical data on innovation subsidies granted within the Spanish region of Valencia (*Comunidad Valenciana*) were analyzed. More specifically, the study focuses on the funds administered by the Valencian Institute for Business Competitiveness (*Instituto Valenciano de Competitividad Empresarial – IVACE*). Given its R&D&i budget, this institution acts as the main reference for innovation in this region. The study period was characterized by economic growth and a strong recovery after the global economic crisis of 2008 to 2012.

First, all the companies that received public R&D grants from the IVACE ([www.ivace.es](http://www.ivace.es)) between 2014 and 2018 were identified by consulting the folder marked *Innovación e I+D* (Innovation and R&D) found under the heading *Descargas* (Downloads). During this period, companies benefited from several lines of support. Once the full list of beneficiaries had been compiled, their details were consulted in the SABI financial database (*Sistema de Análisis de Balances Ibéricos*; <https://sabi.bvdinfo.com>). Numerous variables were collected for each company and each year.  $Assets_{it}$  was the total assets of firm  $i$  in year  $t$ .  $Pp\&e_{it}$  was the value of property, plant and equipment or the long-term physical/tangible assets of company  $i$  in year  $t$ .  $Intang_{it}$  was the monetary value of intangible resources owned by company  $i$  in year  $t$ . Two more specific variables were included to capture the intangible assets of the company:  $Brands_{it}$ , or the total number of registered brands owned by company  $i$  in year  $t$ , and  $Pats_{it}$ , or the total number of registered patents owned by company  $i$  in year  $t$ .  $AggVal_{it}$  indicated the value added of company  $i$  in year  $t$ .  $Totequ_{it}$  was the value of assets less liabilities of company  $i$  in year  $t$ .  $Netinc_{it}$  was the income less expenses, deductions, and taxes of company  $i$  in year  $t$ .  $Empl_{it}$  corresponded to the employee expenses of company  $i$  in year  $t$ .  $Times_{it}$  reflected the successful experience of company  $i$  when applying for innovation subsidies, operationalized as the number of times during the research period that company  $i$  had received a subsidy prior to year  $t$ .  $Internat_{it}$  captured whether company  $i$  was operating outside its home country borders.  $Geogr_{it}$  indicated the geographical location (province) of company  $i$  in year  $t$ . Finally,  $Activity_{it}$  was the general industrial classification of company  $i$  in year  $t$  (0 for service companies and 1 otherwise). These 13 features were selected as explanatory variables. The economic value of the public R&D subsidies that companies received from IVACE was used as the objective variable in the RF model consisting of 500 decision trees.



## Results

Table 1 shows descriptive statistics for research variables. Inspection of the subsidized projects indicates that the average funding received was 19,579 euros. Each company received an average of 1.8 R&D grants. In addition, each company had an average of 1.4 patents and 2.9 brands. Of these projects, 64.3% belonged to companies located in Valencia, 23.8% to companies in Alicante, 10.1% to companies in Castellón, and 1.8% to companies whose main headquarters were outside the region of Valencia.

An advanced method was applied to the above company and innovation subsidy data to generate new knowledge. The optimal explanatory variables were selected using the GA technique and allowed the RF to learn. The learning results were visualized with variable importance plots and partial dependency plots. These results were used to rank the impact or importance of each of the selected variables on the value of innovation subsidies received by each company.

As presented in Fig. 1, internal and external performance metrics were optimized. The fitness function (internal performance metrics) of the GA took combinations of variables in the model and determined how well they performed. The external performance metrics were used to determine the optimal number of iterations in the GA algorithm to prevent overfitting. The number of both the internal and external performance metrics of the estimation decreased after the first few generations.

The GA eliminated the unnecessary variables to reduce variance and improve accuracy. The results of the GA estimation led to the elimination of the following variables: *Brands*, *Assests*, *AggVal*, and *Pats*. The remaining variables were all selected to contribute to a better understanding of the model. The fact that these four variables were found not to be suitable as explanatory variables in the model constitutes the first finding. A company's value added ( $AggVal_{it}$ ) and total assets ( $Assets_{it}$ ) were also found to be unnecessary in the algorithm. These two variables were included as proxies for company size and the value added of the production process, respectively.

A RF model is employed to rank the importance of variables in classification problems. Node impurity is one of the foremost metrics of variable importance. It is defined as the total decrease in node impurities after splitting on the variable, averaged over all trees. This metric was used as the measure of importance. This metric shows how much the model accuracy would decrease if that variable were removed. Therefore, the higher the value of the mean decrease in the Gini score, the greater the importance of the variables in the model. Specifically, the importance of all variables selected in this research by the GA to highlight the factors that determine the value of innovation subsidies received was reported. Figure 2 reports the main results of the RF estimation. Across all the trees considered in the estimation, the variable indicating the value of intangible resources ( $Intang_{it}$ ) was the most important.

After intangible resources, the features with the next largest mean decrease in the Gini index (and, consequently, the key variables to explain success in securing

**Table 1** Mean and standard deviation (in parentheses) of the main variables

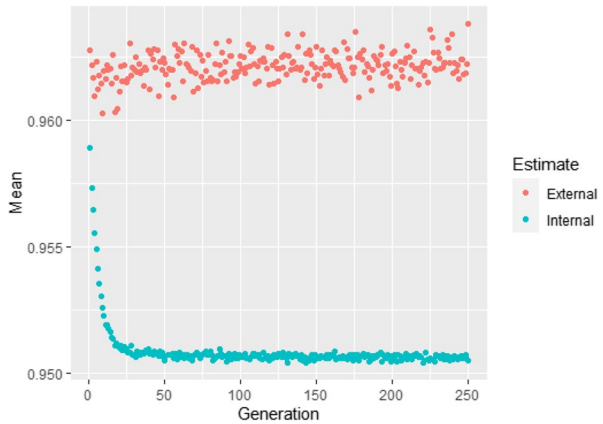
Innovation funds received	Times	Pats	Brands	Intang	PPE	Totequ	Assets	AggVal	Empl	Netinc
19,579 (21,856)	1.8 (1.3)	1.4 (4.3)	2.9 (6.7)	1191 (8881)	5630 (27,401)	6999 (23,307)	14,546 (49,061)	4179 (14,723)	2518 (9726)	832 (3120)

subsidies) are employee expenses ( $Empl_{it}$ ), income less expenses, deductions, and taxes ( $Netinc_{it}$ ), the value of tangible assets ( $Pp\&e_{it}$ ), and total equity ( $Totequ_{it}$ ). Based on the results of the final RF model, the sixth most important variable is successful experience in securing subsidies ( $Times_{it}$ ). The RF results show that international business activity ( $Internat_{it}$ ), geographical location ( $Geogr_{it}$ ), and industrial classification ( $Activity_{it}$ ) are significantly less important.

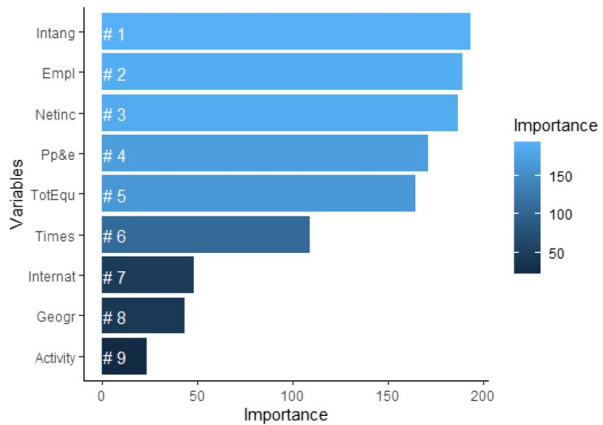
Figure 3 shows the partial dependence plots to check the marginal effect of company characteristics on the economic value of innovation subsidies. As displayed in Fig. 3, the degree of partial dependence on the economic value of innovation subsidies is high for some of the variables. The variables showing the value of intangible resources, the employee costs, and the value of property, plant, and equipment are related to the value of the subsidy to a similar extent and in a similar way. The relationships between these variables and the amount received in subsidies are highly positive for low values of these dependent variables and then remain constant for higher values. There is also a positive relationship with total equity, although this relationship has a slightly different form. Figure 3 reveals that the positive experience of companies requesting innovation subsidies,  $Times_{it}$ , is also important. The degree of partial dependence on the value of the subsidy grows as the number of times that a company has received such subsidies increases. However, the relationship for the variable  $Netinc_{it}$  is negative. Although there are different types of subsidies with different objectives and priorities, it seems that companies with the lowest net income receive larger innovation subsidies.

The last three partial dependence plots refer to categorical variables. Positive values mean that, according to the model, receiving a larger innovation subsidy is more likely for that value of the categorical variable. The values in the y-axis are low for these variables, indicating a low average impact on the amount received in innovation subsidies, as also shown in Fig. 3. Service companies (with a value of 0 in the activity variable) seem to have a stronger influence than industrial companies on the amount received in innovation subsidies. Additionally, the strength of the impact on the amount received in subsidies is highest for companies based outside the region of Valencia, followed by those in Castellón, Valencia, and Alicante, in that order. Finally, regarding the categorical variable  $Internat_{it}$ , companies with no international activity have the highest likelihood of receiving subsidies, followed by importing companies. When this variable takes the category “Export” (i.e., exporting companies) or “Imp/Exp” (companies that engage in both exports and imports), the value is negative. Hence, according to the model, a higher subsidy is less likely for this value of the variable, which

**Fig. 1** Mean errors of the genetic algorithm



**Fig. 2** The RF model with selected variables in descending order of variable importance based on the mean decrease in Gini

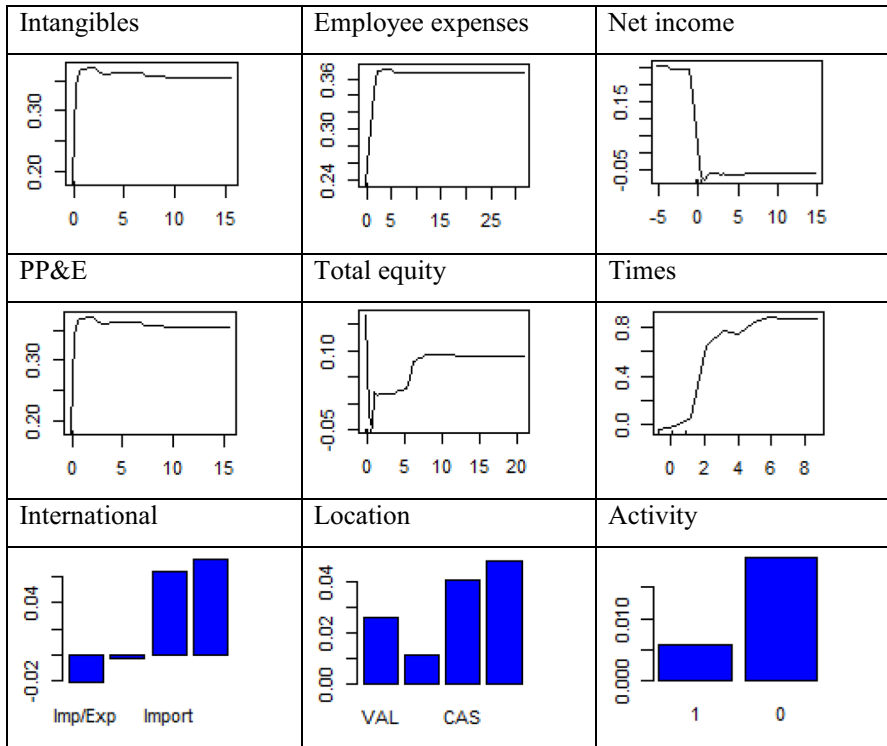


indicates international activity. These results highlight the features that determine the amount received in innovation subsidies by the companies in the sample.

**Discussion**

This paper analyzes the main characteristics that help companies to obtain higher values of public funding for R&D and innovation subsidies. In addition, these characteristics are ranked by importance in this process. The pioneering empirical approach applied in this study entails a combination of an algorithm, GA, and a machine learning process, RF model. By using this approach, this paper makes an important contribution to identify the key drivers for companies participating in R&D&i project selection processes.

The results of the Genetic Algorithm (GA) estimation eliminate *Brands*, *Assests*, *AggVal*, and *Pats*, as they were considered unnecessary variables. The fact that these four variables were found not to be suitable as explanatory variables in the model



**Fig. 3** Partial dependence plots of company characteristics and innovation subsidies

constitutes a significant finding in line with some of the previous literature. Brands and patents are probably among the intangible variables that are most closely related to the outcome of innovation. Therefore, they could be interpreted as a consequence of innovation activities. Although it has been shown that investments in R&D are directly related to public incentives (Becker, 2015; Zúñiga-Vicente et al., 2014), Aiello et al. (2019) showed that public support does not affect the probability of patenting or the number of patents. Following this argument, although some of the public tenders were aimed at subsidising the registration or purchase of patents (e.g. IVACE, 2018), the first conclusion from the analysis is that the number of brands and patents does not significantly affect the amount a company receives in innovation subsidies. Furthermore, this lack of a relationship makes sense in that the patent rates in Spain are below the European average, even though they are growing (EUIPO, 2019).

A company's value added ( $AggVal_{it}$ ) and total assets ( $Assets_{it}$ ) were included as proxies for company size and the value added of the production process, respectively. These two variables were also found to be unnecessary in the algorithm. Therefore, the amount received by firms in innovation subsidies seems to be independent of these two variables. The literature is inconclusive regarding the role

of firm size in determining the probability of receiving public funds. While some authors have reported that larger firms are more likely to secure funds (Almus & Czarnitzki, 2003; Duguet, 2003; Hussinger, 2008; Mardones & Zapata, 2019), others such as Busom (2000) have concluded that subsidies tend to be channelled to smaller firms. In the analysis, most of the grants were awarded to small- and medium-sized enterprises. Therefore, it is difficult to identify significant differences between such firms. In conclusion, size does not seem to influence the amount of innovation funds received by companies.

The remaining variables were all selected to contribute to a better understanding of the model. In this regard, the variable indicating the value of intangible resources ( $Intang_{it}$ ) was the most important across all the trees considered in the RF estimation. Costs and high levels of uncertainty can discourage companies from investing in innovative activities through R&D spending. Some papers (Gök & Edler, 2012; Hall & Lerner, 2010) have shown that, despite their costs, public funding programs stimulate innovation. However, innovation policies should be tailored and region-specific in order to be effective in supporting regional systems to transform their inputs endowment into innovation output (Ganau, 2021). Following this reasoning, the increase in R&D spending and the corresponding increase in the value of intangible resources would be a consequence of public funding. However, the results in this paper suggest that innovation is actually a determinant of receiving public support. Given that almost 70% of innovative activities in Spain are financed with equity capital, intangible resources are particularly relevant for securing public funds (Cotec, 2020). Consistent with the results of Afcha (2012) and Duch-Brown et al. (2011), the model shows that innovation and R&D spending have a significant positive effect on the amount received in innovation subsidies. Firms with more intangibles have more experience in R&D. Consequently, they also have a greater capacity for innovation, which positively influences the possibility of receiving public subsidies.

After intangible resources, the key variables to explain success in securing subsidies are employee expenses ( $Empl_{it}$ ), income less expenses, deductions, and taxes ( $Netinc_{it}$ ), the value of tangible assets ( $Pp\&e_{it}$ ), and total equity ( $Totequ_{it}$ ). Employee expenses ( $Empl_{it}$ ) have a direct, positive impact on the amount companies receive in subsidies. This variable is linked to the number of employees and their qualifications. The processes of developing R&D activities and incorporating external knowledge to meet companies' internal requirements need qualified human resources. Afcha (2012) found that the recruitment of qualified personnel is a determinant of receiving public subsidies. This paper also shows that the number of employees has a direct impact on the probability of receiving a public grant. The combination of these two arguments justifies the result.

$Netinc_{it}$  (income less expenses, deductions, and taxes) is the third most important variable in determining the amount that companies receive in subsidies. Financial solvency has been identified (Hussinger, 2008) as a determinant of governments' choice of candidates. The most promising candidates and the most profitable firms are preferred. These variables are especially relevant in relation to subsidies for the creation of technology-based companies. In such cases, financial viability to ensure the resources and profitability of the project is evaluated in detail. Following this

argument, the value of tangible assets ( $Pp\&e_{it}$ ) and the total equity ( $Totequ_{it}$ ) were also found to be relevant indicators of the amount that companies receive in subsidies. Both variables are directly related to solvency, which again suggests that stronger companies are more likely to receive public funds.

Successful experience in securing subsidies ( $Times_{it}$ ) is the sixth most important variable based on the results of the final RF model. Having already received a grant is cited as a determinant of the probability of receiving another grant in several articles (Antonelli & Crespi, 2013; Duch-Brown et al., 2011; Duguet, 2003). In this case, although it seems to exert some influence, it is not the most important factor in governments' choice of candidates, probably because most R&D projects are undertaken using the company's equity. Therefore, what public organizations value are firms' capabilities.

International business activity ( $Internat_{it}$ ), geographical location ( $Geogr_{it}$ ), and industrial classification ( $Activity_{it}$ ) turn out to be significantly less important in explaining success in securing subsidies according to the RF results. The literature (Afcha, 2012; Góngora-Biachi et al., 2009) cites international business activity as a factor related to public financing. Firms that receive funding often seem to be oriented to international markets. Therefore, this result is surprising given that, in some tenders, companies are asked to provide an internationalization plan (IVACE, 2019) or explain whether their actions will have an international impact. A possible explanation for this finding is that the variable  $Internat_{it}$  is less important in the sample because Valencia is the region with the second highest exports in Spain (EFE, 2020).

The role of geographical location ( $Geogr_{it}$ ) and industrial classification ( $Activity_{it}$ ) raises a number of arguments in the broad debate on how to identify the usefulness of R&D programs (Dimos & Pugh, 2016; Meyer-Krahmer & Montigny, 1989; Zúñiga-Vicente et al., 2014). Some streams of research have reported that firms in the same sector have similar patterns of innovation, whereas other empirical studies have shown considerable variation within sectors. Regarding the amount received in subsidies, Santamaría et al. (2010) concluded that, in Spain, some industries (e.g. aeronautics) have received high levels of support, while others (e.g., automobiles) have received less support. They even found differences between the financial tools that are employed in different sectors. More traditional sectors receive high support through subsidies, whereas new technologies are selectively supported using credit. The model indicates that geographical location and industrial classification have a minor role in determining funding when compared to the above-mentioned variables. Again, this result is somewhat expected given that the representation of the sectors is fragmented.

According to the analysis of the partial dependence plots, the variables related to the value of intangible resources, the employee costs, and the value of property, plant, and equipment have a deep impact on the amount received in subsidies. The positive experience of companies requesting innovation subsidies is also important. Hence, the more times a company has received a subsidy, the higher the value of the innovation subsidy will be. In contrast, the relationship for the variable  $Netinc_{it}$  is negative. This is an interesting finding suggesting that companies are more likely to receive a subsidy when they have a low net income. This result may be a reflection of attempts to support struggling companies that are unable to afford

investment in innovation, suggesting that development and its contribution to society are considered more important than the company's profits. The negative sign for the variable  $Netinc_{it}$  stems from the fact that the more a company innovates, the worse its short-term performance will be because R&D&i requires huge economic investment, and the maturity period for innovation is long. Other partial dependence plots refer to international activity, location and sector of activity. Service companies seem to have a stronger influence than industrial companies on the amount received in innovation subsidies. Additionally, the amount received in subsidies is highest for companies based outside the region of Valencia. Finally, companies with no international activity have the highest likelihood of receiving subsidies, followed by importing companies. Although this result is the opposite of that reported by Bravo-Ortega et al. (2014), it is consistent with those obtained for the net income variable. That is, the amount received in subsidies is considered as a whole, which includes different types of grants with different objectives. The results suggest that most of the subsidies in the present sample target companies with a predominantly regional scope.

### Implications of the Study

The fundamental role of innovation in today's economies highlights the importance of R&D&i, which is a key factor for the creation of high-productivity jobs that enable society to meet its challenges head on (Mei et al., 2020; Wanzenböck & Frenken, 2020). Consequently, public administrations develop and implement public policies to promote R&D, prevent market failures and solve possible problems of appropriability (Leiponen & Byma, 2009). Its importance is of such magnitude that the amounts of funds that most of the advanced countries invest in this concept has been growing significantly during the last few decades. However, despite the importance of these policies, very few studies have examined the factors that help companies access public funds allocated by public administrations. By focusing on a variety of firm-level characteristics, this paper theoretically and empirically extends the understanding of firms' applications for public innovation subsidies.

The stimulation of innovation through public support is a very desirable opportunity for companies' managers to opt for public financial assistance for R&D and innovation. This study makes an important contribution given its usefulness for public innovation policies and for comparing the aspects that are considered important and those that are prioritized. This insight could be helpful to define the profile of target companies, differentiating between regional and national policies, given the differences reported in the literature between the companies that are targeted by each type of policy (García-Quevedo & Afcha Chávez, 2009). Similarly, this study offers a useful guide for other Spanish regions that assign R&D grants based on a first come, first served basis instead of a public tender system.

The results of this paper can help policymakers to further increase their transparency levels regarding the criteria and guidelines that enable companies to access these public funding incentives for innovation. In this regard, the 'Ley de Transparencia, Acceso a la Información Pública y Buen Gobierno' [Law of Transparency,

Access to Public Information and Good Governance] (Ley 19/2013), and the ‘Ley General de Subvenciones’ [General Law of Subsidies] (Ley 38/2003) stipulate that the managing bodies of public funds must publish the name of the beneficiary, the title of the project, and the amount received. Therefore, knowing what criteria determine whether a company can access R&D subsidies would enhance the transparency of public R&D policies (Grimmelikhuijsen & Welch, 2012; Heald, 2006). This aspect is important given that the social impact of public spending aimed at R&D is not currently measured or monitored (Moreno-Izquierdo & Pedreño-Muñoz, 2020). Tomás-Lopes and Marques-Serrasqueiro (2017) have shown a positive correlation between transparency and innovation, indicating that there is a relationship between a country’s level of transparency and R&D spending.

The identification of the main drivers that can lead a company to obtain a public innovation subsidy generates useful information for both company and political managers. Public administrations must optimize the allocation of scarce public resources for innovation in terms of their distribution among the companies considered as optimal candidates. That is, firms which are more likely to achieve profitable outputs in the form of benefits for the company and society. From a company’s point of view, this information provides the guidelines to successfully obtain public financial support for innovation.

### **Limitations of the Study**

This paper has several limitations that could be addressed in future research. The objective of this paper is to determine the relationship between the value of received subsidies and some characteristics of the successful companies receiving public R&D&i funds. Therefore, the study incorporated aggregate data of all R&D&i subsidies granted by IVACE over five years payments. Research is already in progress to clarify whether there are significant differences between different types of subsidies, including those focused on innovation, those focused on R&D and those that support the business models of technology-based companies. This research can be also enhanced by including other companies which could not access to public R&D&i funds. Finally, other lines of research are planned to address in the future relate to the analysis of differences (using the variables studied in this research) between national and regional subsidies and between national and European grants. Further research is required to investigate and to compare criteria in different Spanish regions.

### **Conclusions**

This study adds some evidence to the identification of some significant drivers for companies participating in R&D&i project selection processes and gaining more insights on how public subsidies are allocated among companies. The empirical findings of this research conclude that intangible resources are the most important variables to explain success in obtaining subsidies. Intangible resources are followed



in descending order of importance by the following variables: Employee expenses, net income, tangible assets, total equity, previous successful experience in securing subsidies, international business activity, geographical location, and industrial classification. These results are highly relevant for both companies' managers and policymakers.

Specifically, this research contributes to the innovation subsidy literature in three ways. First, by focusing on a variety of firm-level characteristics, it theoretically and empirically extends the current understanding of the participation of firms in receiving funds from innovation subsidies. The results show that the key variables in determining the size of public subsidies are the value of intangible resources, employee spending, and net income. Unsurprisingly, innovation and R&D investment, which are often accompanied by an increase in human capital, have a significant positive effect on the amount received in innovation grants. Net income is a key determinant of public subsidies, although the direction of the relationship is closely linked to the aim of the grant. A similar phenomenon is observed for international presence. However, this variable has a low significance. Other factors that influence the amount received in public subsidies are tangible assets and equity. These findings imply that, when granting subsidies, IVACE fundamentally values the internal material and human resources and research experience of companies rather than external factors such as geographical location or sector. Finally, although having already received a grant is a determinant of the probability of receiving further grants, governments hardly consider this factor when choosing between candidates.

Second, it uses a sample of companies based in Valencia, Spain (*Valencian Community*), for a 5-year period. Regional systems attributed higher levels of effectiveness. Innovation subsidy studies using samples of companies from small regions are scarce, especially those addressing the determinants of local companies' success in accessing R&D&i policies. According to some estimations from Eurostat (the statistical office of the European Union), Valencian Community is a "moderate innovator +," above the Spanish average and at the same level that other regions like Catalonia. Innovation subsidy studies using samples of companies located in small regions are scarce, especially those examining the key determinants for local companies that successfully benefit from R&D policies. Regional systems are likely to perform better. They are attributed greater transparency, higher efficiency and therefore, higher levels of effectiveness (Holl, 2021; Rodríguez-Pose, 2013; Rodríguez-Pose & Storper, 2006). Regional policies are needed to contribute to the emergence of optimal conditions for regional systems to increase their innovation productivity. Therefore, this study broadens scholars' understanding of the criteria employed by regional administrations to grant public funding.

Third, the analysis method involves a sophisticated combination of techniques. A metaheuristic optimization algorithm (GA) was used to select the key variables. This stochastic method for function optimization is inspired by the processes of natural evolution in a population of individuals, using this idea to produce better and better approximations. To provide an accurate ranking of the importance of the selected variables when companies apply for innovation subsidies, a machine learning process was applied. More specifically, a RF model was used.

Overall, the contribution of innovation to Spanish GDP is still far from the target set by the Spanish Government and that of many other European countries. An in-depth knowledge of the variables that determine access to public subsidies is essential to strengthen R&D systems and encourage companies to continue to invest efficiently in innovation. The research findings can be said to represent a positive attempt to provide useful suggestions to companies acceding to innovation subsidies and to policymakers in the field of innovation policy effectiveness. In this sense, this study provides relevant information to help increase companies' productivity and, consequently, the productivity of Spain as a whole.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

**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

- Acharya, V., & Xu, Z. (2017). Financial dependence and innovation: The case of public versus private firms. *Journal of Financial Economics*, 124(2), 223–243. <https://doi.org/10.1016/j.jfineco.2016.02.010>
- Acosta Ballesteros, J., & Modrego Rico, A. (2001). Public financing of cooperative R&D projects in Spain: The Concerted Projects under the National R&D Plan. *Research Policy*, 30(4), 625–641. [https://doi.org/10.1016/S0048-7333\(00\)00096-2](https://doi.org/10.1016/S0048-7333(00)00096-2)
- Afcha, S. (2012). Analyzing the interaction between R&D subsidies and firm's innovation strategy. *Journal of Technology Management & Innovation*, 7(3), 57–70. <https://doi.org/10.4067/S0718-27242012000300006>
- Aiello, F., Albanese, G., & Piselli, P. (2019). Good value for public money? The case of R&D policy. *Journal of Policy Modeling*, 41(6), 1057–1076. <https://doi.org/10.1016/j.jpolmod.2019.02.006>
- Almus, M., & Czarnitzki, D. (2003). The effects of public R&D subsidies on firms' innovation activities. *Journal of Business & Economic Statistics*, 21(2), 226–236. <https://doi.org/10.1198/07350103288618918>
- Álvarez-Ayuso, I. C., Kao, C., & Romero-Jordán, D. (2018). Long run effect of public grants and tax credits on R&D investment: A non-stationary panel data approach. *Economic Modelling*, 75, 93–104. <https://doi.org/10.1016/j.econmod.2018.06.011>
- Antonelli, C., & Crespi, F. (2013). The “Mathew effect” in R&D public subsidies: The Italian evidence. *Technological Forecasting and Social Change*, 80(8), 1523–1534. <https://doi.org/10.1016/j.techfore.2013.03.008>
- Arrow, K. J. (1962a). The economic implications of learning by doing. *Review of Economic Studies*, 29(3), 155–173.
- Arrow, K. J. (1962b). Economic welfare and the allocation of resources for invention. In R. Nelson (Ed.), *The Rate and Direction of Inventive Activity* (pp. 609–626). Princeton University Press.
- Becker, B. (2015). Public R&D policies and private R&D investment: A survey of the empirical evidence. *Journal of Economic Surveys*, 29(5), 917–942. <https://doi.org/10.1111/joes.12074>

- Berrutti, F., & Bianchi, C. (2019). Effects of public funding on firm innovation: Transforming or reinforcing a weak innovation pattern? *Economics of Innovation and New Technology*, 29(5), 522–539. <https://doi.org/10.1080/10438599.2019.1636452>
- Blanes, J. V., & Busom, I. (2004). Who participates in R&D subsidy programs? The case of Spanish manufacturing firms. *Research Policy*, 33(10), 1459–1476. <https://doi.org/10.1016/j.respol.2004.07.006>
- Bloom, N., Griffith, R., & Van Reenen, J. (2002). Do R&D tax credits work? Evidence from a panel of countries 1979–1997. *Journal of Public Economics*, 85(1), 1–31. [https://doi.org/10.1016/S0047-2727\(01\)00086-X](https://doi.org/10.1016/S0047-2727(01)00086-X)
- Bravo-Ortega, C., Benavente, J. M., & González, A. (2014). Innovation, exports and productivity: Learning and self-selection in Chile. *Emerging Markets Finance and Trade*, 50(1), 68–95. <https://doi.org/10.2753/REE1540-496X5001S105>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Busom, I. (2000). An empirical evaluation of the effects of R&D subsidies. *Economics of Innovation and New Technology*, 9(2), 111–148. <https://doi.org/10.1080/10438590000000006>
- Cantner, U., & Kösters, S. (2012). Picking the winner? Empirical evidence on the targeting of R&D subsidies to start-ups. *Small Business Economics*, 39, 921–936. <https://doi.org/10.1007/s11187-011-9340-9>
- Casanova, L., Klaus Cornelius, P., & Dutta, S. (2018). *Financing entrepreneurship and innovation in emerging markets*. Elsevier. <https://doi.org/10.1016/B978-0-12-804025-6.00002-2>
- Chalioi, E., & Serfes, K. (2017). Strategic incentives for innovations and market competition. *International Journal of Industrial Organization*, 52, 427–449. <https://doi.org/10.1016/j.ijindorg.2017.03.004>
- Chapman, G., & Hewitt-Dundas, N. (2018). The effect of public support on senior manager attitudes to innovation. *Technovation*, 69, 28–39. <https://doi.org/10.1016/j.technovation.2017.10.004>
- Clarysse, B., Wright, M., & Mustar, P. (2009). Behavioural additionality of R&S subsidies: A learning perspective. *Research Policy*, 38(10), 1517–1533. <https://doi.org/10.1016/j.respol.2009.09.003>
- Coad, A. (2009). *The growth of firms: A survey of theories and empirical evidence*. Edward Elgar Publishing Limited.
- Cotec (2020). *Informe Cotec 2020*. <https://cotec.es/presentacion-virtual-del-informe-cotec-2020/>. Accessed June and July 2020.
- Crespi, G., Giuliodori, D., Giuliodori, R., & Rodríguez, A. (2016). The effectiveness of tax incentives R&D+i in developing countries: the case of Argentina. *Research Policy*, 45(10), 2023–2035. <https://doi.org/10.1016/j.respol.2016.07.006>
- David, P. A., Hall, B. H., & Toole, A. A. (2000). Is public R&D a complement or substitute for private R&D? A review of the econometric evidence. *Research Policy*, 29(4–5), 497–529. [https://doi.org/10.1016/S0048-7333\(99\)00087-6](https://doi.org/10.1016/S0048-7333(99)00087-6)
- De Jorge, J., & Suárez, C. (2011). Influence of R&D subsidies on efficiency: the case of Spanish manufacturing firms. *Cuadernos de Economía y Dirección de la Empresa*, 14(3), 185–193. <https://doi.org/10.1016/j.cede.2010.11.001>
- Dimos, C., & Pugh, G. (2016). The effectiveness of R&D subsidies: A meta-regression analysis of the evaluation literature. *Research Policy*, 45(4), 797–815. <https://doi.org/10.1016/j.respol.2016.01.002>
- Dosi, G. (1988). Sources, procedures, and microeconomic effects of innovations. *Journal of Economic Literature*, 26(3), 1120–1171.
- Duch-Brown, N., García-Quevedo, J., & Montolio, D. (2011). The link between public support and private R&D effort: What is the optimal subsidy? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1864192>
- Duguet, E. (2003). *Are R&D subsidies a substitute or a complement to privately funded R&D? Evidence from France using propensity score methods for non-experimental data* (I Cahier de la MSE EUREQua Working Paper No. 2003.75). <https://doi.org/10.2139/ssrn.421920>
- EFE (2020). La Comunidad Valenciana, entre los territorios de mayor internacionalización de sus empresas [Valencian Community, among the territories with the greatest internationalization of their companies]. El periódico de aquí. <https://www.elperiodicodeaqui.com/epda-noticias/la-comunitat-valenciana-entre-los-territorios-de-mayor-internacionalizacion-de-sus-empresas/225598>. Accessed June and July 2020.
- EUIPO (2019). *IPR-intensive industries and economic performance in the European Union, Industry-Level Analysis Report*. [https://euiipo.europa.eu/tunnel-web/secure/webdav/guest/document\\_library/](https://euiipo.europa.eu/tunnel-web/secure/webdav/guest/document_library/)

- observatory/documents/IPContributionStudy/IPR-intensive\_industries\_and\_economicin\_EU/WEB\_IPR\_intensive\_Report\_2019.pdf. Accessed June and July 2020.
- Eurostat (2022). *Gross domestic expenditure on R&D*. [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=R%26D\\_expenditure&oldid=551418#Gross\\_domestic\\_expenditure\\_on\\_R.26D](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=R%26D_expenditure&oldid=551418#Gross_domestic_expenditure_on_R.26D). Accessed June and July 2020.
- Ganau, R., & Grandinetti, R. (2021). Disentangling regional innovation capability: What really matters? *Industry and Innovation*, 28(6), 749–772. <https://doi.org/10.1080/13662716.2021.1904841>
- García-Quevedo, J., & Afcha Chávez, S. (2009). Assessing the impact of public funds on private R&D: A comparative analysis between state and regional subsidies. *Investigaciones Regionales*, 15, 277–294.
- Gök, A., & Edler, J. (2012). The use of behavioural additionality evaluation in innovation policy making. *Research Evaluation*, 21(4), 306–318. <https://doi.org/10.1093/reseval/rvs015>
- Góngora-Biachi, G., Madrid Guijarro, A., Pérez, G., & de Lema, D. (2009). Public support (subsidy) for innovation: Empirical evidence regarding an industrial PYME from south-eastern México. *Innovar*, 19(34), 65–82.
- González, X., & Pazó, C. (2008). Do public subsidies stimulate private R&D spending? *Research Policy*, 37(3), 371–389. <https://doi.org/10.1016/j.respol.2007.10.009>
- Grimmelikhuijsen, S. G., & Welch, E. W. (2012). Developing and testing a theoretical framework for computer-mediated transparency of local governments. *Public Administration Review*, 72(4), 562–571. <https://doi.org/10.1111/j.1540-6210.2011.02532.x>
- Grömping, U. (2015). Variable importance in regression models. *WIREs Computational Statistics*, 7(2), 137–152. <https://doi.org/10.1002/wics.1346>
- Guellec, D., & Van Pottelsberghe De La Potterie, B. (2003). The impact of public R&D expenditure on business R&D. *Economics of Innovation and New Technology*, 12(3), 225–243. <https://doi.org/10.1080/10438590290004555>
- Guisado-González, M., Ferro-Soto, C., & Guisado-Tato, M. (2016). Assessing the influence of differentiation strategy and R&D subsidies on R&D cooperation. *Technology Analysis & Strategic Management*, 28(7), 857–868. <https://doi.org/10.1080/09537325.2016.1180352>
- Guyon, I., Saffari, A., Dror, G., & Cawley, G. (2010). Model selection: Beyond the Bayesian/Frequentist divide. *Journal of Machine Learning Research*, 11, 61–87.
- Hall, B. H., & Lerner, J. (2010). The financing of R&D and innovation. In B. H. Hall & N. Rosenberg (Eds.), *Handbook of the Economics of Innovation* (pp. 609–639). Elsevier. [https://doi.org/10.1016/S0169-7218\(10\)01014-2](https://doi.org/10.1016/S0169-7218(10)01014-2)
- Hao, Z., Qi, W., Gong, T., Chen, L., & Max Shen, Z. J. (2019). Innovation uncertainty, new product press timing and strategic consumers. *Omega*, 89, 122–135. <https://doi.org/10.1016/j.omega.2018.09.011>
- Heald, D. (2006). Varieties of Transparency. In C. Hood & D. Heald (Eds.), *Transparency: The key to better governance?* (pp. 25–43). Oxford University Press.
- Heijs, J. (2002). Justificación de la política tecnológica: un enfoque teórico [Justification of the technological policy: a theoretical approach]. *Revista Madrid, I+D*, 10. <http://www.madrimasd.org/revista/revista10/aula/aulas2.asp>. Accessed June and July 2020.
- Holl, A. (2021). The regional environment and firms' commitment to innovation: empirical evidence from Spain. *Economics of Innovation and New Technology*, 30(6), 565–584. <https://doi.org/10.1080/10438599.2020.1730032>
- Howe, J., & McFetridge, D. G. (1976). The determinants of R&D expenditures. *The Canadian Journal of Economics*, 9(1), 57–71. <https://doi.org/10.2307/134415>
- Hsu, Y. G., Tzeng, G. H., & Shyu, J. S. (2003). Fuzzy multiple criteria selection of government-sponsored frontier technology R&D projects. *R&D Management*, 33(5), 539–551. <https://doi.org/10.1111/1467-9310.00315>
- Hussinger, K. (2008). R&D and subsidies at the firm level: An application of parametric and semiparametric two-step selection models. *Journal of Applied Econometrics*, 23(6), 729–747. <https://doi.org/10.1002/jae.1016>
- IVACE Instituto Valenciano de competitividad. (2018). RESOLUCIÓN de 13 de diciembre de 2018, del presidente del Instituto Valenciano de Competitividad Empresarial por la que se convoca la concesión de subvenciones a empresas para proyectos de I+D y de innovación para el ejercicio 2019. DOCV, No. 8451. [http://www.dogv.gva.es/datos/2018/12/26/pdf/2018\\_12002.pdf](http://www.dogv.gva.es/datos/2018/12/26/pdf/2018_12002.pdf). Accessed June and July 2020.
- IVACE Instituto Valenciano de competitividad. (2019). RESOLUCIÓN de 5 de diciembre de 2019, del presidente del Instituto Valenciano de Competitividad Empresarial (IVACE), por la que se convocan subvenciones para proyectos de creación de empresas de base tecnológica (CREATEC-CV) con

- cargo al presupuesto del ejercicio 2019, con cofinanciación de la Unión Europea a través del Fondo Europeo de Desarrollo Regional. DOCV, No. 8699. [http://www.dogv.gva.es/datos/2019/12/17/pdf/2019\\_11943.pdf](http://www.dogv.gva.es/datos/2019/12/17/pdf/2019_11943.pdf). Accessed June and July 2020.
- Jugend, D., De Camargo Fiorini, P., Armellini, F., & Ferrari, A. G. (2020). Public support for innovation: A systematic review of the literature and implications for open innovation. *Technological Forecasting and Social Change*, 156, 119985. <https://doi.org/10.1016/j.techfore.2020.119985>
- Kang, K. N., & Park., H. (2012). Influence of government R&D support and inter-firm collaborations on innovation in Korean biotechnology SMEs. *Technovation*, 32(1), 68–78. <https://doi.org/10.1016/j.technovation.2011.08.004>
- Kim, K., Choi, S. O., & Lee, S. (2021). The effect of a financial support on firm innovation collaboration and output: Does policy work on the diverse nature of firm innovation? *Journal of the Knowledge Economy*, 12(2), 645–675. <https://doi.org/10.1007/s13132-020-00667-9>
- Klette, T. J., Møen, J., & Griliches, Z. (2000). Do subsidies to commercial R&D reduce market failures? *Microeconomic evaluation studies. Research policy*, 29(4-5), 471–495. [https://doi.org/10.1016/S0048-7333\(99\)00086-4](https://doi.org/10.1016/S0048-7333(99)00086-4)
- Kuhn, M. (2019). *Feature Selection using Genetic Algorithms*. <https://topepo.github.io/caret/feature-selection-using-genetic-algorithms.html>. Accessed June and July 2020.
- Laplane, A. (2021). Market co-creating and shaping through investments in innovation: A comparative analysis of two public funding programmes in Brazil. *Innovation and Development*. <https://doi.org/10.1080/2157930X.2021.1989646>
- Lee, M., & Om, K. (1996). Different factor considered in project selection at public and private R&D institutes. *Technovation*, 16(6), 271–275. [https://doi.org/10.1016/0166-4972\(96\)00006-5](https://doi.org/10.1016/0166-4972(96)00006-5)
- Lee, M., & Om, K. (1997). The concept of effectiveness in R&D project selection. *International Technology Management*, 13(5-6), 511–524. <https://doi.org/10.1504/IJTM.1997.001684>
- Leiponen, A., & Byma, J. (2009). If you cannot block, you better run: Small firms, cooperative Innovation, and appropriation strategies. *Research Policy*, 38(9), 1478–1488. <https://doi.org/10.1016/j.respol.2009.06.003>
- Leiponen, A., & Drejer, I. (2007). What exactly are technological regimes? Intra-industry heterogeneity in the organization of innovation activities. *Research Policy*, 36(8), 1221–1238. <https://doi.org/10.1016/j.respol.2007.04.008>
- Ley 19/2013. De transparencia, acceso a la información pública y buen gobierno. 10 diciembre de 2013. BOE, No. 295. <https://www.boe.es/buscar/doc.php?id=BOE-A-2013-12887>. Accessed June and July 2020.
- Ley 38/2003. General de subvenciones. 18 noviembre de 2003. BOE, No. 276. <https://www.boe.es/buscar/act.php?id=BOE-A-2003-20977>. Accessed June and July 2020.
- Malerba, F. (2002). Sectoral systems of innovation and production. *Research Policy*, 31(2), 247–264. [https://doi.org/10.1016/S0048-7333\(01\)00139-1](https://doi.org/10.1016/S0048-7333(01)00139-1)
- Malerba, F. (2005a). Sectoral systems: how and why innovation differs across sectors. In J. Fagerberg, D. C. Mowery, & R. R. Nelson (Eds.), *The Oxford Handbook of Innovation* (pp. 380–406). Oxford University Press.
- Malerba, F. (2005b). Sectoral systems of innovation: A framework for linking innovation to the knowledge base, structure and dynamics of sectors. *Economics of Innovation and New Technology*, 14(1-2), 63–82. <https://doi.org/10.1080/1043859042000228688>
- Mardones, C., & Zapata, A. (2019). Determinants of public funding for innovation in Chilean firms. *Contaduría y Administración*, 64(1), 1–16. <https://doi.org/10.22201/fca.24488410e.2018.1602>
- Marino, M., Lhuillery, S., Parrotta, P., & Sala, D. (2016). Additionality or crowding-out? An overall evaluation of public R&D subsidy on private R&D expenditure. *Research Policy*, 45(9), 1715–1730. <https://doi.org/10.1016/j.respol.2016.04.009>
- Martin, B. R. (2016). R&D policy instruments -A critical review of what we do and don't know. *Industry and Innovation*, 23(2), 157–176. <https://doi.org/10.1080/13662716.2016.1146125>
- Mei, L., Rodríguez, H., & Chen, J. (2020). Responsible Innovation in the contexts of the European Union and China: Differences, challenges and opportunities. *Global Transitions*, 2, 1–3. <https://doi.org/10.1016/j.glt.2019.11.004>
- Meyer-Krahmer, F., & Montigny, P. (1989). Evaluations of innovation programmes in selected European countries. *Research Policy*, 18(6), 313–332. [https://doi.org/10.1016/0048-7333\(89\)90020-6](https://doi.org/10.1016/0048-7333(89)90020-6)
- Montmartin, B., Herrera, M., & Massard, N. (2018). The impact of the French policy mix on business R&D: How geography matters. *Research Policy*, 47(10), 2010–2027. <https://doi.org/10.1016/j.respol.2018.07.009>

- Moreno-Izquierdo, L., & Pedreño-Muñoz, A. (2020). Europa frente a EE.UU. y China. In *Prevenir el declive en la era de la inteligencia artificial* [Europe versus the US and China. Preventing decline in the age of artificial intelligence]. KDP Publishing. <https://www.prevenireldeclive.com/#autores>
- Ormalá, E. (1989). Nordic experiences of the evaluation of technical research and development. *Research Policy*, 18(6), 333–342. [https://doi.org/10.1016/0048-7333\(89\)90021-8](https://doi.org/10.1016/0048-7333(89)90021-8)
- Pavitt, K. (1984). Sectoral patterns of technical change: Towards a taxonomy and a theory. *Research Policy*, 13(6), 343–373. [https://doi.org/10.1016/0048-7333\(84\)90018-0](https://doi.org/10.1016/0048-7333(84)90018-0)
- Rodríguez-Pose, A. (2013). Do institutions matter for regional development? *Regional Studies*, 47(7), 1034–1047. <https://doi.org/10.1080/00343404.2012.748978>
- Rodríguez-Pose, A., & Storper, M. (2006). Better rules or Stronger communities? On the social foundations of institutional change and its economic effects. *Economic Geography*, 82(1), 1–25. <https://doi.org/10.1111/j.1944-8287.2006.tb00286.x>
- Roessner, D. (1989). Evaluating government innovation programmes: Lessons from the U.S. experience. *Research Policy*, 18(6), 343–359. [https://doi.org/10.1016/0048-7333\(89\)90022-X](https://doi.org/10.1016/0048-7333(89)90022-X)
- Santamaría, L., Barge-Gil, A., & Modrego, A. (2010). Public selection and financing of R&D cooperative projects: Credit versus subsidy funding. *Research Policy*, 39(4), 549–563. <https://doi.org/10.1016/j.respol.2010.01.011>
- Shinkle, G. A., & Suchard, J. A. (2019). Innovation in newly public firms: the influence of government grants, venture capital and private equity. *Australian Journal of Management*, 44(2), 248–281. <https://doi.org/10.1177/0312896218802611>
- Srholec, M., & Verspagen, B. (2012). The voyage of the beagle into innovation: explorations on heterogeneity, selection and sectors. *Industrial and Corporate Change*, 21(5), 1221–1253. <https://doi.org/10.1093/icc/dts026>
- Tomás-Lopes, I., & Marques-Serrasqueiro, R. (2017). The influence of culture and transparency on global research and development intensity: An overview across Europe. *Contaduría y Administración*, 62(4), 1408–1421. <https://doi.org/10.1016/j.cya.2017.06.002>
- Torregrosa-Hetland, S., Pelkonen, A., Oksanen, J., & Kander, A. (2019). The prevalence of publicly stimulated innovations—A comparison of Finland and Sweden, 1970–2013. *Research Policy*, 48(6), 1373–1384. <https://doi.org/10.1016/j.respol.2019.02.001>
- Vanino, E., Roper, S., & Becker, B. (2019). Knowledge to money: Assessing the business performance effects of publicly-funded R&D grants. *Research Policy*, 48(7), 1714–1737. <https://doi.org/10.1016/j.respol.2019.04.001>
- Wanzenböck, I., & Frenken, K. (2020). The subsidiarity principle in innovation policy for societal challenges. *Global Transitions*, 2, 51–59. <https://doi.org/10.1016/j.glt.2020.02.002>
- Zhang, J., & Guan, J. (2018). The time-varying impacts of government incentives on innovation. *Technological Forecasting Social Change*, 135, 132–144. <https://doi.org/10.1016/j.techfore.2018.04.012>
- Zúñiga-Vicente, J. A., Alonso-Borrego, C., Forcadell, F. J., & Galán, J. I. (2014). Assessing the effect of public subsidies on firm R&D investment: A survey. *Journal of Economic Surveys*, 28(1), 36–67. <https://doi.org/10.1111/j.1467-6419.2012.00738.x>

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